2012

Evolution of Grasping Behaviour in Anthropomorphic Robotic Arms with Embodied Neural Controllers

Massera, Gianluca

Evolution of Grasping Behaviour in Anthropomorphic Robotic Arms with Embodied Neural Controllers

by

GIANLUCA MASSERA

A thesis submitted to the University of Plymouth in partial fulfilment for the degree of

DOCTOR OF PHILOSOPHY

School of Computing Communication & Electronics

December 2010

§ Copyright

§ Acknowledgements

All people I have met before and during my PhD route helped me in a way or in another. Often it is the unconscious help that is more important, so I want to express all my gratitude to my colleagues, my friends, my teachers, my bosses, and my relatives and parents.

Thanks all people I knew, I know and I'll know.

§ Author's declaration and word count

At no time during the registration for the degree of Doctor of Philosophy has the author been registered for any other University award without prior agreement of the Graduate Committee. Relevant scientific seminars and conferences were regularly attended at which work was often presented; external institutions were visited for consultation purposes and several papers prepared for publication.

Publications:

E. Tuci, G. Massera, S. Nolfi (2010), **Active categorical perception of object shapes in a simulated anthropomorphic robotic arm**, *IEEE Transaction on Evolutionary Computation Journal*

E. Tuci, G. Massera, S. Nolfi (2009), **On the dynamics of active categorisation of different objects shape through tactile sensors**, *Proceedings of the 10th European Conference of Artificial Life,* ECAL 2009

E. Tuci, G. Massera, S. Nolfi (2009), **Active categorical perception in an evolved anthropomorphic robotic arm**, *IEEE International Conference on Evolutionary Computation* (CEC), special session on Evolutionary Robotics

G. Massera, A. Cangelosi, S. Nolfi (2007), **Evolution of prehension ability in an anthropomorphic neurorobotic arm**, *Frontiers in Neurorobotics*

G. Massera, A. Cangelosi, S. Nolfi (2006), **Developing a Reaching Behaviour in an simulated Anthropomorphic Robotic Arm Through an Evolutionary Technique** in L. M. Rocha et al. (eds) *Artificial Life X: Proceeding of the Tenth International Conference on the simulation and synthesis of living systems,* MIT Press

G. Massera, S. Nolfi (2006), **Evolvere reti neurali per il controllo del posizionamento di un braccio robotico**, *Atti del III Workshop Italiano di Vita Artificiale,* Roma (Italian presentation)

G. Massera, S. Nolfi, A. Cangelosi (2005), **Evolving a Simulated Robotic Arm Able to Grasp Objects** in A. Cangelosi et al. (eds) *Modelling Language, Cognition and Action: Proceeding of the Ninth Neural Computation and Psychology Workshop Progress* in Neural Processing 16, Singapore: World Scientific

G. Massera, S. Nolfi (2005), **Un Controllo Distribuito basato su Reti Neurali per il movimento di un robot esapodo**, *Atti del II Workshop Italiano di Vita Artificiale*, Roma (Italian presentation)

G. Massera (2004), **Exploiting the Physical Agent/Environment Interactions to Evolve Neural Controllers for Autonomous Robots**, *Ninth Neural Computation and Psychology Workshop NCPW9*, University of Plymouth, UK

Presentation and Conferences Attended:

- International Conference on Epigenetic Robotics 2010
- IEEE International Conference on Evolutionary Computation 2009
- ITALK European Project Meetings and Workshops 2008 - 2009 - 2010
- International Conference SAB 2006
- Summer Schools: *"Veni Vidi Veci 2006", "Non-Linear Dynamics and Robots: from Neurons to Cognition"*
- Second & Third Italian Workshop on Artificial Life 2005 - 2006
- Ninth Neural Computation and Psychology Workshop

External Contacts:

Word count of main body of thesis: 35606

Signed: _____

Date: _____

§ Abstract

Gianluca Massera — Evolution of Grasping Behaviour in Anthropomorphic Robotic Arms with Embodied Neural Controllers

The works reported in this thesis focus upon synthesising neural controllers for anthropomorphic robots that are able to manipulate objects through an automatic design process based on artificial evolution. The use of Evolutionary Robotics makes it possible to reduce the characteristics and parameters specified by the designer to a minimum, and the robot's skills evolve as it interacts with the environment. The primary objective of these experiments is to investigate whether neural controllers that are regulating the state of the motors on the basis of the current and previously experienced sensors (i.e. without relying on an inverse model) can enable the robots to solve such complex tasks. Another objective of these experiments is to investigate whether the Evolutionary Robotics approach can be successfully applied to scenarios that are significantly more complex than those to which it is typically applied (in terms of the complexity of the robot's morphology, the size of the neural controller, and the complexity of the task). The obtained results indicate that skills such as reaching, grasping, and discriminating among objects can be accomplished without the need to learn precise inverse internal models of the arm/hand structure. This would also support the hypothesis that the human central nervous system (CNS) does necessarily have internal models of the limbs (not excluding the fact that it might possess such models for other purposes), but can act by shifting the equilibrium points/cycles of the underlying musculoskeletal system. Consequently, the resulting controllers of such fundamental skills would be less complex. Thus, the learning of more complex behaviours will be easier to design because the underlying controller of the arm/hand structure is less complex. Moreover, the obtained results also show how evolved robots exploit sensory-motor coordination in order to accomplish their tasks.

§ Contents

§ List of Figures

§ List of Tables

1 Introduction

Humans use their hands in practically every moment of their lives to explore, perceive, and recognise surfaces and objects. Aside from receiving tactile information, hands are also used for physically manipulating the environment, e.g. to manipulate tools and objects, to perform actions such as grasping, manoeuvring, lifting, carrying, etc. In order to achieve a complete and complex variety of movements, it is important to consider the entire system involved, which consists of the hands and the support given by the arms, forearms and shoulders. This system allows the organism to correctly place, reach for, and manipulate objects or tools. Undoubtedly, the dexterity in humans is related to the complexity of their arm and hand structures.

The arm control and hand movements of human and non-human primates are a fascinating research topic in cognitive science, and especially so in the domain of robotics. Despite the many attempts that have been made, the controllers of robotic manipulators do not show the same dexterity as humans. Indeed, the task of modelling in detail the mechanisms underlying arm and hand movement control in humans and primates, and the task of building robots that are able to perform human-like arm/hand movements still represents an extremely challenging goal (Schaal, 2003). Moreover, despite the progress achieved in robotics through the use of traditional control methods (Gienger et al., 2008), the attempt to develop robots with the dexterity and robustness of humans remains a long-term goal.

In robotics, the design of adaptive robotic systems that are able to perform complex object-manipulation tasks is one of the most important research issues (Schaal, 2003). In cognitive science, the relationship between action control and other cognitive functions has been shown to be important in the study of cognition (Pulvermuller, 2005; Cangelosi et al., 2005). For example, various theories of language evolution have focused on the relationship between the use of the hands, tool–making, and

language evolution (Corballis, 2003).

In neuroscience, research on the cortical areas devoted to arm and hand control have revealed the existence of a group of special neurons (*mirror neurons*) that give the brain's motor-control areas an important role in our cognitive abilities. This corroborates the idea that these areas are not used only to control our limbs, but they also participate in other cognitive skills, such as understanding the actions, goals, and emotions of other people (Rizzolatti & Craighero, 2004; Gallese & Lakoff, 2005; Rizzolatti & Arbib, 1998).

Within the realm of arm control, reaching and grasping behaviours represent key abilities, since they constitute a prerequisite for any object manipulation. Despite the importance of this topic, and despite the large body of available behavioural and neuro-psychological data, as well as the vast number of studies that have been carried out, the issues of how primates and humans learn to display reaching and grasping behaviour still remain under debate, and the findings are often controversial (Schaal, 2003; Shadmehr, 2003). Similarly, while many of the critical aspects that make reaching and grasping behaviours difficult to implement have been identified, experimental research based on different methodologies does not seem to converge toward the identification of a general methodology for robotic arm design.

One of the most controversial areas of contention regards *optimal computational models* (Wolpert & Flanagan, 2003; Kawato, 2003) and *equilibrium point approaches* (Shadmehr, 2003). Optimal computation models are based on the assumption that our brain has areas devoted to encoding our limbs and the external world/objects in what are called internal models. These models are used for forecasting the results of our actions and future states of the external world. These provisions are the 'input' of areas that are devoted to planning and deciding actions.

On the other hand, the equilibrium point hypothesis, which is based upon observations of the musculoskeletal system and the spinal cord neurocircuits, states that as regards the control of our limbs, it is not necessary to assume the existence of such

internal models. Human muscles act like a set of non-linear spring-like actuators (Sandercock et al., 2003; Shadmehr & Wise, 2005a) and their passive properties define their positions of rest. When muscles are stimulated, the arm will assume a posture (which is considered to be an equilibrium point of a dynamical system composed of muscles and spinal cord neurocircuits) that depends upon muscle activation, regardless of the initial posture, and perturbations are damped and corrected by spinal cord neurocircuits that do not need to explicitly involve the central nervous system (CNS). This implies that the CNS does not need an internal model of the limbs in order to control them, but just a method that enables it to shift the equilibrium points of the underlying dynamical systems that act upon our limbs.

This thesis supports the dynamic systems approach and equilibrium point hypothesis. This thesis proposes the use of evolutionary robotics methodology to develop neural controllers that act like dynamical systems that are able to exploit the non-linear interaction between the robot and its environment. In this model, behaviours are determined by trajectories toward the equilibrium points of the systems. Robustness and adaptivity are considered to be properties of evolved dynamical systems in which the equilibrium point changes in response to changing environments or unseen situations, without an explicit need for complex internal models, and by employing very simple neural network structures (Nolfi, 2005a).

1.1 Organisation of the Content

The content of this thesis is divided into 9 chapters and 4 appendices. Chapters 1, 2, and 3 are introductory. They address anatomical issues and general aspects of the human arm and hand. They also provide an overview of problems related to reaching for and grasping objects, from the control engineering point of view. Chapter 4 concerns the methodology used in the experiments and compares it with the methodologies presented in the introductory chapters. It also discusses the state of the art with regard to related research that employs the same methodology, focusing on

the differences with the methodologies described in Chapter 3. The four experiments presented in this thesis are reported in Chapters 5, 6, 7, and 8, respectively. The experiments are presented in the order in which they were performed by the author during his research. The first experiment concerns the realisation of only a reaching behaviour with a simple robotic arm. The second experiment is the first one that addresses the act of grasping by a fully anthropomorphic robotic arm. Following the successful results achieved in this experiment, the third experiment shows how grasping behaviour is exploited by evolved agents that categorise the shapes of two objects. Finally, the fourth experiment investigates the role of linguistic instructions in evolving the ability to perform specific actions with grasped objects. This experiment reveals the facilitation of evolving such behaviour when the controller is supported by linguistic instructions with respect to a condition in which the controller does not have such support. Chapter 9 summarises and discusses the results achieved thus far. In particular, the author presents what he considers to be the contributions to knowledge offered by this thesis. The appendices present details of the models of simulated robotic arms, and a list of publications that are directly related to the results presented here by the author.

2 The Human Arm and Hand

The human arm has three main articulations: the shoulder, elbow, and the radio-ulnar joints (Shadmehr & Wise, 2005b). It has a total of five degrees of freedom (DOFs). Three DOFs are located in the shoulder joint, which act like a ball-and-socket joint that moves the upper arm (via the *humerus* bone). These DOFs allow for *pronation/supination*, *abduction/adduction*, and *flexion/extension* of the humerus (see Figure 2.1). The fourth DOF is located at the elbow, which acts like a hinge and provides *flexion/extension* of the forearm (*radius* and *ulna* bones). The fifth DOF rotates the radius and ulna bones as a single unit, providing *pronation/supination* of the hand. In robotic manipulators, the latter DOF is typically associated with the wrist. Hence, they are likely to consider the arm as a body with only four DOFs, and the wrist as a ball-and-socket joint with the last three DOFs (Bekey, 2006).

The human hand is composed of 27 bones: eight *carpal* bones that compose the wrist, five *metacarpals* in the palm, and fourteen *phalanges* that make up the digits (two in the thumb, and three in each finger) (Page, 1998). There are several joints in the hand, and these are grouped as the *carpometacarpal*, *metacarpophalangeal* (MP), *proximal interphalangeal* (PIP), and *distal interphalangeal* (DIP) joints. The carpometacarpal joints provide the wrist with two dofs that allow for *flexion/extension* and *abduction/adduction* of the palm. The four fingers (index, middle, ring, and little finger) each have a two-DOF joint in the MP (*abduction/adduction* and *flexion/extension*) and hinge joints in PIP and DIP (*flexion/extension*). The thumb does not have a DIP joint because it is composed of only two phalanges, and the MP joint has a considerable axial rotation that could be considered a third DOF, which provides *pronation/supination* of the thumb. This axial rotation allows the thumb to oppose itself to the other digits. This is not considered a true DOF, however, because its movement is highly constrained and it cannot be actively controlled by the muscles. In total, the hand has 21 DOFs: 2 metacarpal (wrist), 10 metacar-

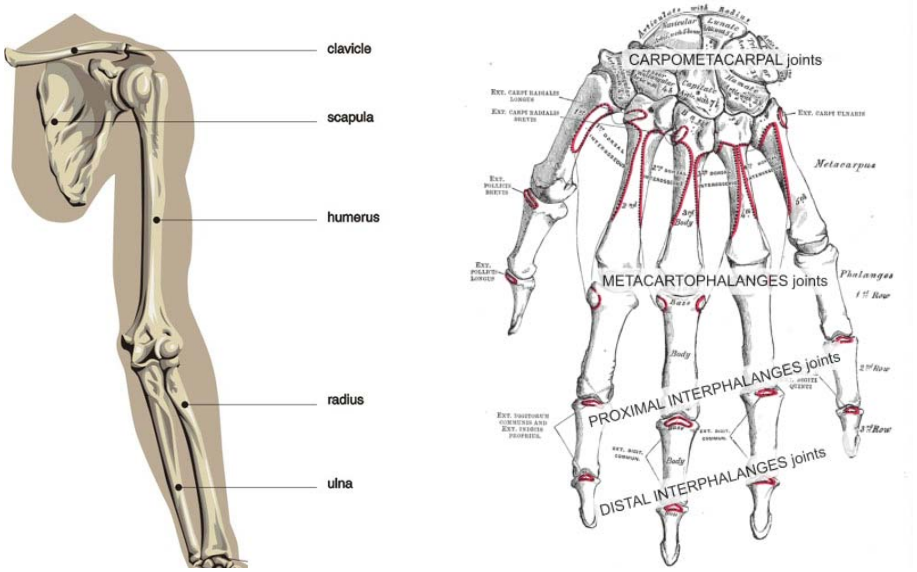


Figure 2.1: Arm and hand bones

pophalangeal (knuckles), and 9 interphalangeal (Jones & Lederman, 2006; Page, 1998).

It is evident that the human arm/hand system is highly redundant in terms of the degrees of freedom provided by the skeleton, and the number of muscles per DOF is also disproportionately large. Due to the non-linearity and redundancy of this system, during the act of reaching to grasp an object there is an infinite number of possible trajectories, arm/hand postures, and force-velocity profiles. Hence, it seems quite unlikely that any two individuals would use the same strategies to achieve the same goal. At the same time, there are strong similarities in approach, not just across individuals, but also across different primate species. These similarities suggest some common organisation of the central nervous system (CNS) among primates that controls the acts of pointing, reaching and grasping (Schaal, 2003). Some pertinent findings in this regard, which can be found in (Jones & Lederman, 2006; Shadmehr & Wise, 2005b; Arbib, 2003), are as follows.

Bell-Shaped Velocity Profile and Trajectory Curvature: When humans perform point-to-point movements, the hand path's Cartesian space (external to the subject) is approximately straight, and the tangential velocity can be rep-

resented as a symmetric bell shape. However, a more detailed examination of trajectories and velocity profiles reveals that the path of the hand makes curved trajectories, depending upon the location of the starting and ending points in the three-dimensional space in which the action takes place, and asymmetric bell shapes, depending upon the accuracy and speed required by the task.

Speed-Accuracy Trade-off: During reaching movements, the amount of time spent approaching the target has been empirically found to depend upon the target distance D from the start point, and the target width W (or equivalently, upon the accuracy required for the task):

$$MT = a + b \cdot log_2 \left(\frac{2D}{W} \right)$$

where a and b are constants. This relationship, which is called *Fitts' law*, (Meyer et al., 1990) is a robust characteristic of human arm/hand movements. Fitts' law is too generic, however, and does not express useful constraints for help in modelling and understanding the mechanism of CNS control. That said, it remains strongly descriptive of the behavioural phenomenon of primate arm movements in reaching and grasping.

Reaching to Grasp: During a reaching movement, the posture of the hand is being adjusted, showing a tight coupling and overlapping in space and time of the hand's preparation for reaching for and grasping an object (Jones & Lederman, 2006). The bell-shaped velocity profile is also valid for changes in the hand posture during the movements involved in reaching to grasp an object. In addition, it has been shown that there is a speed-accuracy trade-off concerning the time required to adjust the hand position and the precision required by the manipulation task (Jones & Lederman, 2006).

Grip Aperture: the amplitude of the maximum grip aperture is highly correlated

with the size of the object to grasp. It has been shown that the hand aperture varies between different objects even when these differences are not consciously perceived. The tactile and proprioceptive sensations of the arm play an important role on modulating the grip aperture. Organisms are able to adjust the hand posture and the reach-to-grasp movements also when the visual information remains the same, regardless of the size of the object to be grasped (Jones & Lederman, 2006).

Continuous Adaptation: Due to continuous changes in body size and properties during development, the CNS has to continuously adapt itself in exerting control over arm/hand movements. Also, the ability to learn new skills is fundamental in biological systems. Primates and human infants show an excellent ability to learn new motor skills to solve novel tasks.

3 An Outline of Reaching and Grasping Solutions in Robotics

Reaching and grasping are the two basic behaviours involved in manipulating an object. In robotics, reaching means moving the arm to create a particular configuration of the joints with which the end-effector of the robot arm is in a desired position. Typically, the end-effector is the palm or a finger of the hand. Grasping means touching an object in a such way that all forces produced at the contact points between the fingers and the object are distributed in such a way as to prevent the object from slipping from the hand.

Industrial and commercial robotic arms and hands are typically designed in such a way as to simplify the reaching and grasping behaviours (Siciliano & Khatib, 2008; Salisbury, 1982). In this regard, the design of anthropomorphic arms and hands is very far from providing the special structures needed to satisfy the principles that would simplify the reaching and grasping behaviours. The main reasons for this difficulty are:

Redundancy: The number of DOFs is redundant, and hence there is an infinite number of trajectories and final postures involved in reaching any given target point. This redundancy potentially allows anthropomorphic arms to reach a target point by circumventing obstacles or by overcoming problems caused by the limits of the DOFs. However, the redundancy of DOFs also implies that the space to be searched during learning is rather large, and a policy for choosing one of all possible postures has to be determined.

Non-Linearity: Anthropomorphic arms are highly non-linear systems. First, small variations in some of the joints could have a great impact on the end-position of the arm, while significant variations of other joints might not have any impact. Second, 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. In conjunction with this, different target positions might require similar trajectories and final postures.

Dynamics: Anthropomorphic arms are articulated structures that are suspended in space. Hence, gravity and inertia play a major role in their dynamics. In the arms of primates, the muscles and associated spinal reflex circuitry seem to confer upon the arm the ability to passively settle into a stable position (i.e. an equilibrium point) independent of its previous position. If this hypothesis were true, the contribution of the central nervous system would simply consist in the modification of the current equilibrium point (Shadmehr, 2003).

Noise and Uncertainty: Sensors and actuators can be slow and noisy. In humans, visual information and proprioceptive information that encodes changes in joint positions are available with a delay of up to $100\,ms$. It may take up to $50\,ms$ for motor commands issued by the central nervous system to initiate muscle contraction (Mial, 2003). Moreover, sensors might provide only incomplete information (e.g. the target point might be partially or totally occluded by obstacles and/or by the arm).

Furthermore, as regards grasping behaviour, the structure of the fingers and the arrangement of the robotic hand's DOFs play an important role in determining the capability of the hand, the possible finger positions, and the possible grasp configurations. In fact, the complexity of grasping behaviour comes from the contacts among the fingers and the object to be grasped.

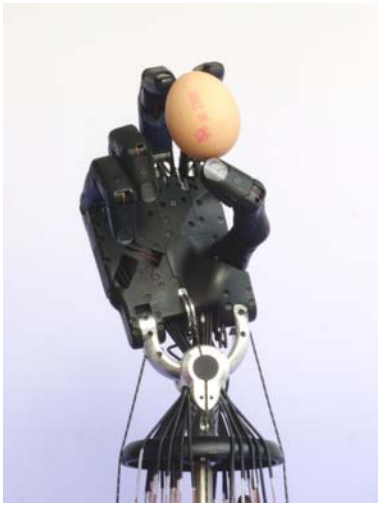
The most famous robotic hand was constructed by Salisbury (1982). The design of Salisbury's hand was not anthropomorphic, but was the result of Salisbury's proof regarding the minimum number of DOFs necessary to produce a dexterous hand. He established that only 9 DOFs are necessary, and the hand he constructed is an example of a 9-DOF hand that is dexterous (Salisbury, 1982).

In the last ten years there has been an increase in the number of designs and developments of anthropomorphic hand structures, some of which are shown in Figure 3.1. The importance of having a five-fingered hand that is similar to a human hand lies in its redundancy and compliance; however, this criterion also adds complexity to the problem. In contexts in which the controller does not have fully detailed information about the object to be grasped and there is noise and uncertainty about the object and hand displacements, the redundancy and compliance of DOFs makes it possible to damp errors more easily. With only three fingers, if one fails to touch the object properly, the grasp fails, while with five fingers, it is more likely that the object will be grasped, even if some of the fingers are incorrectly positioned. The joint's compliance allows for the adjustment of incorrectly positioned fingers by exploiting the passive effect of compliance and its feedback.

Salisbury hand (Salisbury, 1982)

Barrett hand (Townsend, 2010)

DLR hand II (Butterfass et al., 2001)



Shadow hand (Greenhill, 2010)

Figure 3.1: Robotic hands

With an anthropomorphic hand, it is possible to study all the grasp configurations displayed by humans and primates. Due to the dexterity of primates and humans, a grasp taxonomy based upon the position of the fingers is very large, and is dependent upon how the grasp is considered (Feix et al., 2009). In terms of the scope of this thesis, only *power grasps* are taken into account. Power grasps, as the term suggests, are those in which all five fingers are used to firmly grasp an object, e.g. when a glass or bottle is grasped. In Figure 3.1, the Salisbury hand and DLR-II hand are engaged in a power grasp, while the Shadow hand is not. There are two broad classifications of such grasps (Siciliano & Khatib, 2008), which are described below.

Form-Closure (enveloping grasp) occurs when the fingers are arranged to form a cage around the object. leaving no space for it to escape. In this condition, as long as the fingers remain fixed, no external perturbations are able to wrest the object free. An example of this is the Salisbury hand in Figure 3.1, which displays a form-closure grasp.

Force-Closure occurs when the finger positions are not completely wrapped around the object, but are only partially wrapped around it. In this case, some of the object's possible movements can only be blocked by applying a proper force to the object at its contact points. An example of this is the DLR-II hand in Figure 3.1, which is shown grasping a bottle using a force-closure posture. In this case, a proper pressure has to be applied to the bottle to prevent it from slipping out of the hand.

Form-closure and force-closure classify the grasp on the basis of the geometry of the finger positions and contact points. As regards the dynamic aspect of power grasps, the following are the main properties that are desired (Suárez et al., 2006).

Stability: A stable grasp is achieved when all perturbations of the object within a given threshold are automatically damped by the configuration of the contacts. If the form-closure condition is satisfied, then the grasp can withstand all possible perturbations, due the finger positions. In a force-closure grasp, the force

applied by the fingers allows the immobility of the object to be maintained. In this case, the perturbations are dealt with by appropriately controlling the force exerted by the fingers at the contact points.

Equilibrium: This is achieved when the resultant of all forces and torques applied to the object by the fingers and external forces is zero. Or, in simple words, when the object is firmly held in the hand.

Dexterity: This regards the ability to move the grasped object. The definition of dexterity varies depending upon the task to be accomplished after the object is grasped. In general, a grasp is considered dexterous if the object can be moved anywhere that is within reach of the robotic manipulator.

The controllers of a robotic arm have to deal with all the aforementioned problems, and their behaviours must demonstrate the aforementioned properties. To better identify the responsibility of the controllers, and thus, the properties to be achieved and the problems to be overcome, four spaces in which the controllers operate have been distinguished: the *task*, *work*, *joint*, and *actuator* spaces (Torras, 2003). Almost all methods and attempts to control robotic arms can be classified on the basis of the mapping among these spaces. As regards the objective of this thesis, two particular inter-space maps are important: that from the work space to the joint space (i.e. *inverse kinematics*), and that from the work space to the actuator space (i.e. *inverse dynamics*).

The task space encodes tasks in such a way as to give appropriate inputs to the planning module of the controller. Typically, the task space explicitly requires the subdivision of the controller in a planner, which maps the task space input into another space, and a number of other modules that actually perform the task. This further refinement, however, is beyond the scope of this thesis.

The work space represents all locations that can be reached by the end-effector of the arm. This is the 3D Cartesian space in which the robotic arm is able to move.

An input from the work space usually encodes the desired point to be reached, the positions of the objects, and the obstacles.

The joint space represents the posture of the robotic arm, and the actuator space represents the state of the motors acting on the joint. Their interaction depends upon velocities, forces, accelerations, and other parameters, depending upon what types of actuators are used.

Typically, neuroscience models of reaching and grasping address the problem from the task space to the work space. For example, Oztop et al. (2004) developed a neural network that models how infants learn to grasp. The neural network, which is based on the grasping schema proposed in (Arbib et al., 1985; Iberall & Arbib, 1990), is divided into four interconnected layers of encoding: the target location (o), the hand position (h), the wrist rotation (r), and the virtual fingers (v). These four vectors in the work space describe the variables that the *Movement Generation* (MG) module needs in order to move a simulated arm. The MG module is not part of the learning process, but is based upon the ad-hoc inverse kinematics of the simulated arm designed for the experiment. In order to simplify the implementation of MG, there is no simulation of rigid-body dynamics. The object is considered to be in a fixed position in the world, and any contact with the object does not cause it to move. In this way, none of the dynamic aspects of the problem is taken into account.

Also, in (Iossifidis & Schoner, 2006; Schoner & Santos, 2001; Thelen et al., 2001), various models are proposed for reaching and grasping, based upon neuroscientific data about infant learning. In this case, however, the models are implemented using the attractor dynamics approach instead of neural networks. As in the previous example, the task here also consists in the generation of appropriate trajectories in the work space. For this reason, the problem of how the trajectories in the work space are transformed/mapped into joint movements of the arm is not taken into account, but is instead delegated to external routines that implement inverse kinematics.

The works presented in this thesis propose the use of evolved neural networks in order to transform a command in the work space to a command in the joint space. Hence, they are not comparable with above example, but instead, they apply to the lower level acting upon the joints of the robotic arm. More precisely, the aforementioned models generate trajectories in the work space, and the neural networks proposed in this thesis can use those trajectories as inputs that they can then generate as output for the joint movements that put the robotic arm in a specified position in the work space. In this thesis, the neural networks have to control the robotic arm in a dynamic environment in which collisions, inertia, gravity, and other dynamic forces are simulated realistically. This represents an improvement over the MG used in Oztop et al. (2004) as well as the external routines used in (Iossifidis & Schoner, 2006; Schoner & Santos, 2001; Thelen et al., 2001), in which the dynamic aspects of object manipulation are neglected.

3.1 Inverse Kinematics

The problem of finding an arrangement of all the arm's joints, given a desired point $\mathbf{x}$ to be reached by the end-effector, is called *inverse kinematics* (hereafter, IK). In the case of the grasp, instead of a single point $\mathbf{x}$, there is a vector of points consisting of the target contact points for grasping the object. This can be considered as a process of mapping from the work space (all the points reachable by the robotic arm in a given 3D Cartesian frame) to the joint space (all possible configurable positions of the robotic arm).

In general terms, IK consists of the inversion of the *forward kinematics* equation of the robotic arm $\dot{\mathbf{x}} = \mathbf{J}(\theta)\dot{\theta}$:

$$\dot{\theta} = \mathbf{J^{-1}}(\theta)\dot{\mathbf{x}}$$

where $\mathbf{x}$ is the position of the end-effector (or a more general vector that includes

all fingertip positions), θ is a vector $< \theta_1, \theta_2, \ldots, \theta_n >$ of the current joint angles, and $\mathbf{J}$ is a Jacobian matrix that describes the relationship between the work space and joint space. In other words, the forward kinematics computes how much the position of the end-effector, $\dot{\mathbf{x}}$, would change if a given modification to the joint's arm, $\dot{\theta}$, is applied. Inverse kinematics, on the other hand, finds out how much the joint's arm would change, $\dot{\theta}$, if the position of the end-effector must be displaced by a given amount $\dot{\mathbf{x}}$.

While the forward kinematics is easy to derive even for complex and redundant robotic manipulators, the IK is more difficult to calculate due to the $\mathbf{J^{-1}}(\theta)$ term, in that there is no straightforward and closed-form solution for any type of manipulator. In fact, the inverse kinematics for redundant arms (when there are more than 3 DOFs) does not have a unique solution, and approximations are often required in order to calculate $\mathbf{J}$ (Angeles, 2003; Siciliano & Khatib, 2008). For a 6 DOF robotic arm with a special kinematic structure, a closed-form solution for IK can be obtained. Most industrial robotic arms have a kinematic structure that satisfies the requirements for solving the problem of IK (Siciliano & Khatib, 2008) by means of the following design features.

1. Three consecutive revolute joint axes intersect at a common point (as in a spherical wrist).

2. Three consecutive revolute joint axes are parallel.

When a closed-form solution exists, there are two basic approaches to finding it.

Algebraic: These methods involve algebraic manipulation of the equations of the forward kinematics. A common strategy is to find the relevant joint variables, reduce them to a transcendental equation, and then invert them into single variable equations.

Geometric: By analysing the geometric structure of the robotic arms, it is possible to ascertain the points on the manipulator at which the problem can be

decomposed. For example, the two conditions for the existence of a closed-form solution make it possible to divide the problem into inverse orientation kinematics (first condition), and inverse position kinematics (second condition). Following this division, an algebraic method is commonly applied to each sub-problem.

Even when the two conditions for the existence of a closed-form solution are met, if the joint structure of the manipulator is quite complex, it can be a very time-consuming process to find a solution to the inverse kinematics using the above methods (Angeles, 2003; Siciliano & Khatib, 2008).

If the structure of the robotic arms does not satisfy the conditions required for there to be a closed-form solution, as in the case of anthropomorphic structures, numerical methods must be used in order to find out solutions to an IK problem. These methods do not depend on a particular configuration of the robotic arms, and they can be applied to any kind of kinematic structure. Unfortunately, numerical methods are typically slower and they only approximate a solution; hence, they do not compute all possible positions of the arm.

There are various techniques for resolving an IK problem numerically (Siciliano & Khatib, 2008). Iterative methods are an example of this. In such methods, an initial sub-optimal solution is created using an empiric (and very fast) algorithm, and then it is iteratively refined in order to converge to an optimal solution. The performance of iterative methods is strongly affected by the quality of the initial sub-optimal solution, and they also converge to only one solution, depending upon the starting point.

As regards the aims of this thesis, it is important to note that neural networks have also been used to find a solution to IK problems (Kokera et al., 2004; Manocha & Zhu, 1994; Toal & Flanagan, 2002; Williamson, 1998; Li & Leong, 2003; Oyama et al., 2001; Martìn & del R Del Milla, 1998; Rathbone & Sharkey, 1999; Krose & der smagt, 1993a; Bekey, 2006). One of the crucial points here is that neural networks

have been used as an arbitrary function approximation mechanism to directly derive an approximation of the IK equation. For example, in (Kokera et al., 2004; Manocha & Zhu, 1994; Toal & Flanagan, 2002; Williamson, 1998; Li & Leong, 2003; Oyama et al., 2001; Martìn & del R Del Milla, 1998; Rathbone & Sharkey, 1999; Krose & der smagt, 1993a; Bekey, 2006), a neural network is taught to submit output for the final position of the arm in order for it to reach a given point, and then a procedure was developed by the authors in order to actually move the arm. In this way, the neural network does not directly control the trajectory, and it also cannot receive feedback regarding what is happening as the arm moves. As proposed in this thesis, the neural networks directly control all the movements of the arm, and they have the capacity to adjust its trajectory as the arm moves, by means of sensor feedbacks. In some ways this can be considered a neural implementation of an iterative method, but with the difference that neural networks are non-linear, while iterative methods are linear.

As regards the use of IK for solving the process of grasping, in addition to the problems involved in resolving IK equations, there is also the problem of finding the optimal finger positions for a power grasp. This challenge consists in finding the positions of the contact points on the object's surface that satisfy the aforementioned grasp properties. This challenge entails two problems: how to define the contact points, and how to define a mathematical formulation of the desired grasp properties.

The first problem has already been solved, and various definitions have been advanced, that depend on the properties of the fingertips (Bicchi & Kumar, 2000; Siciliano & Khatib, 2008). As for defining a mathematical formulation of the desired grasp, many attempts have been made, but they have not produced convergence to a common framework. While it is easy to visually judge whether or not a robot hand grasps an object well, it is difficult to find out how to formulate it with a function of contact points that is usable in an optimisation algorithm in order to control the robot. However, depending upon the context and the information available about the object, engineers have developed many functions for measuring the quality of a

grasp, and hence can constrain all possible solutions in order to find the best grasp configuration for a given task (Suárez et al., 2006).

From a review of such functions (Suárez et al., 2006), it seems unlikely that primate or human brains use some of these functions in order to plan, learn, or simply understand if they have grasped an object appropriately. The reason for this is that all the grasping-quality metrics that are used to develop controllers in industrial manipulators need precise information about the geometry of the contact points and of the object's shape — information that is unlikely to be derived from visual-tactile sensory information alone.

The main objective of grasping quality functions is to obtain a very good grasp. However, when some of these metrics are used to evaluate the quality of human grasps (Veber et al., 2005), the humans do not perform as well as one would expect. Hence, it seems that there are some peculiarities in how humans choose a grasp configuration that are not 'grasped' by the quality functions that have thus far been proposed.

The above analysis demonstrates the complexity of the task. In fact, for a given object there are infinite possible contact point configurations, and in addition, in reaching for an object, there are infinite trajectories. Hence, the solution must be searched for in a very huge space. For this reason, it is unlikely that there exists an analytic or closed-form solution for general cases or for non-engineered environments.

3.2 Inverse Dynamics

Inverse kinematics takes into account only the geometric structure of the arm's joints. In real-world situations and when an adaptive robotic controller is required, however, this is not enough. In fact, when a manipulator moves the joints quickly, or when it is attached to a moving platform or in other similar situations, it is not possible to ignore inertial effects, frictions, and gravitational forces. In inverse

dynamics, then, the controller also takes into account the velocities, acceleration, and other forces that are applied to the joint's actuators (Angeles, 2003; Torras, 2003).

This approach can be viewed as a process of mapping that involves the actuators' space. Given a desired point in the work space, the controller generates a sequence of motor commands (forces, velocities) that drive the arm to a final position in which the position of the end-effector coincides with the desired point in the work space. Hence, in addition to the previous equation, the following has to be considered (and inverted, in order to find a solution using analytic approaches):

$$\tau = \mathbf{I}(\theta)\,\ddot{\theta} + \mathbf{V}\left(\theta, \dot{\theta}\right) + \mathbf{B}\left(\theta, \dot{\theta}\right) + \mathbf{G}(\theta)$$

where τ is the torque vector, $\mathbf{I}$ is the $n \times n$ inertia matrix of the arm, n is the number of DOFs, $\mathbf{V}$ represents the centrifugal and Coriolis acceleration terms, $\mathbf{B}$ is the friction terms, and $\mathbf{G}$ is the vector of the gravity terms.

This equation is very complex to invert because the inertia, centrifugal, and gravity terms also depend on time due to the arm movements, and they are also influenced by the controller's commands. Thus, closed-form solutions that use algebraic or geometric methods might not exist even for simple robotic structures, and numerical approaches are often the only way to obtain an approximation of the inverse dynamics (Angeles, 2003; Siciliano & Khatib, 2008).

Using the inverse dynamics for grasping is often quite complex, even in a simple case. Hence, a typical constraint that is used to simplify the solution for grasping is to avoid any movements of the object from the moment of first contact to the completion of the grasp. In addition, in situations in which the manipulator is not provided with detailed information about the object, it is important to exploit the interactions with the object and to use sensory feedback in order to adjust the fingers. The rationale behind this is that the hand has to adjust itself in order to grasp the object. This approach does not consider the possibility that the hand can

rotate and move the object to orient it with the palm, which adds a lot of complexity to the problem.

Felip & Morales (2009), for example, have proposed an algorithm that can govern interaction with the object to be grasped. The algorithm is based on three phases, or procedures, which are controlled by sensory feedback. The main aim of the algorithm is to produce stable grasps of unknown objects without having detailed information about them. Briefly, the three phases are as follows.

Alignment: The palm is systematically rotated $\pm 2^{\circ}$ until a straight movement of the palm, when one or more fingers are in contact with the object, will not produce axial torques of the palm.

Parallel Face Detection: The hand is moved up and down about $5\,mm$. The variations in the grip aperture that are needed to keep the fingers in contact with the object are used to rotate the palm in order to align it with the object's lateral faces.

Force Adaptation: The fingers begin to close and the object is lifted. If any slippage of the object is perceived, the force applied to the object is increased until the slippage stops.

It has been demonstrated that this algorithm works with a variety of objects, with a $40\,s$ average time for grasping an object. Compared with the time needed by humans to grasp the same objects, this is a very long time. Most of this time is taken up by the first and second phases, because the movements that align the palm and the fingers are very slow, to avoid moving the object.

4 Background to Experiments

All the experiments reported in this thesis concern the synthesis of a neural controller for anthropomorphic robots that are able to manipulate objects through an automatic design process based upon artificial evolution. The scenarios of the experiments presented here are of significant complexity, due to the difficulty of controlling a system with many degrees of freedom, the need to master the effects produced by gravity, inertia, collisions, etc. (the role of gravity, inertia, collision, and noise are taken into account by accurately simulating the physical laws of motion and the effect of collisions). In the experiment on "Reaching and Grasping" (Chapter 6), the task is further complicated by the need to grasp two objects that differ in shape, mass, and inertia, while in the experiment on "Manipulation and Object Discrimination" (Chapter 7), the task is complicated by the great similarity among the objects to be distinguished from one another.

The thesis has been structured in order to correspond with the relevant lines of research that the author has pursued, which began in 2005 and ended in 2010. The chapters are thus presented in both the chronological and logical order that was followed by the author during his research. It is important to note that during the course of these Ph.D. studies, most of the results were published for conferences and in journals, and that the chapters largely follow the order of the author's publications (copies of which are bound in the thesis). Hence, the relevance of the discussion and the contributions of the knowledge presented here should be evaluated with respect to the year in which the work was done and was published. The following table presents that order of publication, and summarises the main characteristics and differences between the experiments conducted and the corresponding publications.

	Task	Robot	Controller
Chapter 5 (Massera et al., 2006)	Reaching a target position	Arm controlled in terms of velocity without hand (App.A)	Neural network without internal neurons

Chapter 6 (Massera et al., 2007)	Reaching for and grasping two different objects: a sphere and a cylinder	Arm controlled by muscle actuators and five-fingered hand (App.B)	Neural network with recurrent internal neurons and some direct input-output connections

In Parallel

	Chapter 7(Tuci et al., 2010)	**Chapter 8**(Massera et al., 2010)
Task	Categorisation and discrimination between two similar objects: a sphere and an ellipsoid	Performing a sequence of actions (reach, grasp and lift) on the basis of incoming 'linguistic' instructions
Robot	Arm controlled by muscle actuators and five-fingered hand (App.B)	Arm controlled by muscle actuators and five-fingered hand (App.C)
Controller	Continuous time recurrent neural network	Neural network with recurrent internal neurons

For each experiment, the neural network architecture was chosen based upon the particular setup of the experiment, in an effort to keep it as simple as possible. In the extreme case of the experiment in Chapter 5, the neural network is so simple that it does not have any hidden neurons. For the other experiments, various characteristics were added when called for by the experimental setup. In the case of the experiment regarding distinguishing between two objects, Elio Tuci (the first author of the corresponding publication) selected and designed the neural network and part of the experimental setup, and the author of this thesis designed the way in which the controller categorises objects, which is one of the most important results of thate experiment.

In this thesis, the attempt to design robots that develop their skills autonomously by evolutionary robotics makes it possible, at least in principle, to delegate the solutions

to some of these aspects to the adaptive process itself, and to apply the empirical evidence that indicates that the manipulation skills learned early on by infants are acquired through self-learning mechanisms, rather than by imitation (Oztop et al., 2004). In fact, the evolutionary algorithm applied in this thesis uses only a mutation operator and elitism, which allows it be viewed as a kind of self-learning mechanism. To explain this point of view, let us suppose that n generations are done. Due to elitism, it is always possible to trace back from the best individual of generation n to an individual of the first generation, thus obtaining a single sequence of mutations that produced the last individual. This sequence of mutation can be viewed as being similar to how a robot acquires its skills through a process of trial and error, during which random variations in the free parameters of the robots' neural controller (which are initially assigned randomly) are retained or discarded on the basis of their effect at the level of the overall behaviour exhibited by the robot in its interaction with the environment. More precisely, the effect of variations is evaluated using a set of utility functions that determine the extent to which the robot manages to manipulate a target object. The use of this adaptive algorithm and utility functions leaves the robot free to discover, during the adaptive process, its own strategy to achieve the goals set by the experimenter. This in turn allows the robot to exploit sensory-motor coordination (i.e. the possibility of acting in a certain way in order to later experience useful sensory states), as well as the properties arising from the physical interactions between the robot and the environment.

One of the main characteristics of the models presented in this thesis is that the robot controller adjusts its output on the basis of the available sensory feedback by directly updating the forces exerted on the joints (see Schaal et al., 2005 for related approaches). The importance of the sensory feedback loop has been emphasised in other works in the literature. For example, in Felip & Morales (2009), the authors describe an experiment in which a three-fingered robotic arm displays a reliable grasping behaviour through a series of routines that continually modify the relative positions of the hand and fingers on the basis of the current sensory feedback. These

movements tend to optimise a series of properties such as hand-object alignment, contact surface, finger position symmetry, etc.

The robot controllers that process sensory and propriosensory information and control the state of the arm/hand actuators are modelled through the use of dynamical recurrent neural networks. The architectures of the artificial neural networks currently employed are not inspired by the characteristics of the neuroanatomical pathways of the human brain. Also, many of the features of the neurons and synapses are not taken into account (see Oztop et al., 2004, for an example of works that do emulate some of the anatomical characteristics of the human brain). The use of artificial neural networks as a robot controller provides several advantages with respect to alternative formalism, such as robustness, graceful degradation, generalisation, and the possibility of processing sensory-motor information in a way that is quantitative both in state and time (Bar-Yam, 1997a; Haykin, 1999). These characteristics also make neural networks particularly suitable for use with evolutionary robotics, in which a suitable configuration of the free parameters is obtained through a process that operates through the accumulation of small variations.

Most of the characteristics of the simulated robots that have been implemented are the same as those in all the experiments reported here, but some are different. Appendices A, B, andC report the complete details of the models of the anthropomorphic robotic arms that are implemented, highlighting their differences and improvements with respect to the others. More specifically, the morphological characteristics of the human arm and hand are taken into account by using a robot that approximately reproduces the morphological characteristics of a 3.5 year-old in terms of size, shape, articulations, degrees of freedom, and relative limits (Shadmehr & Wise, 2005b; Sandini et al., 2004; Jones & Lederman, 2006). In all the robotic arms implemented, some of the properties of the human musculoskeletal system have been incorporated into the model by using muscle-like actuators that are controlled by antagonistic motor neurons (Shadmehr, 2004; Shadmehr & Wise, 2005c; Gialias & Matsuoka, 2004). For the sake of simplicity, the segments forming the arm, the

palm, and the fingers are simulated as completely rigid bodies. However, the way in which the fingers are controlled endows the hand with a certain level of compliance.

In the experiments reported in this thesis, the robots are equipped with proprioceptive and tactile sensors (when necessary), and with a vision system that only provides information concerning the position of the object, and not information about its shape and orientation. The vision system has been simplified to a greater degree than the proprioceptive and tactile sensors, because the experimental evidence on young infants indicates that they rely heavily on somatosensory and tactile information to carry out reaching and grasping actions, and that they use vision to elicit these actions (Rochat, 1998). While visual information about the shape and orientation of an object (which is employed to prepare the grasping behaviour or to adjust the position of the hand) only starts to play a role 9 months following birth (McCarty et al., 2001), the capacity for precision grasping is developed after 12–18 months (Hofsten, 1982, 1984, 1991). In this thesis, the focus is upon the power grasp and the experiments presented here show that it is possible to successfully develop strategies for grasping different objects without the use of a more sophisticated vision system.

4.1 Evolutionary Robotics

In the experiments presented here, the *Evolutionary Robotics* methodology has been employed to develop neural networks that are able to control an anthropomorphic robotic arm to reach, grasp, and distinguish one object from another. Evolutionary Robotics (hereafter, ER) is inspired by Darwin's theory of evolution (Darwin, 1859), in its development of autonomous agents. Robotic agents compete with each other for survival, and only the best individuals are allowed to reproduce in the next generation. Indeed, the basic idea is quite simple: the robots with the best performance are selected, and through a simulated reproduction process, they generate offspring that are largely similar to themselves, but slightly different, having incorporated a

number of random variations. Although the majority of random variations lead to robots with a lower level of performance, some of them lead to robots that perform better than their parents. These positive variations are retained in the selection process, while the negative ones are discarded, and the cyclic repetition of this selection and reproduction process tends to produce a population of robots that are able to perform the task that is required. Although Darwin's theory of evolution is more than a century old, the idea of using an evolutionary approach in the design of autonomous robots is more recent. In 1984, Braitenberg imaged the experiments reported in (Braitenberg, 1984), and the term *Evolutionary Robotics* was coined at the beginning of 1990 (Cliff et al., 1993). In this period, various researchers began to use evolutionary techniques in robotics research (Cliff et al., 1993; Floreano & Mondada..., 1994; Nolfi et al., 1994; Nolfi & Floreano, 2000; Harvey et al., 2005; Floreano et al., 2008).

One of the peculiarities of ER is its generative paradigm. The evolutionary process produces many possible solutions, generation after generation. When the evolutionary process ends, many good solutions are generated that represent agents that are able to accomplish the task in an efficient manner. In ER, the behaviour of these agents becomes the object of post-evolution analysis, whose goal is to understand the different solutions, the characteristics they have in common, and the solutions that have been discovered. These analyses are then likely lead to new ideas for further evolution and increased understanding of how adaptive solutions emerge (Nolfi, 2005a).

Another important characteristic of ER is the possibility it offers to evolve complex dynamical systems in which the strategy for solving the required task is an emergent property of such dynamical systems. The strength of a complex dynamical system lies in the non-linearity of the interactions among its components and the external environment. These interactions have the potential to produce very powerful and reliable solutions for complex task, starting from simple components (Bar-Yam, 1997b). The ER paradigm simplifies the design of complex systems, because the

fitness rewards the overall global behaviour of a complex system, which can be easily evaluated, while there is no constraint on the possible underlying interactions.

Furthermore, with the ER paradigm there is no need to precisely specify the desired output, as is required in error-minimising learning. This is what makes it possible to tackle the inverse dynamic problem for a redundant anthropomorphic arm, as was done in the experiments reported in Chapters 6, 7, and 8 of this thesis (Massera et al., 2007, 2010; Tuci et al., 2010). In fact, in supervised approaches for a given input, the controller has to generate a sequence of forces to apply that are difficult to calculate a priori, and difficult to learn using error-minimising procedures or reinforcement learning.

4.2 Reaching and Grasping

The main difference among the approaches reported in Chapter 3 and the ER approach used in this thesis regard how a neural network is considered. In the studies referenced in Chapter 3, a neural network is used to generate correct mapping between two different spaces; i.e. in inverse dynamics, from the work space to the joint space. From this point of view, neural networks are likely to be considered as an arbitrary function approximation mechanism. In the ER approach, on the other hand, the neural controller is seen as an internal dynamical system (Bar-Yam, 1997c,a) that interacts with environments via the agent's body. There is no explicit mapping between spaces, but such mapping emerges from the continuous, minute interactions among the controller, the body, and the environment. From this point of view, the agent's behaviour is an emergent property of these minute interactions. The evolutionary process is able to exploit the potential of simple architectures via dynamical interaction, and is likely to lead to complex adaptive behaviour starting from minimal agents. Hence, some of the weaknesses of other approaches can be resolved by the use of ER.

In particular, it has been shown that due to the global effect of weights, an accurate

mapping of inverse kinematics using feed-forward neural networks is extremely difficult to obtain (Krose & der smagt, 1993b; Torras, 2003). That said, the experiment reported in Chapter 5 was designed using an er approach, keeping in mind that the controller is a dynamical system that is directly and continuously acting upon the dynamics of agent-environment interactions. In this way, it is possible to exploit a simple architecture, such as a multi-layer perceptron, to learn inverse kinematics/dynamics solutions in order to solve the problems involved in reaching (Massera et al., 2006, 2005; Bianco & Nolfi, 2004).

Remaining within the field of ER, there have been a few previous attempts to use an ER approach to develop a controller for a robotic arm.

Nolfi & Marocco (2002) studied the case of a simulated robotic "finger", which was evolved to distinguish between the shapes of spherical versus cubic objects (anchored to a fixed point) of different sizes and orientations. The robotic finger consists of an articulated structure made of three segments that are connected via motorised joints with six DOFs, six corresponding actuators, six proprioceptive sensors that encode the current position of the joints, and three tactile sensors placed on the three corresponding segments of the finger. The authors observed that the adapted robots solve their problems through the use of simple control rules that cause the robot to scan for an object by moving horizontally from the left to the right side, and by moving slightly upward as a result of collisions between the finger and the object. These simple control rules cause two different behaviours to be displayed:

- With spherical objects, the robotic finger fully extends itself on the left side of the object after following the object's surface.
- With cubic objects, the robotic finger remains fully bent, close to one of the corners of the cube.

These two behaviours correspond to well differentiated activation of the proprioceptive sensors. These differences are used by the finger to distinguish between the two

types of objects. We would note that although the discrimination cue that is necessary in order to categorise is available in each single sensory pattern experienced after the display of the appropriate behaviour, this cue results from a dynamical process that arises as a result of several robot/environment interactions.

Bianco & Nolfi (2004) also used an ER approach in the autonomous design of a neural controller for a simulated robotic arm with a two-fingered hand and nine DOFs that give it the ability to grasp objects with different shapes. The arm was provided only with tactile sensors. The evolved robots displayed the ability to grasp objects with different shapes and different orientations that were located in a limited number of various positions within a limited area. These robots, however, were not able to deal with larger variations of the objects' positions.

Buehrmann & Paolo (2004) evolved a control system for a simulated robotic arm with three DOFs that has the ability to reach a fixed object placed on a plane and to track moving objects. The arm was equipped with two pan-tilt cameras, each consisting of a two-dimensional array of laser range sensors that were placed above the robot arm and on the end-point of the robotic arm. The controller consisted of several separate neural modules that receive different sensory information and control different motor joints. The networks are evolved separately to create the ability to produce distinct elementary behaviours (e.g. change the orientation of the camera so it will focus on the object; move the first joint, which determines the orientation of the arm in order to orient it towards the object; approach the object by controlling the second and third joints, etc.). The authors show that the evolved neural controller can exploit its own actions in order to self-select stimuli that facilitate the accomplishment of spatial and temporal coordination.

Marocco et al. (2003) used an arm model with 6 DOFs to evolve the ability to touch or avoid objects depending upon their shape. In addition to the capability of distinguishing among objects, the robots were also evolved for their ability to "name" the object (or the action) with which they are interacting. This permitted the analysis of different social interaction protocols, which was then used to investigate

the social and cognitive factors that support the evolutionary emergence of shared lexicons. Although this model used a very simplified arm model and a limited object set and location, it was the first attempt to use a neurorobotic model to study the links between action and linguistic representations from an evolutionary perspective.

Bongard (2010) co-evolved both a neural controller and certain aspects of a robot's morphology to manipulate objects, using multi-objective optimisation based upon evolutionary algorithms (Deb, 2001). In this case, the task consisted of the optimisation of various abilities: grasping, lifting, and distinguishing among objects. The neural controllers had to be able to enact all these capacities simultaneously, and seven different evolutionary setups were compared. In one of these, only the neural network's parameters were evolved, while in the others, an increasing number of the robot's morphological characteristics were co-evolved using the parameters of the neural network. The goal of Bongard (2010) was not to demonstrate that a robot could accomplish these tasks, but rather, to show that there is a positive correlation between the likelihood of successfully performing the tasks, and the extent of the robot's morphology that was placed under evolutionary control.

4.3 Active Categorical Perception

The scenario of the experiment in Chapter 7 was designed in order to investigate the perceptual skills of an autonomous agent that are required in order to actively categorise un-anchored spherical and ellipsoid objects that are placed in different positions and orientations on a planar surface.

Categorical perception can be considered as the ability to divide continuous signals received by the sense organs into discrete categories whose members resemble one another more than they do the members of other categories. Categorical perception represents one of the most fundamental cognitive capacities displayed by natural organisms, and it is an important prerequisite for the display of several other cognitive skills, as discussed in (Harnad, 1987). Not surprisingly, categorical perception has

been extensively studied both in natural sciences such as Psychology, Philosophy, Ethology, Linguistics, and Neuroscience, and in artificial sciences such as Artificial Intelligence, Neural Networks, and Robotics; see (Cohen & Lefebvre, 2005) for a comprehensive review of this research field. In the great majority of cases, however, researchers have focused their attention on categorisation processes that are passive and instantaneous. Passive categorisation processes take place in experimental setups in which the agents cannot influence the experienced sensory states through their actions. In instantaneous categorisation processes, the agents are required to categorise the current experienced sensory state, rather than a sequence of sensory states distributed over a certain time period.

The experiment in Chapter 7 focuses upon categorisation processes that are active and that are eventually distributed over time (Beer, 2000; Nolfi, 2009). This task is achieved by exploiting the properties of autonomous embodied and situated agents. An important consequence of being situated in an environment consists in the fact that the sensory stimuli experienced by an agent are co-determined by the action performed by the agent itself. That is, the actions and the behaviour exhibited by the agent later influence the stimuli that it senses, their duration in time, and the sequence in which they are experienced. This implies that:

1. categorical perception is strongly influenced by an agent's action; see also(Gibson, 1977; Noe, 2004) as regards this issue; and

2. sensory-motor coordination (i.e. the ability to act in order to sense stimuli or sequences of stimuli that allow an agent to perform its task) is a crucial aspect of perception and more generally of situated intelligence, see (Pfeifer & Scheier, 1999).

A growing body of literature in robotics is increasingly devoted to the effort to obtain discrimination among material properties (e.g. hardness, texture) and object shapes using the sense of touch in artificial arms. Many of these works, such as

the one described in (Takamuku et al., 2007), draw their inspiration from human perceptual capability in order to develop highly elaborated touch sensors.

In (Takamuku et al., 2007), the authors describe a tendon-driven robotic hand covered with artificial skin made of strain gauges sensors and polyvinylidene films. The strain gauge sensors mimic the functional properties of Merkel cells in human skin in detecting strain. Polyvinylidene films mimic the functional properties of the Meissner corpuscles in detecting the velocity of strain. The artificial hand, through the execution of squeezing and tapping procedures, manages to discriminate among objects based on their hardness.

In a similar vein, a research group at Lund University developed three progressively more complex versions of a robotic hand (LUCS Haptic Hand I, II, and III) that was designed for tasks involving haptic perception (Johnsson et al., 2005; Johnsson & Balkenius, 2006, 2007a). The perceptual capabilities of the three versions of LUCS, which differ in both their morphology and sensory capabilities, were tested during the execution of a grasping procedure using objects made of different materials (e.g. plastic and wood). The authors showed that the sensory patterns generated in the hands' interactions with objects are rich enough to be used as a basis for haptic object categorisation (Johnsson & Balkenius, 2007b).

Other robotics systems have combined visual and tactile perception to carry out fairly complex object discrimination tasks (Dario et al., 2000; Natale & Torres-Jara, 2006; Stansfield, 1991).

Generally speaking, in spite of the heterogeneity of hardware and control design, the abovementioned research works focus upon the characteristics of the tactile sensory apparatus and/or categorisation algorithms. In these works, the way in which the sensory feedback affects the movement of the hand is determined by the experimenter on the basis of his/her intuition. Moreover, the discrimination phase follows the exploration phase, and is performed by elaborating sensory data gathered during the manipulation of objects (i.e. the data collected during the exploration

phase cannot influence the agent's behaviour during that phase).

The work described in Chapter 7 differs from the abovementioned literature in two significant ways:

- The way in which the agent interacts with the environment is not designed by the experimenter, but is adapted in order to facilitate the categorisation task, and
- The agent is left free to generate its motor behaviour on the basis of previously experienced sensory states.

Rather than studying the performance of particularly effective tactile sensors or of specific categorisation algorithms, the focus is on the development of autonomous actions for distinguishing the shapes of objects via coarse-grained binary tactile sensors and proprioceptive sensors.

The issue of how a robot can actively develop categorisation skills has been already investigated in a few recent research works. In general terms, these works demonstrate how adapted robots exploit their actions in order to self-select stimuli that facilitate and/or simplify the categorisation process, and show how this leads to solutions that are parsimonious and robust (Scheier & Lambrinos, 1996; Scheier et al., 1998; Nolfi, 2002; Beer, 2003).

Unlike in the case of the experiments described in (Nolfi & Marocco, 2002; Bianco & Nolfi, 2004; Buehrmann & Paolo, 2004; Marocco et al., 2003), sensory-motor coordination does not always guarantee the perception of well-differentiated sensory states in different contexts corresponding to different categories. In these circumstances, the agents can actively categorise their perceptual experiences by integrating ambiguous sensory information over time. A few studies have already shown that evolved wheeled robots compensate for sensory patterns that are unreliable due to their coarse sensory apparatus by acting and reacting to temporally distributed sensory experiences in such a way as to bring forth the necessary regularities that

enable them to associate a stimulus with its category (Tuci et al., 2004; Gigliotta & Nolfi, 2008).

In the case of the task of discrimination (Chapter 7), the evolved robots act in such a way that they experience the regularities that enable them to appropriately categorise the shapes of objects. However, sensory-motor coordination does not seem to guarantee the perception of fully differentiated sensory states that correspond to different categories. The problem caused by the lack of clear categorical evidence is solved through the development of the ability to integrate ambiguous information over time through a process of accumulation of evidence.

4.4 Language and Action

The aim of the study presented in Chapter 8 is to investigate whether the use of linguistic instructions facilitates the acquisition of a sequence of complex behaviours. Neural networks are evolved to produce the ability to manipulate spherical objects located on a table by reaching for, grasping, and lifting them under two different conditions:

- While receiving as input a linguistic instruction that specifies the type of behaviour to be exhibited during the current phase, and
- Without receiving such input.

The obtained results shown that the linguistic instructions facilitate the development of the required behavioural skills.

One assumption behind the research presented in Chapter 8 is that the activity of developing robots that display complex cognitive and behavioural skills should be carried out by taking into account the empirical findings in psychology and neuroscience that show that there are close links between the mechanisms of action and those of language. As shown in (Cappa & Perani, 2003; Glenberg & Kaschak, 2002;

Hauk et al., 2004; Pulvermuller, 2002; Rizzolatti & Arbib, 1998), action and language develop in parallel, influence each other, and are based upon each other. If applied to the world of robotics, the co-development of action and language skills might make it possible to transfer the properties of the knowledge represented by action to linguistic representations, and vice versa, thus making possible the synthesis of robots with complex behavioural and cognitive skills (Cangelosi et al., 2007, 2010).

Another assumption is that the behavioural and cognitive skills of embodied agents are emergent dynamical properties that have a multi-level and multi-scale organisation. Behavioural and cognitive skills arise from a large number of fine-grained interactions that take place within the robot's body, its control system, and the environment, as well as among these three realms (Nolfi, 2005a). Handcrafting the mechanisms underlying these skills can be a difficult task. This is due to the inherent difficulty in figuring out, from the point of view of an external observer, the detailed characteristics of the agent that, as a result of the interactions between the elementary parts of the agent and of the environment, lead to the exhibition of the desired behaviour. The synthesis of robots that can display complex behavioural and cognitive skills can instead be obtained through an adaptive process. In such a process, the detailed characteristics of the agent are subjected to variations which are then retained or discarded on the basis of their effects at the level of the overall behaviour exhibited by the robot in its environment (Nolfi, 2005a). Therefore, the role of the designer should be limited to the specification of the utility function, which determines whether variations should be preserved or discarded, and eventually to the design of the environmental conditions in which the adaptive process takes place (Weng et al., 2001; Weng, 2004; Nolfi, 2005b).

5 Reaching

The first experiment presented in this chapter is a preliminary study of the development of the control system for an anthropomorphic robotic arm with 4 DOFs using an evolutionary robotic technique (Nolfi & Floreano, 2000). The control system consists of a simple neural network that directly controls the direction and the intensity of the velocities that are applied to the motorised joints. The 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).

An analysis of the evolved robots indicates that they are able to solve the assigned task and that they then generalise their skill in applying it to different target positions and to moving targets. Overall, the obtained results demonstrate that an effective reaching behaviour can be developed without relying upon internal models that perform direct and inverse mapping.

5.1 The Robot

The simulated robot consists of cylindrical segments articulated by revolute joints, as illustrated in Figure 5.1. More details about the robotic arm used in this experiment can be found in Appendix A.

5.2 The Neural Controller

In this model, the neural network controls the 4 DOFs in order for the arm to reach a given point in the given space. The neural controller consists of a feed-forward neural network with 3 sensory neurons that are directly connected to 4 motor neurons, as shown in Figure 5.2.

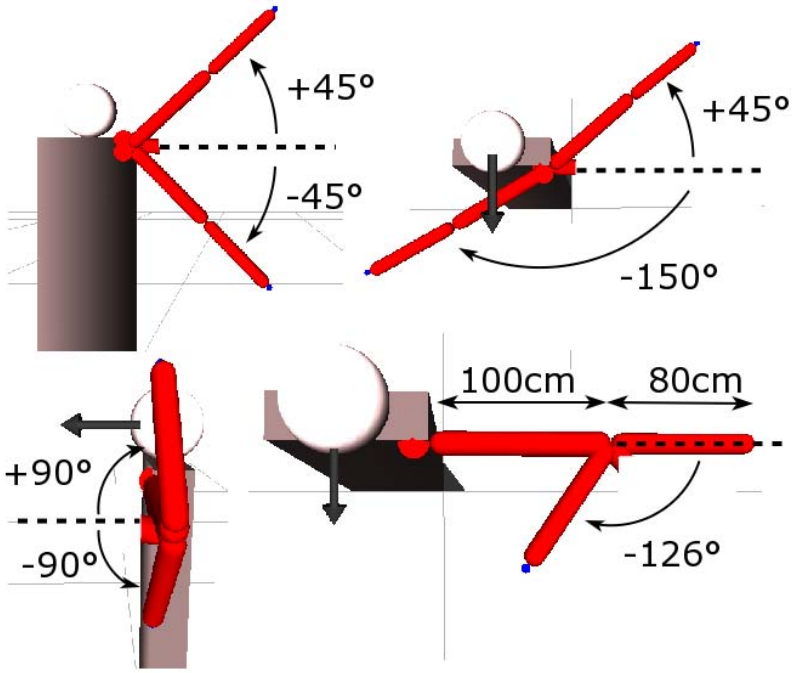


Figure 5.1: Robot structure for the reaching experiment. The four DOFs of the simulated robotic arm. The two diagrams at the top illustrate the abduction/adduction (left) and extension/flexion (right) of the shoulder joint. The two bottom figures illustrate the rotation of the shoulder (left) and the extension/flexion of the elbow (right). In all the diagrams, the arrows indicate the frontal direction of the robot.

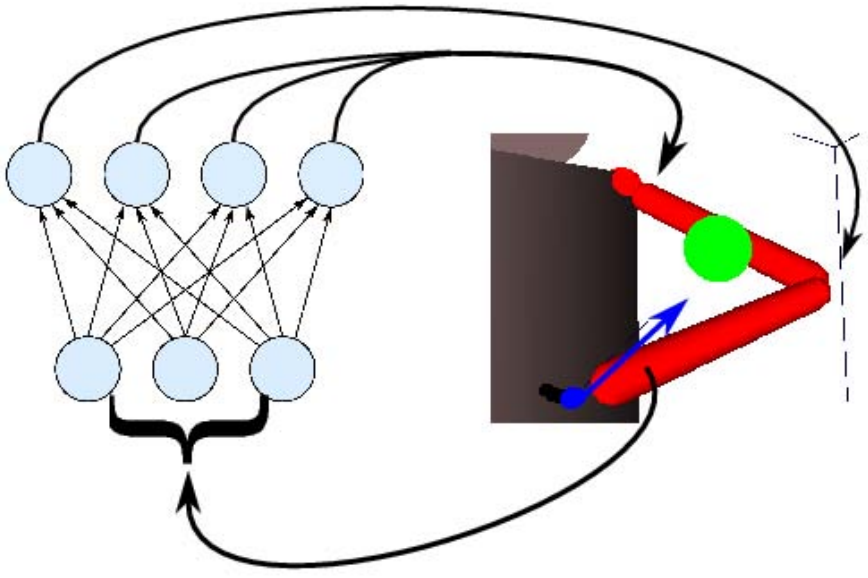


Figure 5.2: The Neural Network controlling the robotic arm. The bottom three circles represent the input neurons, and the blue arrow is the distance vector that is fed into the input neurons. The top four circles are the output neurons, which set the velocity of the associated joint, as shown by the bold black arrows.

The 3 **Sensory Neurons** can be seen as the output of a vision system (which has not been simulated) that computes the relative distance of an object from the hand up to a distance of $80\,cm$, and normalised in the range of $[-1, +1]$ over three orthogonal axes.

The 4 **Motor Neurons** encode the angular velocity of the four corresponding motorised joints. Each motor neuron receives one incoming synapse from each internal neuron, and their output is updated every $0.015\,s$ on the basis of the following equation:

$$
\begin{aligned}
A_i &= \sum_{j=1}^{3} w_{ji}\sigma_{1.0}\left(x_j\right) \\
y_i &= \begin{cases} -890 & if\ A_i < -890 \\ A_i & otherwise \\ +890 & if\ A_i > +890 \end{cases}
\end{aligned}
$$

where y_i is the output of the i-th motor neuron, and also the velocity expressed in rpm (revolutions-per-minute) to set on the corresponding joint. A_i is the net activation of motor neuron i, and it is clamped into $[-890, +890]$ in order to prevent overly rapid movement of the joint's arm. x_j is the output of sensory neuron j. w_{ji} is the synaptic weight that connects the sensory neuron j to the motor neuron i, and $\sigma_\lambda(x) = (1 + e^{-x})^{-\lambda}$ is the standard logistic function.

5.3 The Evolutionary Process

The connection weights of the neural controller were evolved as reported in (Nolfi & Floreano, 2000). The genotype of evolving individuals encodes the connection weights of the neural controller. Each connection weight is encoded with $16\,bits$ and is normalised in the range of $[-10, +10]$, making a total of $12 \cdot 16 = 192\,bits$ for

each genotype. The size of the evolved population is 100. The 20 best individuals in each generation were allowed to reproduce by generating 5 copies, with 1.5% of their bits replaced with a new randomly-selected value (the reproduction is asexual). The evolutionary process lasted for 1,000 generations. The experiment was replicated 10 times, starting from different, randomly generated genotypes.

In this simulation, the evolving controller was evolved to produce the ability to reach the target as fast as possible and stay on it. In order to obtain neural networks that are able to arrive at targets that are distributed anywhere in the reachable arm space, each individual was tested for 16 trials that differed in terms of the initial arm posture. In detail, the joint space of the arm was divided into 16 non-overlapping sub-spaces, and in each trial, the joint's initial configuration was taken from one of these sub-spaces. In all 16 trials, the target was positioned in front of the robot, and each trial lasted 4.5 s (i.e. 300 steps of simulation).

An incremental fitness function was developed in order to avoid local-optima in the reaching ability:

$$F = \frac{1}{16 \cdot 300} \sum_{i=1}^{16} \sum_{t=0}^{300} dist\,(x, r) \tag{5.1}$$

Expressed in words, the fitness function is the average of all steps of all trials of the following function:

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

where x is the Euclidean distance between the end-effector of the arm and the target point, and r is a threshold that is initially set to 10 cm. The fitness function ranges from 0 to 100. During the evolutionary process, the threshold r is progressively reduced every time the average fitness of the individuals exceeds 78. The threshold r represents the requested precision of reaching. Hence, if the threshold is high (i.e.

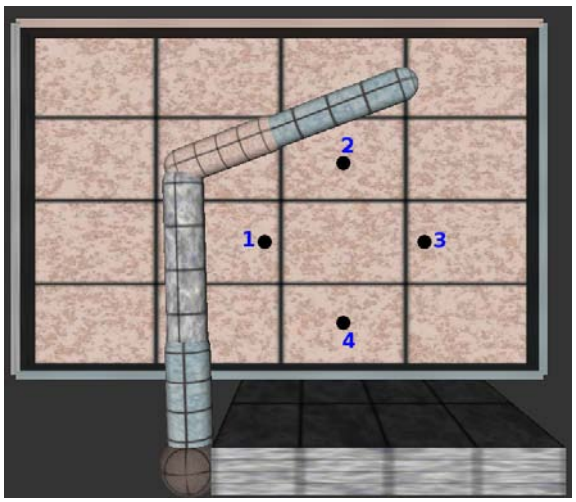


Figure 5.3: Scenario used to explain what local-optima are avoided by the fitness function in equation 5.1. The four points represent four different target positions that the robot should reach. See the text for details.

$10\,cm$) the task is quite easy, and when the threshold becomes increasingly smaller, the task becomes increasingly more complex. Thus, the incremental fitness is about increasing the difficulty of reaching (reducing the threshold r) when almost all the individuals become good enough (with an average fitness above 78).

This particular fitness formulation also helps to avoid local-optima. To explain what kinds of local-optima are avoided, let us suppose the evaluation of two individuals, A and B, during four trials in which the initial posture is fixed and the target point displacement changes, as shown in Figure 5.3. Let us further suppose that agent A reaches targets 1 and 3 with $1\,cm$ of error, and targets 2 and 4 with $9\,cm$ of error, while agent B reaches all the targets with $4\,cm$ of error. A non-incremental error-minimising function (i.e. $dist(x, 0)$ in equation 5.1) will assign a fitness value of 30.88 to A and 15.53 to B. On the other hand, the proposed fitness function in which r equals 3 will assign a value of 52.49 to A and 60.65 to B. In the first stage of evolution, the selection of B against A makes it possible to evolve individuals that are not focused upon specific areas (target points 1 & 3), but that are able to arrive roughly at every target displaced in the reachable space. The gradual reduction in r increases the pressure on the agents to perform reaching with more and more precision.

The agent A is a local-optima, because the majority of paths that lead to better individuals pass through agents whose performance on targets 1 and 3 is a bit worse, while its performance improves on targets 2 and 4. For instance, if we suppose that an offspring of A reaches targets 1 and 3 with 1.5 *cm* of error, and 2 and 4 with 6 *cm* of error, it nonetheless seems that a good improvement in the non-incremental function will assign 26.10, less than A.

5.4 Results

Ten different replications of the evolutionary set-up were run, starting from different randomly generated populations of genotypes. In all of the above, the evolved agents displayed the ability to reach the target object with precision, even with the randomness of the initial arm posture.

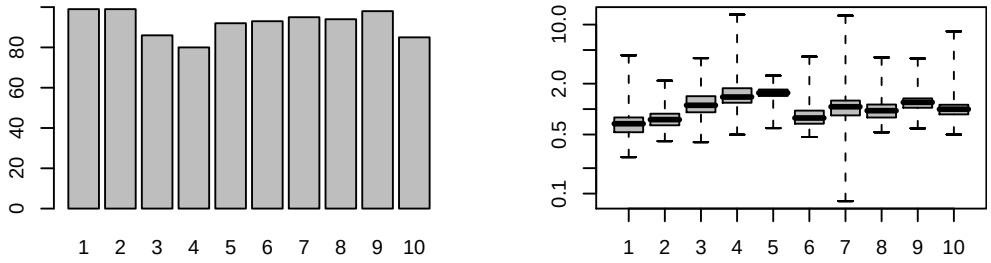


Figure 5.4: Performance on reaching a fixed target. **a)** Percentage of trials in which the distance between the endpoint of the arm and the target was less than 1 *cm* at the end of the trial. **b)** Average distance between the endpoint of the arm and the target at the end of the trials. Each column represents the performance obtained by testing the best evolved individual in each replication for 100 trials. The bold lines, grey histograms, and bars indicate the average performance, variance, and minimum and maximum values, respectively.

For each replication, the best individual was tested in over 100 trials in which the target was placed in a fixed position and the initial arm posture varied. Figure 5.4-a shows, for each replication, the percentage of trials in which the distance between the target and the endpoint of the arm was less than 1 *cm* (which is considered successful reaching). The best performance rate was 92.1% of reaches that were

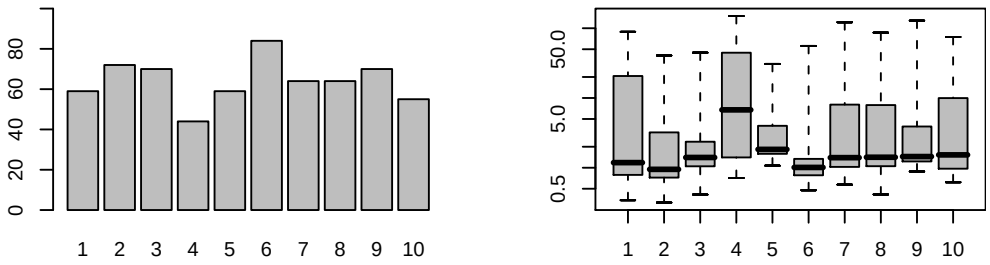


Figure 5.5: Performance on reaching a randomly positioned target. **a)** Percentage of trials in which the distance between the endpoint of the arm and the target was less than 1 *cm* at the end of the trial. **b)** Average distance between the endpoint of the arm and the target at the end of the trials. Each column represents the performance obtained by testing the best evolved individual in each replication for 100 trials. The bold lines, grey histograms, and bars indicate the average performance, variance, and minimum and maximum values, respectively.

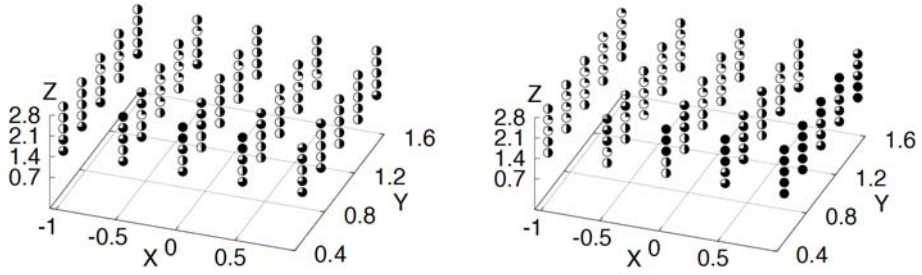
successful. Figure 5.4-b shows, for each replication, the average distance between the target and the endpoint of the arm at the end of the trials. In the figure, the bold lines, grey histograms, and bars indicate the average performance, variance, and minimum and maximum values registered, respectively.

The evolved ability also generalises to different positions of the target and to moving targets. Figure 5.5 shows the performance of evolved robots tested with the target placed in randomly selected locations (within 200 *cm* of the fixed location of the target used during the evolutionary process). As shown in Figure 5.5, the performance varied significantly in different replications of the evolutionary process. In the case of the best replication, however, the performance is only slightly worse, at 84%, with respect to the normal condition. Indeed, the average performance was still good, with a 64.1% rate of successful reaching of randomly distributed targets within the reachable space.

The performance of the best individuals was also measured on 125 target points that were evenly distributed in front of the robot on a $5 \times 5 \times 5$ grid. Figure 5.6 presents the results obtained for two individuals; all the other individuals are not shown due to their similarity to these two cases. For each target point, the individuals were tested for 5 trials, starting from different randomly assigned initial positions.

The filled area of each bullet in Figure 5.6 indicates the average distance between the target area and the endpoint of the arm in the following intervals: $< 1\,cm$ ◔, $[1, 10]\,cm$ ◑, $[10, 50]\,cm$ ◕, and $> 50\,cm$ ●; thus, the greater the degree to which a bullet is filled, the worse is the performance at that point.

The individual whose performances are shown in Figure 5.6-a behaved slightly better in the central and distant areas than in the near area. At the same time, the individual whose performances are shown in Figure 5.6-b had a close-to-optimal reaching ability in the left area, and significantly worse performance in the right area.



a) Best Individual of Seed 2 **b) Best Individual of Seed 8**

Figure 5.6: Performance obtained by testing with 125 targets points evenly distributed in front of the robot on a $5 \times 5 \times 5$ grid area. Graphs **a)** and **b)** present the results obtained 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: $< 1\,cm$ ◔, $[1, 10]\,cm$ ◑, $[10, 50]\,cm$ ◕, $> 50\,cm$ ●. The axes indicate the position of the target points along the vertical and horizontal dimensions in meters.

These qualitatively different performances can be explained by considering that the four DOFs are strongly interdependent. This clearly means that strategies that treat each joint as an independent entity (that must be moved so to reduce its distance from the target independent of the current position of the other joints) are inadequate. Evolving robots should select control strategies that minimise the problems resulting from the high interdependence among the DOFs.

Although evolving robots were selected for their ability to reach a static target, the controller generalises its ability to follow mobile targets quite well. Figure 5.7 shows the behaviour produced by one of the best evolved individuals that tries to

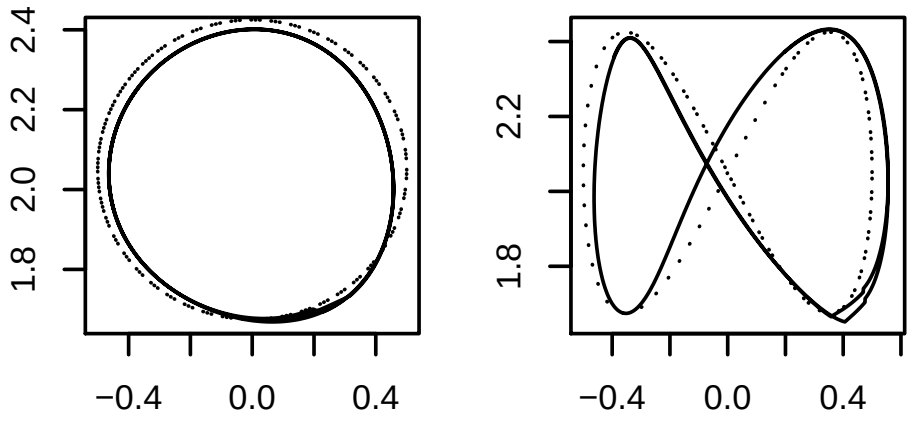
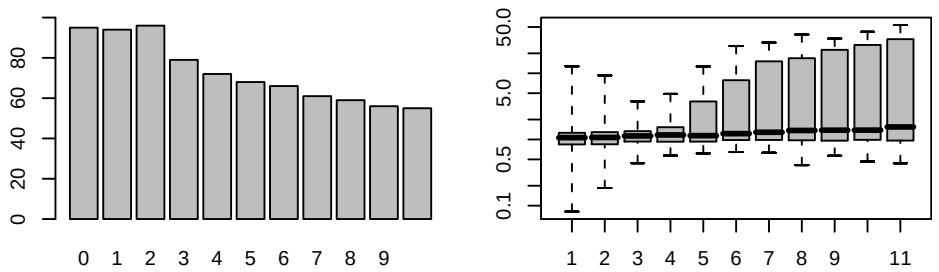


Figure 5.7: Performance on following a mobile target. Trajectory produced by the endpoint of the arm and by a moving target (solid and dotted lines, respectively). The results were obtained in two tests in which the target moved by displaying a circular and a figure-eight-shaped trajectory (**a)**, and **b)** picture, respectively). The vertical and horizontal axes indicate the respective positions of the target and of the end-point of the arm, in meters.

reach a moving target by following a circular and a figure-eight-shaped trajectory. In Figure 5.7 the dotted lines represent the trajectories of the mobile target, and the solid lines, the trajectories of the endpoint of the arm.

Furthermore, evolved agents were tested under the condition in which the updating of the sensory neurons was delayed. The performance in this situation decreased gradually as the delay increased from 60 to 150 ms. The percentages of successful trials for different lengths of delays are shown in Figure 5.8-a. The delays are expressed in multiples of 15 ms; for example, 2 indicates a delay of 30 ms with respect to the normal condition. Surprisingly, the level of performance increased with a delay of 30 ms and remained almost constant with a delay of 15 ms. Figure 5.8-b shows a box-plot of the distances between the endpoint of the arm and the target point, at the end of the trials. The median values are near 1 cm for all lengths of delay, which demonstrates that with a long delay of the sensor neurons, the best individuals performed quite well.

Finally, ten additional replications of the evolutionary process were carried out in which the update of the sensory neurons was delayed $7 \cdot 15 = 105\, ms$. The results of



a) **b)**

Figure 5.8: Performance obtained by testing robots evolved in a normal condition in a test condition in which the updating of the sensory neurons was delayed. **a)** Percentage of trials in which the distance between the endpoint of the arm and the target was less than $1\,cm$ at the end of the trial. **b)** Average distance between the endpoint of the arm and the target at the end of the trials. Each column represents the performance obtained by testing the best evolved individual in each replication for 100 trials. The bold lines, grey histograms, and bars indicate the average performance, variance, and minimum and maximum values, respectively. The x axis indicates the sensory delay (in multiples of $15\,ms$) in both graphs.

these new evolutions showed levels of performance that were quite similar to those obtained without a delay. In fact, the percentage of trials in which the distance between the endpoint of the arm and the target was less than $1\,cm$ is 91.2%, and the average distance between the target and the endpoint of the arm was $1.34\,cm$. Without sensory delay, the results are 92.1% and $1.34\,cm$, respectively. In addition, evolved robots generalise their ability to reach targets that are randomly located in the reachable space. The average percentage of successful reaching behaviours was 62.7%, and the average distance between the endpoint of the arm and the target was $6.56\,cm$. These results are similar to the performance obtained without sensory delay, which gave values of 64.1% and $9.81\,cm$, respectively.

5.4.1 Analysing Evolved Trajectories

The fitness function rewards the individual that shows rapid and precise trajectories. Once the individuals are evolved, it is possible to analyse to what degree the trajectories are good from this point of view. In fact, by taking one of the best individuals, it is possible to create a handcrafted trajectory that moves all joints at the maximum speed possible ($890\,rpm$). The procedure to generate such a handcrafted

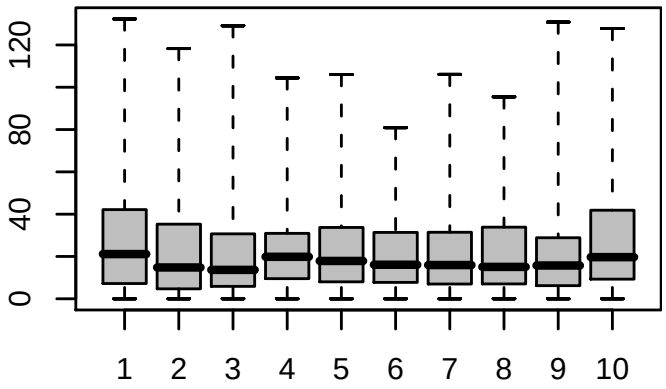


Figure 5.9: Comparison between trajectories produced by neural network and handcrafted ones. Average distance in *cm* between the trajectories produced by an evolved neural controller and the trajectories produced by manually setting the desired position of the joints on the basis of the final postures produced by the evolved neural controller. Each column indicates the results obtained for the best individual in a corresponding replication of the experiment. The bold line, grey boxes, and dotted lines indicate the average the variance, and the minimum and maximum values, respectively.

trajectory is as follows:

1. A starting posture of the arm is randomly selected and recorded.

2. The neural controller moves the arm until it reaches the target point.

3. The final posture of the arm is recorded.

4. Starting from the same starting posture as in point 1, the joints are moved at the maximum speed until they reach the same final posture as in point 3.

Figure 5.9 shows box-plots of the differences in *cm* between the trajectories produced by the ten best evolved individuals and the corresponding trajectories produced by the handcrafted procedure given above. Each box-plot in the Figure 5.9 is a summary of the data from 16 repetitions of the procedure, starting from different and random initial postures.

The fact that the differences are rather small (Figure 5.9) indicates that the trajectories produced by evolved robots are quantitatively similar to those that can be obtained by minimising the movement of the joints.

5.5 Discussions

Notwithstanding their simple architecture, evolved controllers display the ability to effectively produce a reaching behaviour. However, due to the minimal information provided by the sensory neurons, the neural controller cannot develop the ability to select different trajectories. Indeed, the global system is like a dynamical one in which the equilibrium-point is the target position (Shadmehr, 2003).

In a context in which reaching is executed without any obstacles, there is no need to generate and select different trajectories. Such problems arise, and the trajectories are determined when grasping behaviour is studied. In fact, even if there are no obstacles in the environment, robots need to reach for objects in different ways, depending upon the kind of object to be handled and the purpose of the manipulations.

In the model described, all the arm joints were actuated by a velocity-based controller, in which the desired velocity is set in the input and a PID[1] controls the stability. Muscles, however, which more closely resemble the actual structure of the human arm, provide a more stable and efficient way to actuate robotic arms. Controlling a joint via two antagonist muscles provides a number of useful features: greater stability, independent control of position and stiffness/compliance, perturbation damping, greater robustness, faster movements, and faster reactions to external perturbations (Shadmehr & Wise, 2005a,c; Buehrmann & Paolo, 2006).

The problem of controlling a robotic arm is often approached by assuming that the robot should possess, or should acquire through learning, an internal model that: (a) predicts how the arm will move and the sensations that will arise, given a specific motor command (*direct mapping*), and (b) transforms a desired sensory consequence into the motor command that will achieve this (*inverse mapping*) — for a review of this, see (Torras, 2003).

The aim of this experiment is not to deny that primates rely upon internal models of

[1] **P**roportional–**I**ntegral–**D**erivative controller, see http://en.wikipedia.org/wiki/PID_controller

this kind to control their motor behaviour. However, this does not necessarily imply that elementary movements are learned on the basis of a detailed description of the sensory-motor effects of any given motor command, or of a detailed specification of the desired sensory states. Direct and inverse mapping might operate at a higher level of organisation; it might, for example, 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 detailed direct and inverse mapping at the level of micro-actions (i.e. at the level of those elements that constitute elementary behaviours). this 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 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 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, which can typically be predicted only to a certain extent. For these reasons, the role of 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 moment.

This leaves open the question of how simple elementary behaviours 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 results of fine-grained interactions among 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 parameter variations on the agent/environment dynamics (Sternad & Schaal, 1999).

6 Reaching and Grasping

The second experiment presented in this chapter concerns the evolution of a neural network that is required to control a robot that performs the acts of reaching for and grasping objects placed on a table. The robot is a full anthropomorphic manipulator with a five-fingered hand attached to a 7-DOF arm. The actuation of the arm's joints is performed by muscle-like actuators. The fingers are controlled collectively in order to reduce the number of actuators, and consequently, the size of the neural controllers.

The obtained results demonstrate how the evolved robots manage to solve problems using solutions that are rather parsimonious, from the point of view of the robot's neural controller. The post-evaluation of the robot's performance under new conditions not experienced during the evolutionary process indicates that evolved robots generalise rather well with respect to the shape of an object, and relatively well with respect to the position of the object on a table. Overall, the obtained results demonstrate that effective reaching and grasping skills can be developed without relying upon internal models that perform direct and inverse mapping.

An analysis of the behaviour exhibited by evolved robots indicates that the chosen approach allows for the synthesis of solutions that exploit the morphological properties of the robot's body (i.e. its anthropomorphic shape, the elastic properties of its muscle-like actuators, and the compliance of its actuated joints) as well as the physical interaction between the robot and the environment, in ways that are not easy to derive using analytic methods.

6.1 The Robot

All the details of the robot's structure are presented in Appendix B. In the following sections, only a broad overview of the arm and hand structure is given.

6.1.1 Arm Structure

The arm (Figure 6.1) consists mainly of three elements (the arm, the forearm, and the wrist) that are connected through articulations that are distributed in the shoulder, the arm, the elbow, the forearm, and wrist. It is an enhancement of a previous 4-DOF model that included a wrist comprised of another 3-DOF joint. The wrist adds the ability to produce pitch, yaw and roll of the five-fingered hand. In Figure 6.1, the cylinders graphically represent the rotational dofs. The axes of the cylinders indicate the corresponding axis of rotation, and the links among the cylinders represent the rigid connections that make up the arm structure.

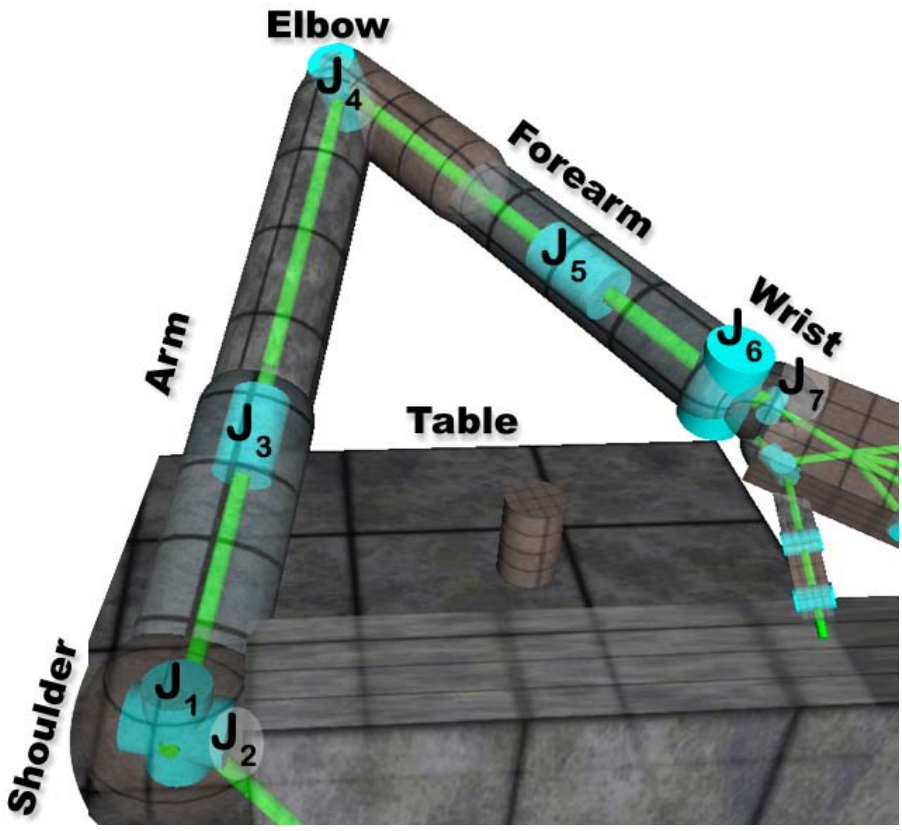


Figure 6.1: The kinematic chain of the arm. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation. The links among the cylinders represent the rigid connections that make up the arm structure.

6.1.2 Arm Actuators

The joints of the arm are actuated by two simulated antagonist muscles that are implemented accordingly to Hill's muscle model (Sandercock et al., 2003; Shadmehr & Wise, 2005b). More precisely, the total force exerted by a muscle is the sum of three forces $T_A(\alpha, x) + T_P(x) + T_V(\dot{x})$ which depend upon the activity of the corresponding motor neuron (α), the current elongation of muscle (x) and the muscle contraction/elongation speed ($\dot{x}$) which are calculated on the basis of the equations B.1 (for details see Appendix B).

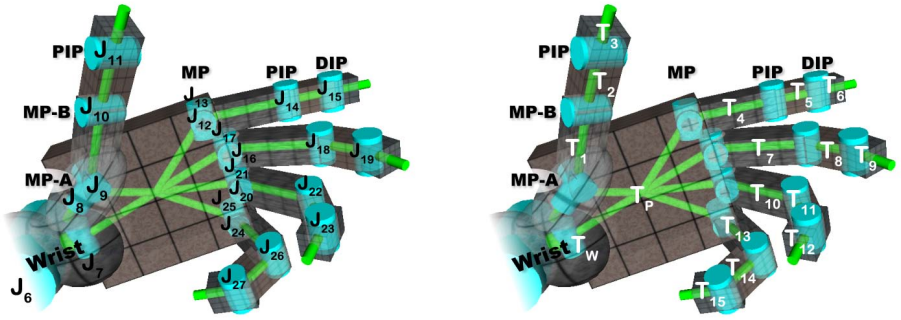
The active force T_A depends on the activation of muscle α and on the current elongation/compression of the muscle. When the muscle is completely elongated/compressed, the active force is zero regardless of the activation α. When the muscle is at its resting length, the active force reaches its maximum, which depends on the activation α.

The passive force T_P depends only on the current elongation/compression of the muscle. T_P tends to elongate the muscle when it is compressed less than it does when it is at its resting length, and also tends to compress the muscle when it is elongated beyond its resting length. T_P differs from a linear spring by virtue of its exponential trend, which produces a strong opposition to muscle elongation and little opposition to muscle compression.

T_V is the viscosity force. It produces a force proportional to the velocity of the elongation/compression of the muscle.

6.1.3 Hand Structure

The hand was added to the robotic arm just below the wrist (at joint G in Figure 6.1). The robotic hand (Figure 6.2) is composed of a palm and 14 phalange segments that make up the digits, which are connected through 15 joints, making a total of 20 DOFs (see Appendix B for details).



a) **b)**

Figure 6.2: The hand structure. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation, and the labels on the cylinders in **a)** are the names of the joints. The links among the cylinders represent the rigid connections that make up the hand structure. The white labels on the links in **b)** are the names of the tactile sensors.

6.1.4 Hand Actuators

The joints can be controlled independent of one another by specifying the desired position. One of the most important features of the hand's joints is their compliance in order to facilitate the grasping of objects. For all the details of this, see Appendix B.

6.1.5 Hand Tactile Sensors

The hand is equipped with tactile sensors that are distributed over the wrist, palm, and all five fingers. Figure 6.2-b show where the tactile sensors are placed. The white labels indicate the names of the tactile sensors. Each tactile sensor simply counts the number of contacts that take place on the corresponding part on which it is placed. The contacts that result from the humanoid touching itself are not counted. For example, in the case of T_P, it reports all contacts between the palm and another object, but not the contacts between the palm and fingers.

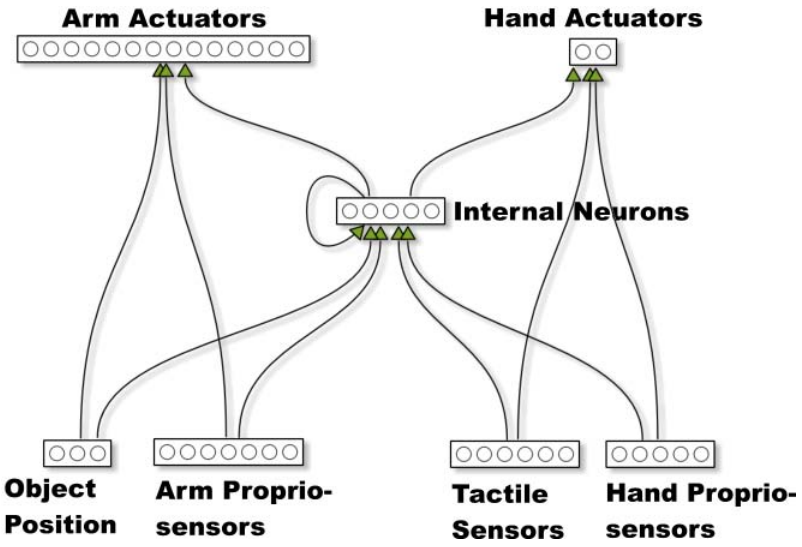


Figure 6.3: Architecture of the neural controllers. The arrows indicate blocks of fully connected neurons.

6.2 The Neural Controller

The robot is equipped with neural controllers, as shown in Figure 6.3 which include 21 sensory neurons $(x_1, \ldots x_{21})$, 5 internal neurons $(h_1, \ldots, h_5)$ with recurrent connections, and 16 motor neurons $(o_1, \ldots o_{16})$.

The neurons are divided into seven blocks in order to facilitate the description of their functionality and connectivity:

- **Object Position** (x_1, x_2, x_3): This layer can be seen as the output of a vision system (which has not been simulated) that computes the relative distance of the object with respect to the hand up to a distance of $80\,cm$ normalised in the range of $[-1, +1]$ over three orthogonal axes.

- **Arm Propriosensors** $(x_4, \ldots, x_{10})$: These encode the current angles of the 7 corresponding DOFs located on the arm and on the wrist normalised in the range of $[-1, +1]$.

- **Tactile Sensors** $(x_{11}, \ldots, x_{16})$: These measure whether the 5 fingers and the unit constituted by the palm and wrist are in physical contact with another object. More precisely, the output of each tactile sensor is calculated with the

following equation:

$$\begin{aligned} x_{11}(t) &= \delta_{11}\left(\sigma_{0.2}\left(T_P + T_W\right)\right) + \left(1 - \delta_{11}\right) x_{11}(t-1) \\ x_i(t) &= \delta_i \left(\sigma_{0.2}\left(\sum_{s=1}^{3} T_{s+3(i-12)}\right)\right) + \left(1 - \delta_i\right) x_i(t-1) \;\; for\, i = 12, \ldots, 16 \end{aligned}$$

where T_i is the value of the corresponding tactile sensor as described above. $\sigma_\lambda(x) = (1 + e^{-x})^{-\lambda}$, and δ_i is a coefficient that range in the interval $[0, 1]$. The value of δ_i represents the dependence of the output of the tactile sensors on the previous one. These equations are similar to those for leaky integrators (Nolfi & Marocco, 2001; Beer, 1995), and the idea behind them is quite simple: the total activation is the sum of the δ_i values for all the tactile sensors on a finger and the $1 - \delta_i$ values for the tactile sensations of the previous step. In this way, when there are no further contacts, the activation of the neuron leaks a small amount over time, starting from the value of the previous step, instead of going to zero instantaneously as do normal sensory neurons. The time needed to reach zero is proportional to the δ_i value, and for this reason, δ_i is considered a time constant.

- **Hand Propriosensors** $(x_{17}, \ldots, x_{21})$: These encode the current *extension/flexion* state of the five corresponding fingers in the range of $[0, 1]$ where 0 means fully extended and 1 means fully flexed. More precisely, the output of each hand's propriosensors is calculated with the following equation:

$$\begin{aligned} x_{17} &= map_{[0,1]}\left(\frac{ang\left(J_{10}\right) + ang\left(J_{11}\right)}{2}\right) \\ x_i &= map_{[0,1]}\left(\frac{\sum_{s=13}^{15} ang\left(J_{s+3(i-18)}\right)}{3}\right) \;\; for\, i = 18, \ldots, 21 \end{aligned}$$

where $ang\left(J_i\right)$ is the angular position expressed in $radians$ of the joint J_i, and $map_{[0,1]}$ maps the possible angle values of joints into the interval $[0, 1]$. Due

to the compliance of the finger's joints when the hand collides hits with an object and due to the fact that the state of the three corresponding DOFs is summarised in a single variable, the same sensory state might correspond to a different states of the MP and DIP joints.

- **Internal Neurons** ($h_1, \ldots h_5$): Each internal neuron has a bias

- **Arm Actuators** ($o_1, \ldots, o_{14}$): The values of o_i directly indicate the state of activation of the 14 motor neurons that control the corresponding muscles of the arm.

- **Hand Actuators** (o_{15}, o_{16}): The value of o_{15} is the desired *extension/flexion* angle of the thumb; where 0 means fully extended and 1 means fully flexed. Also the value of o_{16} is the desired *extension/flexion* angle of all the other four fingers. It is important to note that in this experiment setup, not all the DOFs of the fingers are controlled by the neural network. In fact, the positions of the joints are controlled by a limited number of variables via a velocity-proportional controller (with the maximum joint velocity set to $0.30\,rad/s$). More precisely, the force exerted by the MP, PIP and DIP joints (MP-A, MP-B, and PIP in the case of the thumb), which determine the extension/flexion of the corresponding finger, are controlled by a single variable θ that ranges from $[-90°, +0°]$. The desired position of the three joints was set to θ, θ and $2/3 \cdot \theta$, respectively. In the case of the thumb, the supination/pronation is also controlled by θ, by setting the desired angle to $-2/3 \cdot \theta$. The DOF that determines the abduction/adduction of the first phalanx of each finger is controlled by a second variable, which was set to a constant value of $0\,rad$ for this experiment. This simplification of the control of the fingers is justified by the fact humans also have a very limited control of each phalanx of fingers during power grasps. Except for very fine movements, humans make use of all 3 DOFs of their fingers at the same time, while maintaining the fingers in a natural posture (Jones & Lederman, 2006; Page, 1998). Hence, the activation of the

motor neurons mapped into the θ variable is then mapped into a natural posture that corresponds to the natural constraints of human fingers (Yasumuro et al., 1999).

The state of the sensors, the desired state of the actuators, and the internal neurons were updated every $10\,ms$, using the following equation:

$$h_i = \sum_{j=1}^{21} w_{ji}\sigma_{0.5}\left(x_j\right) + \beta_i$$

$$o_i = \sum_{j=1}^{10} w_{ji}\sigma_{1.0}\left(x_j\right) + \sum_{j=1}^{5} w_{ji}\sigma_{1.0}\left(h_i\right)\ for\, i = 1, \ldots, 14$$

$$o_i\left(t\right) = \delta_i \left(\sum_{j=11}^{21} w_{ji}\sigma_{1.0}\left(x_j\right) + \sum_{j=1}^{5} w_{ji}\sigma_{1.0}\left(h_i\right) + \beta_i \right) + \left(1 - \delta_i\right) o_i\left(t-1\right)\ for\, i = 15, 16$$

where x_i are the output of the sensory neurons as described above. h_i and o_i are the output of the internal and actuator neurons, respectively. w_{ji} is the synaptic weight from neuron j to neuron i. β_i is the bias of the i-th neuron, and δ_i is the coefficient for implementing leaky neurons as proposed in (Nolfi & Marocco, 2001) for the hand actuators.

This particular sensory system configuration was chosen in order to be able to study situations in which the vision and tactile sensory channels need to be integrated. In isolation, each of the two types of sensors does not provide enough information to perform the task.

The proprioceptive neurons, in addition to the sensory neurons described in the previous experiment in Section 5.1, play an essential role. In the case in which the robot arm almost reaches the object, the lack of these proprioceptive neurons results in the sensory neurons shifting to zero, and furthermore, due to the perceptron architecture, the output neurons also tend to zero. This behaviour would not permit the evolution of the neural network for muscle-actuated joints, because in order to maintain a configuration, the muscle's activation must stay at a value other than

zero. In fact, with muscle-actuated joints, the arm postures are encoded with different activation of antagonistic muscles. and when the activation tends to zero, the arm tends to assume a position of rest, depending upon the muscle's passive/length properties.

This neural control model is a step forward in control systems that is consistent with the equilibrium point hypothesis (Shadmehr, 2003). The governing concept is that the controller's inputs are parameters of the dynamical system represented by the controller that modifies its equilibrium points. Hence, the inputs of the neural network that are dedicated to reaching do not encode an output motor sequence, but rather, an equilibrium point of the system that leads to the correct posture that is required in order to reach the desired point. In more detail, the neural network produces a dynamical system that depends upon the incoming inputs, and the dynamics of the system produce a behaviour that ends up in a configuration at the equilibrium at which the robot grasps the object.

This approach offers the advantage of having a high robustness to perturbation. because if some variation or perturbation occurs, the equilibrium point will not change and the dynamical system will tend to the same final position, regardless.

6.3 The Evolutionary Process

The free parameters of the neural controller, i.e. the connection weights (w_{ji}), the biases (β_i) of the internal neurons and hand actuators, and the time constant (δ_i) of the leaky-integrator neurons, were adapted using an evolutionary robotics method (Nolfi & Floreano, 2000).

The initial population consisted of 100 randomly generated genotypes, which encode the free parameters of 100 corresponding neural controllers. Each parameter was encoded with 16 bits. Each genotype contained 6,096 bits corresponding to 381 free parameters: 366 connection weights, 7 biases normalised in the range of $[-10, +10]$

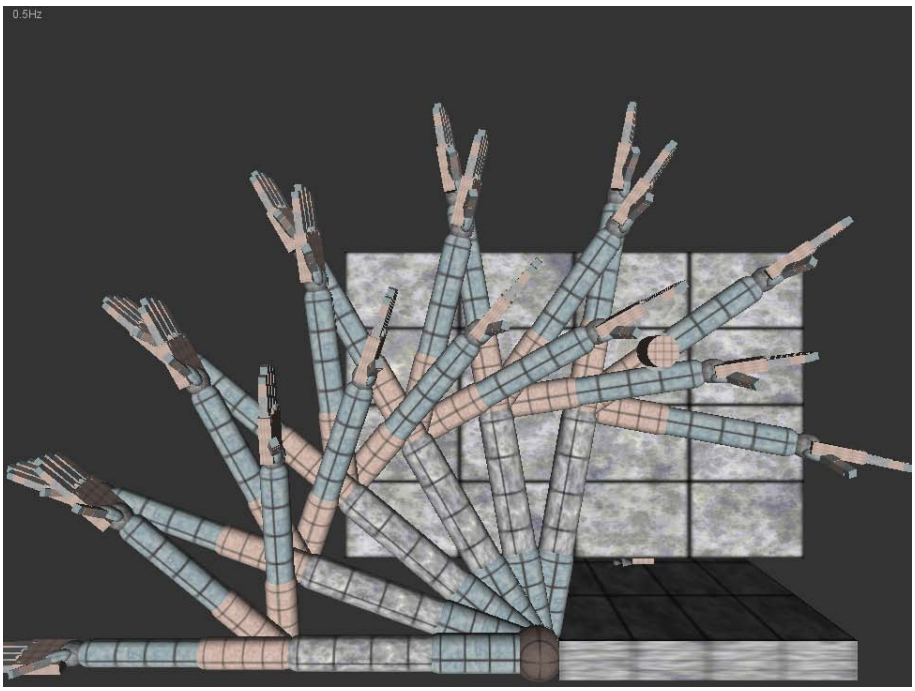


Figure 6.4: The 18 predefined initial postures of the arm. The postures were obtained by systematically setting the elbow joint at three different angles, and one joint of the shoulder in six different positions; all others joints were set to 0.

and 8 time constants normalised in the range $[0.0, 1.0]$.

The 20 best genotypes of each generation were allowed to reproduce by generating five copies each. Four out of the five copies were subjected to mutations and one copy was left intact. During mutation, each bit of the genotype had a 1.5% probability of being replaced by a new, randomly selected value. The evolutionary process was continued for 400 generations (i.e. the process of testing, selecting and reproducing robots was iterated 400 times).

The experiment was replicated 10 times. The robot was adapted in order to possess the ability to grasp spherical and cylindrical objects on a table that was placed in front of it. The objects could move freely and could even fall off the table (Figure 6.1). During the adaptive process, each genotype was translated into a corresponding neural controller, which was embodied in the simulated robot and tested for 18 trials. Each trial lasted $4\,s$, which corresponds to 400 steps. At the beginning of each trial, the arm was set in the i-th posture of the 18 corresponding predefined postures shown in Figure 6.4. The target object was placed in a fixed position in the central

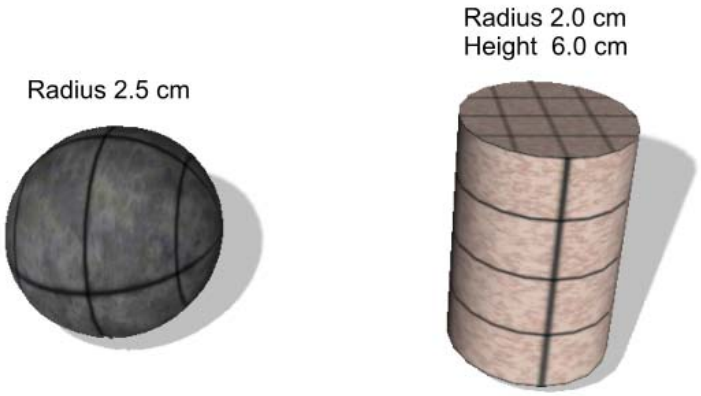


Figure 6.5: The two objects to be grasped. The sphere has a radius of 2.5 cm and a weight of 32.72 g; the cylinder has a radius of 2.0 cm, a height of 6.0 cm. and a weight of 37.70 g.

area of the table-top. The spherical objects had a radius of 2.5 cm and a weight of 32.72 g, and the cylindrical objects had a radius of 2.0 cm, a height of 6.0 cm, and a weight of 37.70 g (see Figure 6.5).

The evolving robots were evaluated on the basis of the following fitness function, which rewards successful reaching and grasping behaviours:

$$F = \frac{1}{1803600} \sum_{t=1}^{18} \sum_{s=200}^{400} \left(\frac{1}{1 + 0.25 \cdot dist} + 500 \cdot grasp \right)$$

where $dist$ encodes the distance between the barycentre of the hand and the object. The term $grasp$ encodes whether an object has been successfully grasped (i.e. $grasp$ is 1 when the target object is elevated with respect to the table and is in physical contact with the robot hand, and is 0 otherwise). t is the current trial, and s is the current time step. To allow the robot to reach for and grasp the object, the fitness is calculated only in the second-half of each trial (i.e. from time step 200 to time step 400). The constant at the beginning of the function, which corresponds to the maximum fitness that can be gathered by grasping each object during the first phase of each trial and by holding the object above the plane for the rest of the trial, was used to normalise the fitness value in the range of $[0, 1]$.

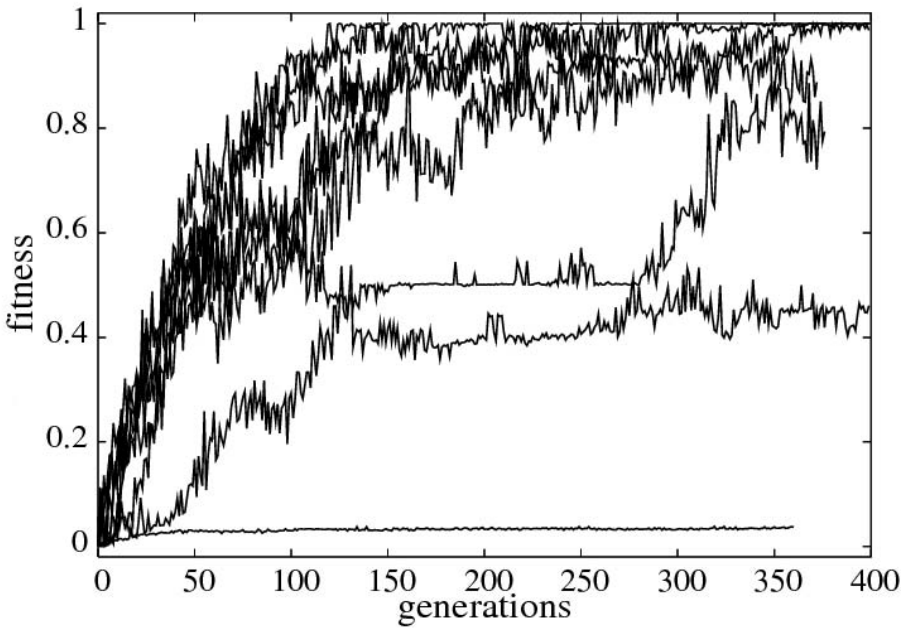


Figure 6.6: Fitness of the best individuals throughout generations for the 10 replications of the experiment.

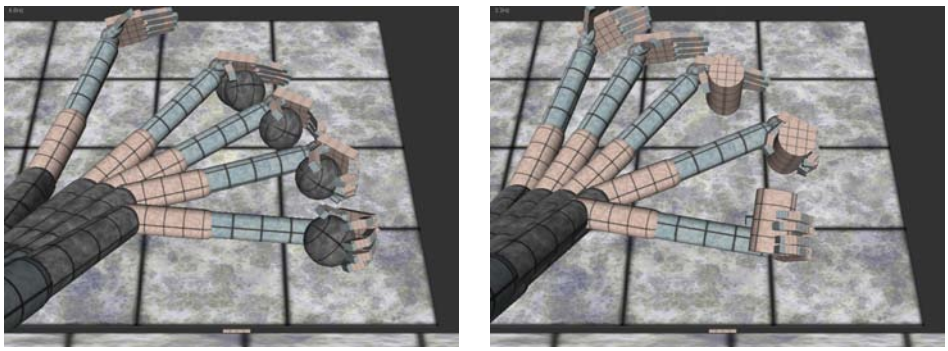
6.4 Results

By analysing the behaviour of the evolved robots throughout multiple generations, 8 out of 10 replications of the experiment developed robots with the ability to reach and grasp objects. As can be clearly seen in Figure 6.6, the best evolved robots displayed close to optimal performance.

The best individual from one of the most successful replications successfully grasped the two types of objects using all of the 18 initial postures shown in Figure 6.4. As shown in Figure 6.7, the behaviour displayed by this individual can be divided into three phases:

1. An initial phase in which the arm moves towards the object with increasing speed. When the hand is near the object, the robot begins to slow down the speed of the arm and initiates flexion of the hand.

2. A second phase, in which the tactile sensors begin to be activated. The arm stays almost still. The robot flexes the fingers and the wrist to encircle the object.



a) Robot grasping a sphere **b) Robot grasping a cylinder**

Figure 6.7: Snapshots of the grasping behaviour. Five superimposed snapshots of the grasping behaviour displayed by one of the best evolved robots.

3. A final phase, in which the arm and the wrist rotate so that the palm is face-up. The robot moves the arm to lift the object from the table. The rotation of the palm reduces the risk of the object falling from the hand while it is being lifted.

A set of videos showing the behaviour of evolved robots can be accessed on the web page http://laral.istc.cnr.it/esm/arm-grasping/.

The best individuals displayed remarkable generalisation abilities when tested in conditions that were different from those experienced during evolution. As regards the position of the object, Figure 6.8 shows the average performance of the best evolved robots from three of the best replications in the experiment, in which the positions of the objects on the table were systematically varied. Each robot was tested in 120 different conditions corresponding to 60 different positions of the object on the table, and to the two types of objects (spherical and cylindrical objects). For each testing condition, the robot was tested for 18 trials corresponding to the 18 different starting positions of the arm. In the figure, the colours of the rectangles indicate the average performance for the corresponding location. In each picture, the left and right areas correspond to the left and right areas of the table with respect to the robot. The top and bottom areas correspond to the proximal and distant areas of the table with respect to the robot.

Although different individuals varied with respect to their generalisation capabilities,

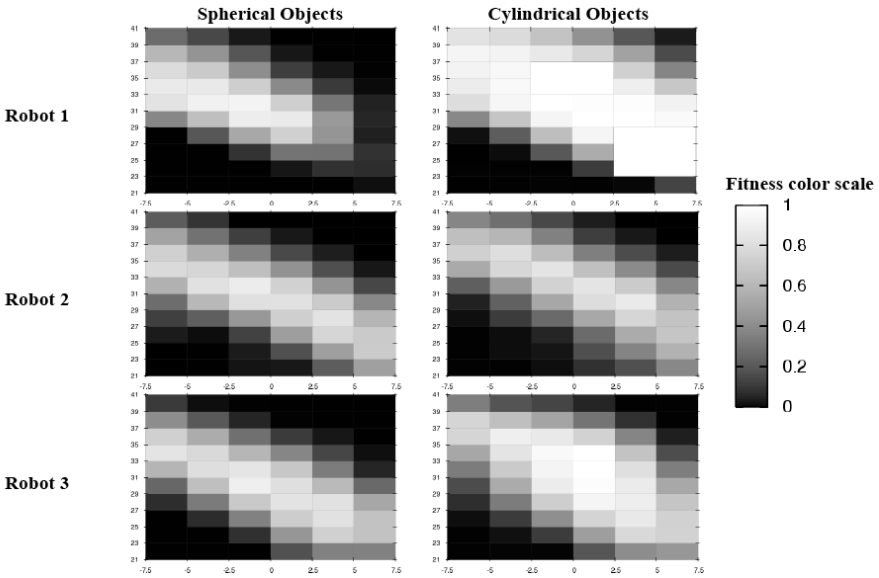


Figure 6.8: Performance of the best evolved robots from the three best replications of the experiment. The coloured areas in the map graphs represent the average performance of the robot, indicated by the row, upon grasping the object indicated by the column. The average represents over 18 trials corresponding to the 18 different starting positions of the arm. For each map graph, the left and right positions correspond to the left and right areas of the table with respect to the robot, and the top and bottom positions correspond to the proximal and distant areas of the table with respect to the robot.

they all displayed rather good performance on the central diagonal area, which corresponds to the preferential trajectory followed by the arm in normal conditions (i.e. when the objects were placed in the central area of the table-top). The decrease in performance on the top-right and bottom-left parts of the table can be explained by considering that the grasping of objects located in these areas requires postures that differ significantly from those that the robots assume to grasp objects in the central area of the table-top.

The best individuals also displayed a remarkable ability to grasp objects that differ in shape and size, and that are placed in locations different from those experienced during evolution. The objects used in these tests are shown in Figure 6.9.

The results of these tests are summarised in Figure 6.10. In the figure, the bars represent the average performance over all 60 different positions for all 18 initial postures of the arm, as described for the previous test. Figure 6.11 reports the

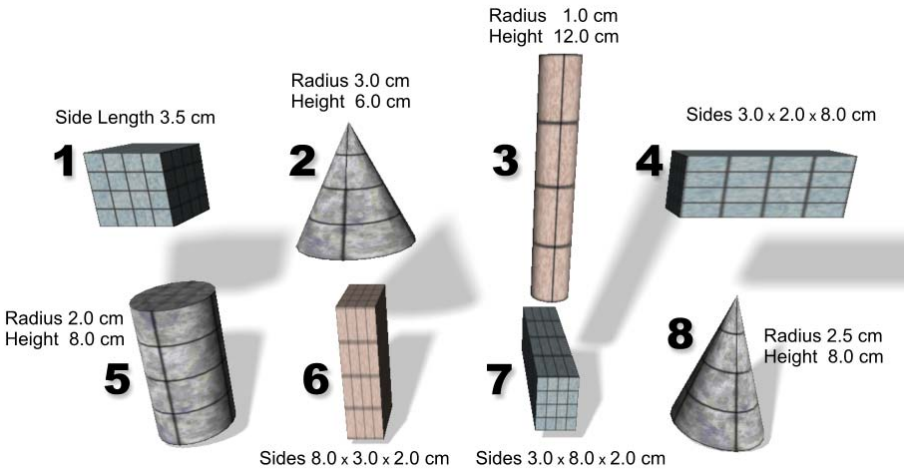


Figure 6.9: Objects used for testing robots' generalisation ability with respect to object shape and size. The dimensions of the objects are specified in the figure, and the bold numbers on the left identify the objects as referenced in Figures 6.10 and 6.11.

average performance for each position of an object for each of the best individuals.

The differences in performance among the individual robots from different replications of the experiment are due to the different behavioural strategies displayed by evolved individuals, with particular reference to the second and third phases of their behaviour in which the robots grasp and lift the objects (for more information, see the video available from the web page http://laral.istc.cnr.it/esm/arm-grasping/).

For example, the fact that the best individual from replication 1 displayed poor performance with objects 2, 4, 6, and 7 as compared to the other evolved individuals is due to the fact that it flexes its fingers very quickly. This type of strategy actually denies this robot the possibility of exploiting the adjustment of the relative position of the fingers with respect to the objects, which arises spontaneously in time as a result of the effects of the forces exerted by the hand, collisions between the fingers and the object, and the compliance of the hand.

The poor performance of the best individuals from replication 7 on objects 2, 3 and 4 can be explained by considering that the way in which this individual lifts objects after the grasping phase tends to produce collisions with the plane in the case of large objects, which might cause the object to fall from the hand.

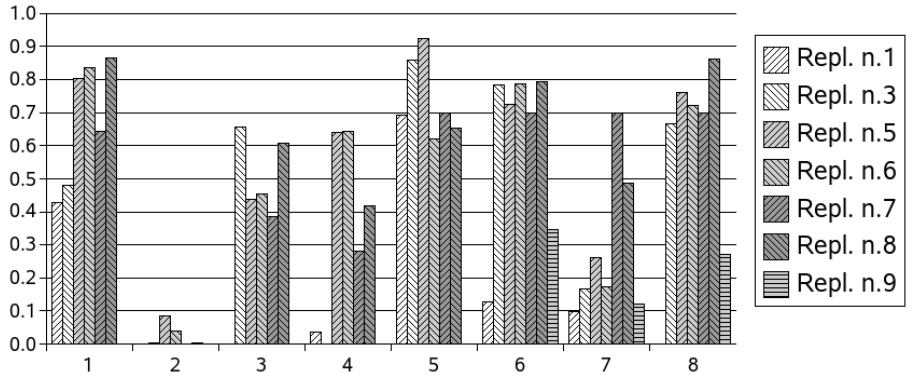


Figure 6.10: Performance for grasping eight different objects. Performance of the evolved robots from the seven best replications of the experiment, as observed by testing them with the eight objects shown in figure 6.9. The bars represent the average performance over all 60 different positions for all 18 initial postures of the arm. The positions are located as shown in the map graphs in Figure 6.8.

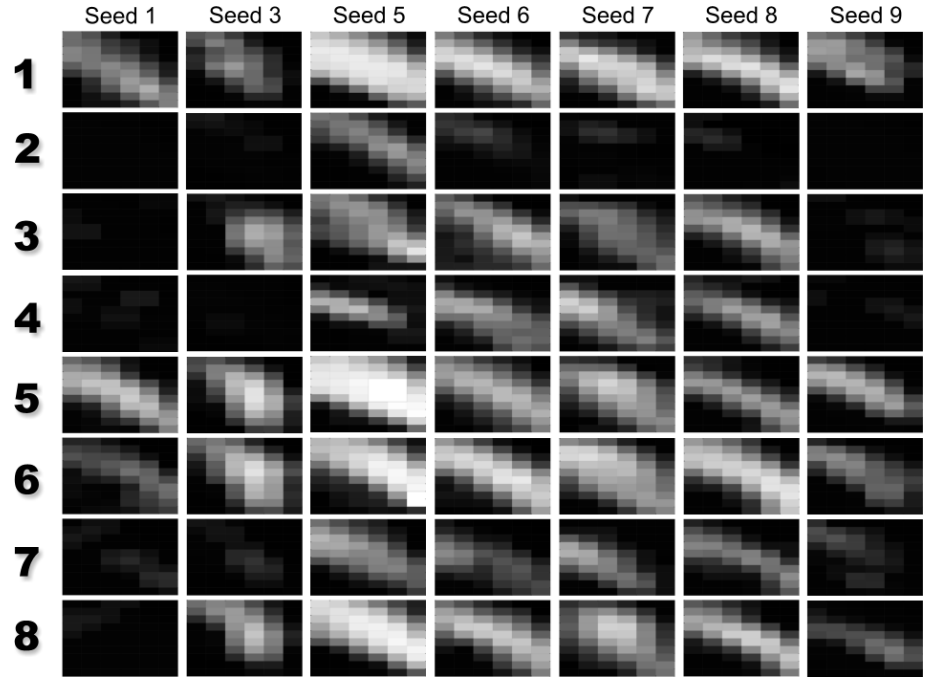


Figure 6.11: Performances for grasping eight different objects. The coloured areas in the map graphs represent the average performance of the best robot from the replication, as indicated by the column, upon grasping the object indicated by the row. The number of the row identifies the object, as in Figure 6.9. The average represents 18 trials corresponding to the 18 different starting positions of the arm. For each map graph, the left and right positions correspond to the left and right areas of the table with respect to the robot, and the top and bottom positions correspond to the proximal and distant areas of the table with respect to the robot.

Finally, the good performance of replication 8 can be explained by this robot's ability to control the thumb, which is crucial for grasping difficult, slippery objects. Also, this robot produces little rotation of the arm and wrist during the lifting phase, which minimises the risk of collisions with the plane after the objects have been grasped.

Overall, these results suggest that certain behavioural strategies might be effective for a large variety of objects, and that the limited differences in the shapes and sizes of the objects to be grasped does not necessarily have an impact on the rules that regulate robot/environmental interactions.

An important role that contributes to the generalisation ability is played by the muscle-like properties of the actuators of the arm and by the compliance of the actuators of the fingers, both of which are exploited in the evolution process. In fact, the compliance of the fingers simplifies the problem of adapting the postures of the fingers to the shape of the object. Another important role that contributes to the ability to grasp an object in different positions is the choice to encode the object position extracted by the vision system with respect to the hand position, instead of to the fixed frame of the robot.

6.5 Discussion

The work presented in this chapter shows how effective reaching and grasping behaviours exhibited by an anthropomorphic robotic arm can be developed through a process of evolution. Evolution is like a trial–and-error process in which the variants of the free parameters are retained or discarded on the basis of their effects upon the level of global behaviour. However, the free parameters encode the control rules that regulate the fine-grained interaction between the robot and the environment. Hence, the robots are left free to choose the way in which the problem is solved during the adaptation process, since they are rewarded only with respect to their ability to approach and lift objects. The particular trajectory used to approach

objects, the postures of the arm and hand, and the ways in which different motor actions produced by the robot interact with the environment, are all irrelevant from the point of view of the fitness function employed for rewarding the robots.

The experimental setup presented is significantly more advanced than that of previous works based on similar adaptive techniques (Bianco & Nolfi, 2004; Buehrmann & Paolo, 2004; Gomez et al., 2005; Massera et al., 2006; Bongard, 2010). The morphology of the anthropomorphic arm and hand with 27 DOFs is rather more complex than the arm models cited. Hence, the size of the neural controller and the dimensions of the corresponding search space are greater. Also, the task involves the ability to reach for and grasp freely moving objects with different shapes placed on a table.

The obtained results demonstrate how the proposed methodology and the exploitation of the properties that arise from the physical interaction between the robot and the environment allow effective behaviours to be produced on the basis of a parsimonious control system. For example, the effects of the collisions between the fingers of the robotic hand and the objects being grasped, combined with the compliance of all the finger joints, enable the robot hand to spontaneously conform to the shape of an object, which in turn allows the robot to effectively grasp objects with different shapes and orientations without the need for control mechanisms to regulate the movement of the arm and hand on the basis of the characteristics of the objects.

This line of research is also consistent with recent cognitive robotics approaches, such as those in the field of developmental robotics (Lungarella & Metta, 2003). Developmental robotics, also known as epigenetic robotics, is an interdisciplinary approach to robot design. Developmental robots are characterised by a prolonged develop- mental process in which varied and complex cognitive and perceptual structures emerge as a result of the interaction of an embodied system with a physical and social environment. Lungarella & Metta (2003) show that although most current investigations of developmental robotics have focused on sensorimotor control

(e.g. reaching) and social interaction (e.g. gaze control), future cognitive robotics research needs to go beyond the limited sphere of behaviours such as these. In order to design truly autonomous behaviour, future robotics research needs to integrate motor control with improved sensory and motor apparatus, more refined value-based learning mechanisms, and means of exploiting neural and body dynamics.

This approach also has a potential relevance to computational neuroscientific research on motor control (Shadmehr & Wise, 2005b). The current architecture of the robot's neural controller has not been restricted to any specific brain region known to be involved in limb control. Therefore, the current model and simulation results cannot be used to speculate upon its relevance to neuroscientific research. However, the development of future extensions of the model might specifically focus on investigating the role of the structure of the neural network controller and its mapping onto brain regions and circuits (e.g. the cerebellum, motor areas) that are known to be involved in prehension ability (Jones & Lederman, 2006; Kawato, 2003). This would also make it possible to test current theories of minimisation criteria, such as energy minimum, jerk minimum, and stability maximisation for the generation of voluntary movements, and a comparison between robotic model results and the results in the literature of limb neurophysiology (Shadmehr, 2003).

7 Manipulation and Object Discrimination

The experiment in this chapter investigates the perceptual skills of an anthropomorphic robotic arm with a five-fingered hand controlled by an artificial neural network that is given the task of actively categorising un-anchored spherical and ellipsoid objects placed in different positions and orientations over a planar surface. The task requires that the agent produce different categorisation outputs for objects with different shapes, and similar categorisation outputs for objects with the same shape.

The aim of this study is to prove that, in spite of the complexity of the experimental scenario, the ER approach can be successfully employed to design neural mechanisms that allow the robotic arm to perform such a perceptual categorisation task. Indeed, the best individuals synthesised by artificial evolution techniques develop a close-to-optimal ability to discriminate among the shapes of objects, as well as an ability to generalise their skill in new circumstances. Moreover, specific analysis was carried out on the best neural controllers in order to discover:

- how the robot acts in order to bring forth the sensory stimuli that provide the regularities necessary for categorising objects, in spite of the fact that sensations themselves may be extremely ambiguous, incomplete, partial, and noisy;
- the dynamical nature of sensory flow (i.e., how sensory stimulation varies over time and the time rate at which significant variations occur);
- the dynamical nature of the categorisation process (i.e., whether the categorisation process occurs over time, as the robot interacts with the environment); and

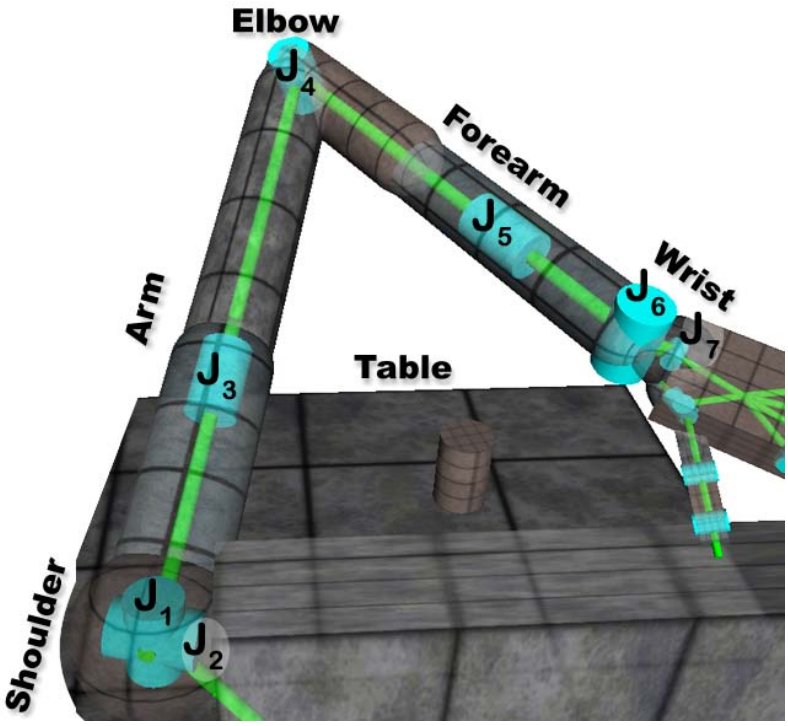


Figure 7.1: The kinematic chain of the arm. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation. The links among the cylinders represent the rigid connections that make up the arm structure.

- the role of qualitatively different sensations originating from different sensory channels in the accomplishment of the categorisation task.

7.1 The Robot

The robot that is the subject of the experiment presented in this chapter is the same as the one used in the previous experiment, and all the details of its structure are given in Appendix B. In the following sections, only a broad view of the arm and hand structure is given.

7.1.1 Arm Structure

The arm (Figure 7.1) consists mainly of three elements (the arm, the forearm and the wrist) that are connected through articulations distributed in the shoulder, the

arm, the elbow, the forearm and wrist. It is an enhancement of a previous 4-DOF model to which has been added a wrist comprised of another 3-DOF joint. The wrist adds the ability to produce pitch, yaw and roll of the five-fingered hand. In Figure 7.1, the cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation, and the links among the cylinders represent the rigid connections that make up the arm structure.

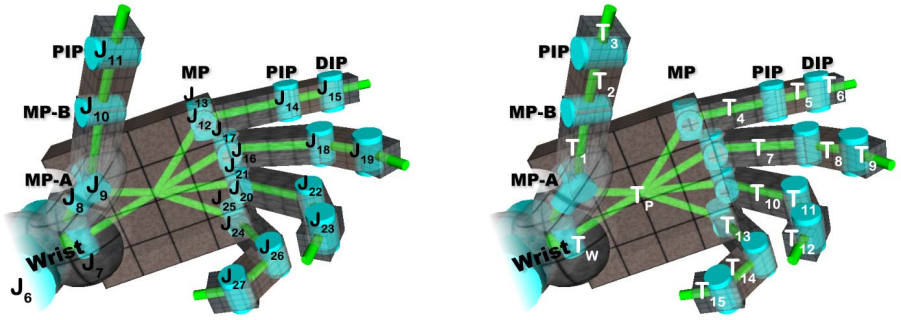
7.1.2 Arm Actuators

The joints of the arm are actuated by two simulated antagonist muscles that are implemented accordingly to Hill's muscle model (Sandercock et al., 2003; Shadmehr & Wise, 2005b). More precisely, the total force exerted by a muscle is the sum of three forces $T_A(\alpha, x) + T_P(x) + T_V(\dot{x})$ which depend on the activity of the corresponding motor neuron (α), the current elongation of the muscle (x) and the muscle contraction/elongation speed ($\dot{x}$), which are calculated on the basis of the equations B.1 (for details see the appendix B).

The active force T_A depends upon the activation of muscle α and on the current elongation/compression of the muscle. When the muscle is completely elongated/compressed, the active force is zero, regardless of the activation α. At the resting length of the muscle, the active force reaches its maximum, which depends upon the activation α.

The passive force T_P depends only on the current elongation/compression of the muscle. T_P tends to elongate the muscle when it is compressedto less than its resting length, and tends to compress the muscle when it is elongated beyond its resting length. T_P differs from a linear spring in that its exponential trend produces a strong opposition to muscle elongation and little opposition to muscle compression.

T_V is the viscosity force. It produces a force that is proportional to the velocity of the elongation/compression of the muscle.



a) b)

Figure 7.2: The hand structure. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation, and the labels on the cylinders in **a)** are the names of the joints. The links among the cylinders represent the rigid connections that make up the hand structure. The white labels on the links in **b)** are the names of the tactile sensors.

7.1.3 Hand Structure

The hand was added to the robotic arm just below the wrist (at joint G, shown in Figure 7.1). The robotic hand (Figure 7.2) is composed of a palm and 14 phalange segments that make up the digits, which are connected through 15 joints, making a total of 20 DOFs (see Appendix B for details).

7.1.4 Hand Actuators

The joints can be controlled independent of one another by specifying the desired position. One of the most important features of the hand's joints is their compliance in order to facilitate the grasping of objects. For all details, see Appendix B.

7.1.5 Hand Tactile Sensors

The hand is provided with tactile sensors that are distributed over the wrist, the palm, and all five fingers. Figure 7.2-b shows where the tactile sensors are placed. The white labels indicate the names of the tactile sensors. Each tactile sensor simply counts the number of contacts that take place on the corresponding part it is placed on. Contacts made by the humanoid parts are not counted. For example, in the case

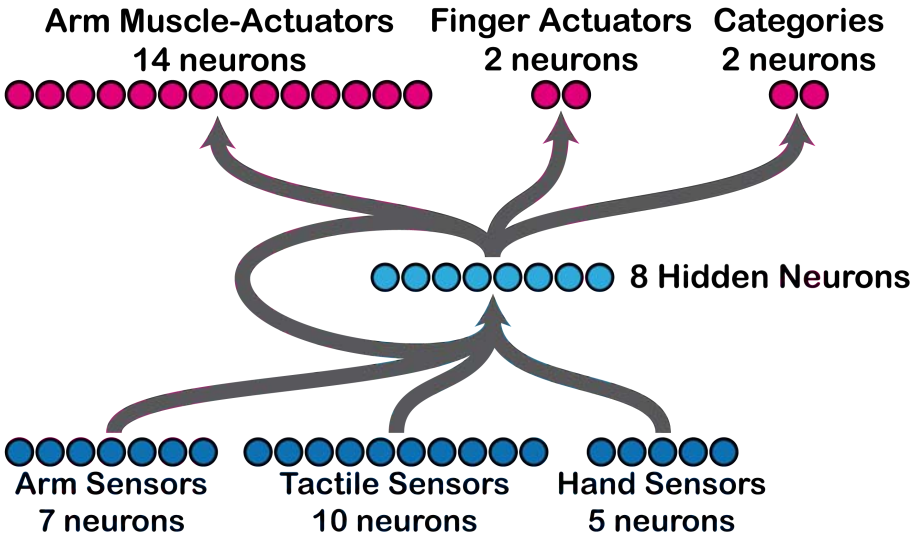


Figure 7.3: The architecture of the neural controllers. The arrows indicate blocks of fully connected neurons.

of T_P, it reports all contacts between the palm and other objects, but not contacts between the palm and the fingers.

7.2 The Neural Controller

The robot is equipped with the neural controllers shown in Figure 7.3 which include 22 sensory neurons $(x_1, \ldots x_{22})$, 8 internal neurons $(h_1, \ldots, h_8)$ with recurrent connections and 18 motor neurons $(o_1, \ldots o_{18})$. The neurons are divided into the following seven blocks in order to facilitate the description of their functionality and connectivity:

- **Arm Propriosensors** $(x_1, \ldots, x_7)$: The activation values x_i of the arm's propriosensory neurons encodes the current angles of the 7 corresponding DOFs located on the arm and the wrist normalised in the range of $[-1, 1]$.

- **Tactile Sensors** $(x_8, \ldots, x_{17})$: The activation values x_i of the tactile sensory neurons are updated on the basis of the state of the tactile sensors distributed over the hand. Each tactile sensor is associated with only one of all available sensors. Hence, not all tactile sensors shown in Figure 7.2-b are used. The

tactile sensors used are T_P, T_3, T_4, T_6, T_7, T_9, T_{10}, T_{12}, T_{13} and T_{15}, which are associated, respectively, to neurons $x_8, \ldots, x_{17}$. The activation of x_i is 1 if the corresponding tactile sensor T reports any contacts, and 0 if there are no contacts.

- **Hand Propriosensors** ($x_{18,} \ldots, x_{22}$): The activation values x_i of the hand's sensory neurons encodes the current extension/flexion of the 5 corresponding finger's joints (see joints J_8, J_9, J_{10}, J_{11}, and J_{12} in Figure 7.2-a), normalised in the range of $[0, 1]$ (with 0 for a fully extended, and 1 for a fully flexed joint).

- **Internal Neurons** ($h_1, \ldots, h_8$): Each internal neuron has a bias.

- **Arm Actuators** ($o_1, \ldots, o_{14}$): The firing rate $\sigma_{1.0}(o_i+\beta_i)$ of the motor neurons determines the state of the simulated muscles of the arm (see eq. 7.1).

- **Hand Actuators** (o_{15}, o_{16}): the firing rate $\sigma_{1.0}(o_{15} + \beta_{15})$ is the desired *extension/flexion* angle of the thumb, where 0 means fully extended and 1 means fully flexed. Further, the firing rate $\sigma_{1.0}(o_{16} + \beta_{16})$ is the desired *extension/flexion* angle of the other four fingers. It is important to note that in this experiment setup not all the DOFs of the fingers are controlled by the neural network. In fact, the positions of the joints are controlled by a limited number of variables through a velocity-proportional controller (the joints' maximum velocity is set to $0.30\,rad/s$). More precisely, the force exerted by the MP, PIP and DIP joints (MP-A, MP-B and PIP in the case of the thumb), which determines the *extension/flexion* of the corresponding fingers, are controlled by a single variable θ that ranges between $[-90°, +0°]$. The desired position of the three joints is set to θ, θ and $2/3 \cdot \theta$, respectively. In the case of the thumb, the *supination/pronation* is also controlled by θ by setting the desired angle to $-2/3 \cdot \theta$. The DOF that govern the *abduction/adduction* of the first phalanx of each finger is controlled by a second variable, which was set to a constant value of $0\,rad$ in this experiment.

  This simplification of the control of the fingers is justified by the fact humans

also have a very limited control of each phalanx of fingers during power grasps. Except when making very fine movements, humans make use of all 3 DOFs of the fingers at the same time, while maintaining the fingers in a natural posture (Jones & Lederman, 2006; Page, 1998). Hence, the activation of motor neurons mapped into the θ variable is then mapped into a natural posture in accordance with the constraints of human fingers (Yasumuro et al., 1999).

- **Categories** (o_{17}, o_{18}): Their firing rates are used to categorise the shape of the object; i.e. to produce different output patterns for different object types.

The internal neurons are fully connected. In addition, each internal neuron receives one incoming synapse from each sensory neuron. Each motor neuron receives one incoming synapse from each internal neuron. There are no direct connections between the sensory and motor neurons.

To take into account the fact that sensors are noisy, tactile sensors x_i return, with a 5% probability, a value different from the computed value, and 5% uniform noise was added to proprioceptive sensors x_i.

The values of the neurons were updated using the following equation:

$$\begin{aligned} 0.01 \cdot \dot{S}_i &= -S_i + g \cdot x_i \\ \tau_i \dot{h}_i &= -h_i + \sum_{j=1}^{22} w_{ji} \sigma_{1.0} \left(S_j + \beta_j\right) + \sum_{j=1}^{8} w_{ji} \sigma_{1.0} \left(h_j + \beta_j\right) \\ 0.01 \cdot \dot{o}_i &= -o_i + \sum_{j=1}^{8} w_{ji} \sigma_{1.0} \left(o_j + \beta_j\right) \end{aligned} \tag{7.1}$$

with $\sigma_\lambda(x) = (1 + e^{-x})^{-\lambda}$.

In these equations, using terms derived from an analogy with real neurons, S_i, h_i, o_i represents the cell potential, τ_i the decay constant, g is a gain factor, x_i the intensity of the sensory neuron i, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, β_j the bias term, $fr\left(y_j\right) = \sigma(y_j + \beta_j)$ the firing rate. All decay constants

τ_i, all the network connection weights ω_{ij}, all biases β_j, and g are the genetically specified parameters of the networks. The biases β_j of the sensory neurons are all equal and are genetically determined.

7.3 The Evolutionary Process

A simple generational genetic algorithm was employed to set the parameters of the networks (Mitchell, 1996). The initial population contained 100 genotypes. Generations following the first one are produced by a combination of selection with elitism and with mutation. For each new generation, the 20 highest-scoring individuals from the previous generation, the elite, are retained unchanged. The remainder of the new population is generated by making 4 mutated copies of each of the 20 highest-scoring individuals. Each genotype is a vector comprising 420 parameters. Each parameter is encoded with 16 bits. Initially, a random population of vectors is generated. In the process of mutation, there is a 1.5% probability that each bit of the genotype can be flipped. The genotype parameters are linearly mapped to produce network parameters with the following ranges: biases $\beta_i \in [-4, -2]$, weights $\omega_{ij} \in [-6, 6]$, gain factor $g \in [1, 10]$; decay constants τ_i of the hidden layer are exponentially mapped into $[10^{-2}, 10^{0.3}]$ with the lower bound corresponding to the integration step-size used to update the controller and the upper bound, arbitrarily chosen, corresponding to about half of a trial length (i.e., $2\,s$). The cell potentials are set to 0 when the network is initialised or reset, and equations 7.1 are integrated using the forward Euler method[1].

During evolution, each genotype is translated into an arm controller and then evaluated 8 times starting from position A, and 8 times starting from position B, for a total of $K = 16$ trials (see figure 7.4). In position A, the angular positions of joints $\langle J_1, .., J_7 \rangle$ are $\langle -50°, -20°, -20°, -100°, -30°, 0°, -10° \rangle$, and for position B they are $\langle -100°, 0°, 10°, -30°, 0°, 0°, -10° \rangle$. For each position, the arm experiences the el-

[1] http://en.wikipedia.org/wiki/Euler_method

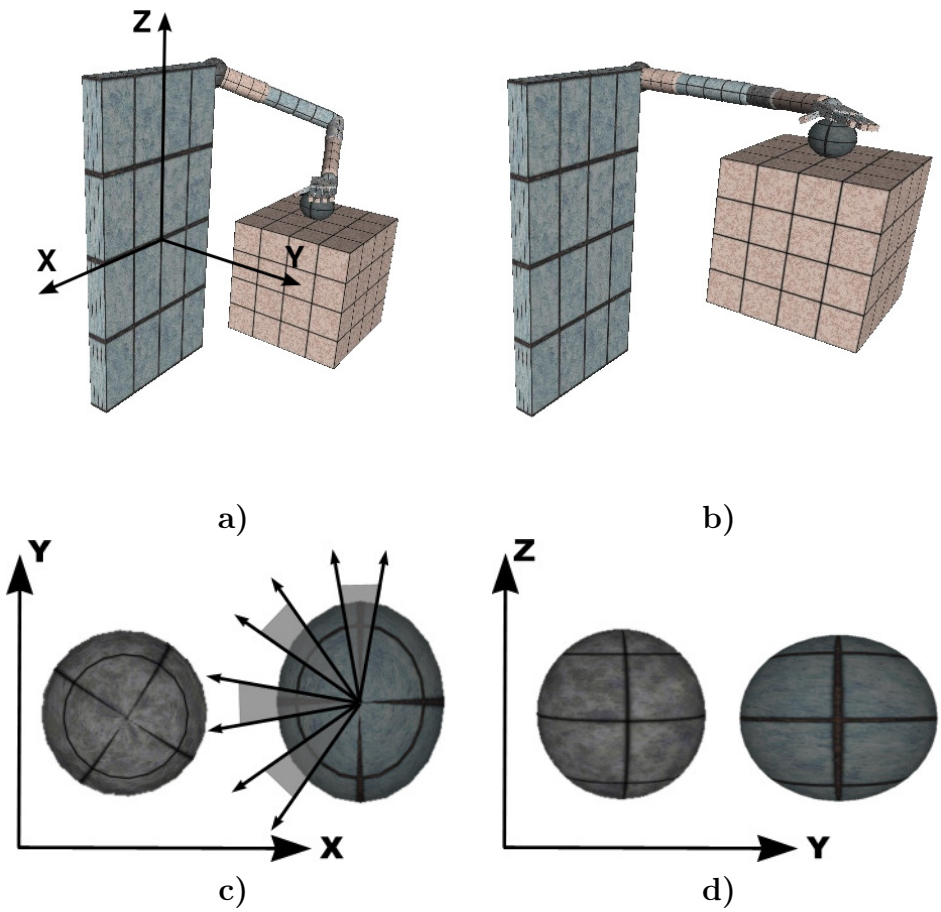


Figure 7.4: Initial positions of the arm and the objects. **a)** Position A for the arm, in which the angles of joints $\langle J_1, .., J_7 \rangle$ are $\langle -50°, -20°, -20°, -100°, -30°, 0°, -10° \rangle$. **b)** Position B for the arm, in which the angle of joints $\langle J_1, .., J_7 \rangle$ are $\langle -100°, 0°, 10°, -30°, 0°, 0°, -10° \rangle$. **c)** The sphere and the ellipsoid viewed from above. **d)** The sphere and the ellipsoid viewed from the side. The radius of the sphere is $2.5\,cm$. The radii of the ellipsoid are 2.5, 3.0 and $2.5\,cm$. In **c)** the arrows indicate the intervals within which the initial rotation of the ellipsoid is set.

lipsoid 4 times and the sphere 4 times. The radius of the sphere is $2.5\,cm$. The radii of the ellipsoid are 2.5, 3.0 and $2.5\,cm$. Moreover, the rotation of the ellipsoid with respect to the z-axis is randomly set in the range of $[350°, 10°]$ in the first presentation, $[35°, 55°]$ in the second presentation, $[80°, 100°]$ in the third presentation, and $[125°, 145°]$ in the fourth presentation (see also Figure 7.4-c: the arrows indicate the intervals within which the initial rotation of the ellipsoid is set).

At the beginning of each trial, the arm is located in the corresponding initial position (i.e., A or B), and the state of the neural controller is reset. A trial lasts 4 simulated seconds ($T = 400$ time step). A trial is terminated earlier in the case that an object

falls off the table.

In each trial k, an agent is rewarded by an evaluation function that seeks to assess its ability to recognise and distinguish the ellipsoid from the sphere. We would note that, rather than imposing a representation scheme in which different categories are associated with an a priori determined state/s of the categorisation neurons, the robot is left free to determine how to communicate the results of its decisions. That is, the agents can develop whatever representation scheme they might choose, as long as each object category is clearly identified by a unique state/s of the categorisation neurons. This system also has an advantage in that it scales up to categorisation tasks with objects of more than two categories, without having to introduce structural modifications to the agent's controller. More precisely, the agents are rewarded on the basis of the extent to which the categorisation outputs produced for objects of different categories are located in non-overlapping regions of a two-dimensional categorisation space $C \in [0,1] \times [0,1]$. The categorisation and the evaluation of the agent's discrimination capabilities is performed in the following way:

- In each trial k, the agent represents the experienced object (i.e., the sphere S or the ellipsoid E) by associating to it a rectangle R_k^S or R_k^E whose vertexes are:

$$\text{the bottom left vertex:}$$
$$(\min_{0.95T<t<T} fr(o_{17}), \min_{0.95T<t<T} fr(o_{18}))$$
$$\text{the top right vertex:}$$
$$(\max_{0.95T<t<T} fr(o_{17}), \max_{0.95T<t<T} fr(y_{18}))$$

- The sphere category, referred to as C^S, corresponds to the minimum bounding box of all R_k^S.

- The ellipsoid category, referred to as C^E, corresponds to the minimum bounding box of all R_k^E.

The final fitness FF attributed to an agent is the sum of two fitness components F_1 and F_2:

- F_1 rewards the robots for touching the objects, and corresponds to the average distance over a set of 16 trials between the centre of the palm and the experienced objects

$$F_1 = \frac{1}{16} \sum_{k=1}^{16} \left(1 - \frac{d_k}{d_{max}} \right)$$

where d_k is the Euclidean distance between the object and the centre of the palm at the end of trial k; d_{max} is the maximum distance that can be achieved between the centre of the palm and the object when located on the table.

- F_2 rewards the robots for developing an unambiguous category representation scheme on the basis of the position in a two-dimensional space of C^S and C^E:

$$F_2 = \begin{cases} 0 & if\ F_1 \neq 1 \\ 1 - \frac{area\left(C^S \cap C^E\right)}{min\{area(C^S), area(C^E)\}} & otherwise \end{cases}$$

note that $F_2 = 1$ if C^S and C^E do not overlap (i.e., if $C^S \cap C^E = \emptyset$).

The fact that, for each individual, F_1 must be 1 to be rewarded with F_2, constrains the evolution process to work on strategies in which the palm is in continuous contact with the object. This condition was introduced on the assumption that it represents a prerequisite for the ability to perceptually distinguish between the shapes of objects. However, alternative formalism, which encodes different evolutionary selective pressures, may work as well.

7.4 Results

Ten evolutionary simulations, each using a different random initialisation, were run for 500 generations.

Figure 7.5 shows the fitness of the best agent of each generation for the five evolutionary runs that managed to generate the highest-scoring individuals for at least 10 consecutive generations. The other five runs failed to achieve this first objective.

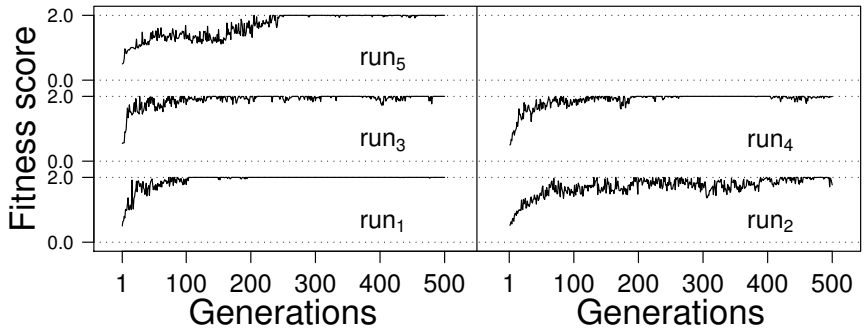


Figure 7.5: Fitness curves of the best agents. Graph showing the fitness of the best agents of each generation of the five evolutionary runs that managed to generate the highest-scoring individuals for at least 10 consecutive generations: run_1, run_2, run_3, run_4, and run_5

A quick glance at the curves in the figure shows that run_1 reaches very quickly (in about 100 generations) a plateau of the highest fitness score, and that it keeps on generating highest-scoring agents until the end of the evolution. run_2, run_3, run_4, and run_5 also generate highest-scoring agents but they need more generations, and the solutions seem to be more sensitive to the effect produced by those parameters of the task that were randomly initialised and/or by noise. Although all the agents with the highest fitness are potentially capable of accomplishing the task, the effectiveness and the robustness of their collective strategies need to be further estimated based upon more severe post-evaluation tests.

The next section presents the results of a series of post-evaluation tests whose aim was to estimate the robustness of the best-evolved discrimination strategies chosen from run_1, run_2, run_3, run_4, and run_5. Following these results, Section 7.4.2 presents the results of post-evaluation tests whose aim is to estimate the role of different sensory channels in categorisation. Finally, Section 7.4.3 presents analyses of the dynamics of the categorisation strategies of the best evolved agents.

It is important to note that although all the post-evaluation analyses were carried out on all the best evolved agents, for the sake of space, for several tests only the results concerning the performance of one of these agents are reported in the thesis. An ex-

haustive description of the analyses carried out on all the best evolved agents, the results of the tests that are not shown here, further simulations, as well as videos of the best evolved strategies, can be found at http://laral.istc.cnr.it/esm/active_perception

7.4.1 Robustness

To verify to what extent the robots were able to distinguish between two types of objects, regardless of the initial orientation of the ellipsoid object, specific post-evaluation tests were conducted (referred to as test P), in which the ellipsoid's initial orientation was systematically changed. More precisely, in test P, an agent was required to distinguish between the two objects placed in position A 360 times, and those placed in position B 360 times. In each position, the agent experienced the sphere half of the times (i.e. for 180 trials), and the ellipsoid half of the times (i.e. for 180 trials). Moreover, trial after trial, the initial orientation of the ellipsoid around the z-axis was changed by 1°, from 0° in the first trial, to 179° in the last trial. For each run, 10 agents chosen from among those with the highest fitness were post-evaluated. It is important to recall that these agents were selected from evolutionary phases in which the run managed to generate the highest-scoring individuals for at least 10 consecutive generations. Table 7.1 shows the results for the best agent A_j chosen from run_i, with $j, i = 1, ..., 5$.

Note that, compared to the evolutionary conditions in which the agents were allowed to perceive the ellipsoid only 4 times in 4 different initial orientations, P is a severe test. The results thus unambiguously tell us whether or not the five selected highest-fitness agents are capable of distinguishing the ellipsoid from the sphere and categorising them in a much wider range of initial orientations of the former object. For each selected agent, test P was repeated 5 times (i.e. P_i with $i = 1, .., 5$), with each repetition seeded differently in order to guarantee random variations in the noise added to the sensor readings.

The performance of agent A_j in test P_i was quantitatively established by considering

	A_1		A_2		A_3		A_4		A_5	
	R_k^E	R_k^S	R_k^E	R_k^S	R_k^E	R_k^S	R_k^E	R_k^S	R_k^E	R_k^S
P_1	357	360	310	351	340	358	347	356	355	354
P_2	359	360	311	347	342	358	356	358	356	355
P_3	356	360	312	349	343	356	348	355	356	354
P_4	357	360	304	353	341	355	342	354	354	355
P_5	358	360	303	348	349	356	349	355	353	353
Tot. (%)	3587 (99)		3288 (91)		3498 (97)		3520 (98)		3545 (98)	

Table 7.1: Results of post-evaluation tests P_i.

all the responses given by A_j over 3600 trials (i.e., 720 trials per test P_i, repeated 5 times, with $i, j = 1, ..., 5$). In each post-evaluation trial, the response of the agent was based on the firing rates of the categorisation neurons (o_{17} and o_{18}) during the last 20 time steps (i.e., $0.95T < t < T$) of each trail k. In particular, the lowest and the highest firing rates recorded by both neurons were used to define the bottom left and the top right vertexes of a rectangle, as illustrated in Section 7.3. At the end of each test P_i, this process generated 360 rectangles associated with the trials in which the agent experienced the sphere (i.e., rectangles R_k^S with $k = 1, .., 360$), and 360 rectangles associated with the trials in which the agent experienced the ellipsoid (i.e., rectangles R_k^E with $k = 1, .., 360$). At the end of the five post-evaluation tests P_i, from all the rectangles collected to that point, we calculated the highest number of R_k^S and R_k^E rectangles that could be included in two non-overlapping minimal bounding boxes C_i^S and C_i^E. These two rectangles then represent the sphere category (C_i^S) and the ellipsoid category (C_i^E) of agent A_i for all successive tests reported. The quantitative estimation of the robustness of an agent categorisation strategy is expressed as the percentage of the included rectangles R_i^S and R_i^E with respect to the excluded ones (which are considered as errors in categorisation).

The last row in Table 7.1 reports, for each agent A_i, the total number of rectangles that can be included in the minimal bounding-boxes without breaking the non-overlapping rule. These numbers are extremely high, showing an over 97% rate of success. These five agents are quite good at discriminating between and categorising

the sphere and the ellipsoid, with a much wider range of initial orientations of the ellipsoid. Only agent A_2 displayed slightly worse performance, and it was excluded from all further post-evaluation tests. The agents that performed higher than 95% on the first test P (i.e. A_1, A_3, A_4, and A_5) underwent a further series of tests in circumstances in which:

- the length of the longest radius of the ellipsoid progressively increased/decreased (see figure 7.6);
- the length of the radius of the sphere progressively increased/decreased (see figure 7.7);
- the initial position of the object and of the hand varied (see figure 7.8).

In these tests as well as all the other post-evaluation tests described from this point forward concerning A_1, A_3, A_4, and A_5, a trial k is considered:

- **Successful**: if, at the end, the rectangle containing the last 20 responses of neurons o_{17} and o_{18} falls completely within the corresponding bounding-box for the object experienced during the test; i.e. if it falls in C_i^S for the sphere and if it falls in C_i^E for the ellipsoid;
- **Unsuccessful** with a wrong response: if at the end the rectangle containing the last 20 responses of neurons o_{17} and o_{18} falss completely within the wrong bounding box during the test; i.e., if it falls in C_i^E for the sphere and if it falls in C_i^S for the ellipsoid;
- **Unsuccessful** with no response: if at the end the rectangle containing the last 20 responses of neurons o_{17} and o_{18} falls completely outside the bounding-boxes C_i^S and C_i^E.

As far as it concerns tests in which the length of the longest radius of the ellipsoid progressively increased/decreased, the distortions that further increased the longest

ellipsoid radius up to $1\,cm$ were rather well tolerated by the agents; both A_1 and A_5 managed to reliably differentiate between the two objects with a success rate higher than 90%. Distortions that tend to reduce the longest radius of the ellipsoid were clearly disruptive for all the agents, producing an expected 50% success rate when the ellipsoid was reduced to a sphere. In tests in which the ellipsoids had a radius that became progressively shorter that the radius of the sphere, the performance of all the agents was quite disrupted (see Figure 7.6).

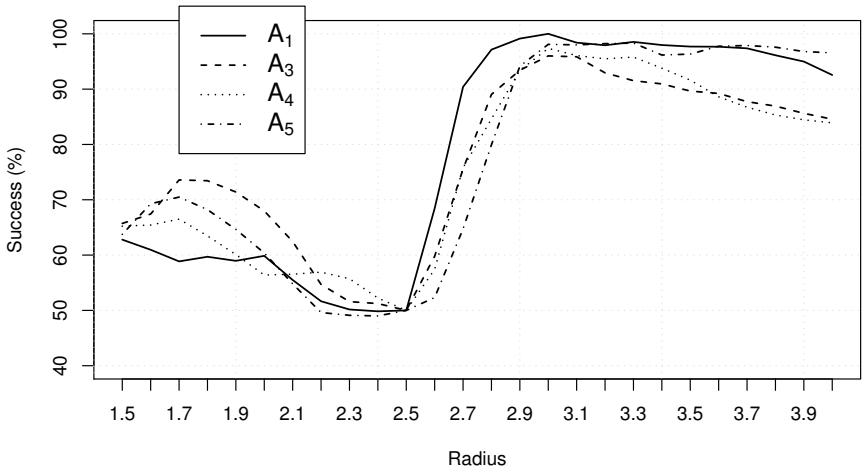


Figure 7.6: Performance on changing the radius of ellipsoid object. Graph showing the percentage of success in post-evaluation tests in which the length of the longest radius of the ellipsoid progressively increased/decreased

As far as it concerns tests in which the length of the radius of the sphere progressively increased/decreased, these distortions were particularly disruptive for all the agents except for A_5. This agent was also not disrupted to as great a degree as the other agents in tests in which the sphere became progressively smaller, and was very successful in tests in which the radius of the sphere was at least $7\,mm$ longer than the longest radius of the ellipsoid (see Figure 7.7).

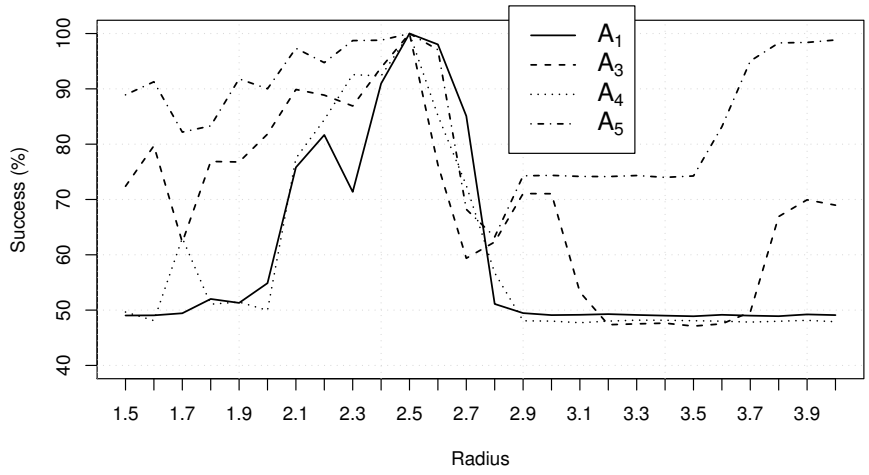


Figure 7.7: Performance with different radii of the sphere object. Graph showing the percentage of success in post-evaluation tests in which the length of the radius of the sphere progressively increased/decreased.

A further series of post-evaluation tests was performed in order to estimate the robustness of the best evolved strategies, in which the initial positions of the object and of the arm were changed. To simplify the analysis, the test focused upon only those circumstances in which the movements of the arm with respect to the initial positions experienced during evolution are produced by displacements of only one joint at a time (see Figure 7.8; the heights of the black and grey areas correspond to the percentage of success of the agent tested, upon changing the initial position of the joint indicated on the left; black represents position A, and grey position B). Although the results are quite heterogeneous, a number of features are shared by all the agents. First, displacements of joint J_1 for position A were tolerated quite well. Second, the wider the displacement, the greater the drop in performance, with the exception of J_4 for agents A_1, A_3, and A_4, in which displacements that tend to bring the hand/object progressively closer to the body resulted in better performance for both positions. It is important to note that A_4 is particularly sensitive to disruptions to joint J_1 and J_2 for position B, and to joint J_6 for position A.

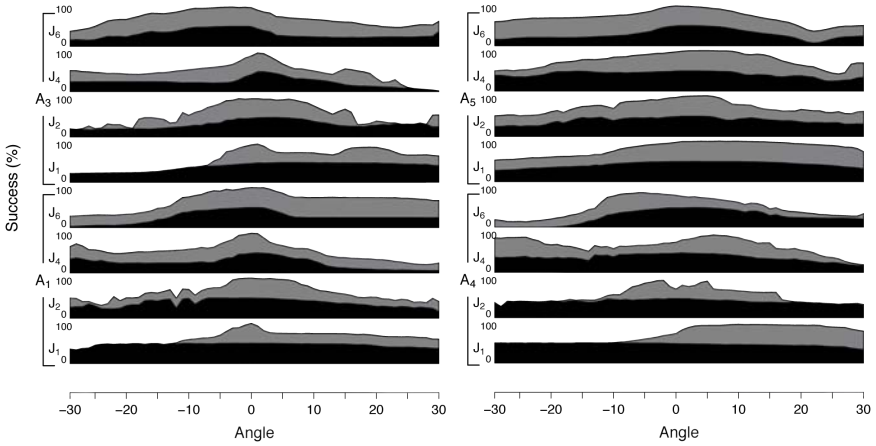


Figure 7.8: Performance of the agents with changes in the initial position of the arm, in which just one joint at a time was moved with respect to the initial positions used during evolution. The heights of the black and grey areas correspond to the percentage of success of agent A_i tested with changes in the initial position of joint J_i by the number of degrees indicated on the x axis; black represents position A, and grey position B.

7.4.2 The Role of Different Sensory Channels for Categorisation

To understand the mechanisms that enable agents A_1, A_3, A_4, and A_5 to solve their tasks, we first established the relative importance of the different types of sensory information that were available through the **arm propriosensors**, **tactile sensors**, and **hand propriosensors**. This was accomplished by measuring the performance displayed by the agents in a series of *substitution tests*, in which one type of sensory information experienced by each agent during its interaction with an ellipsoid was replaced with the corresponding type of sensory information that was previously recorded in trials in which the agent was interacting with a sphere. In these tests, each agent experienced the ellipsoid in all its initial rotations (i.e. from 0° to 179°), excluding those for which, given the randomly chosen seed for the tests, its responses turned out to be wrong in the absence of any type of substitution (i.e. the rectangle R_k^E did not fall within any of the five bounding-boxes C_i^E in the results of test P_i described above). For each ellipsoid's initial orientation, each substitution test was repeated 180 times. The rationale behind these tests is that any drop in performance caused by the substitution of a different type of sensory information provides an indication of the relative importance of that sensory channel in the

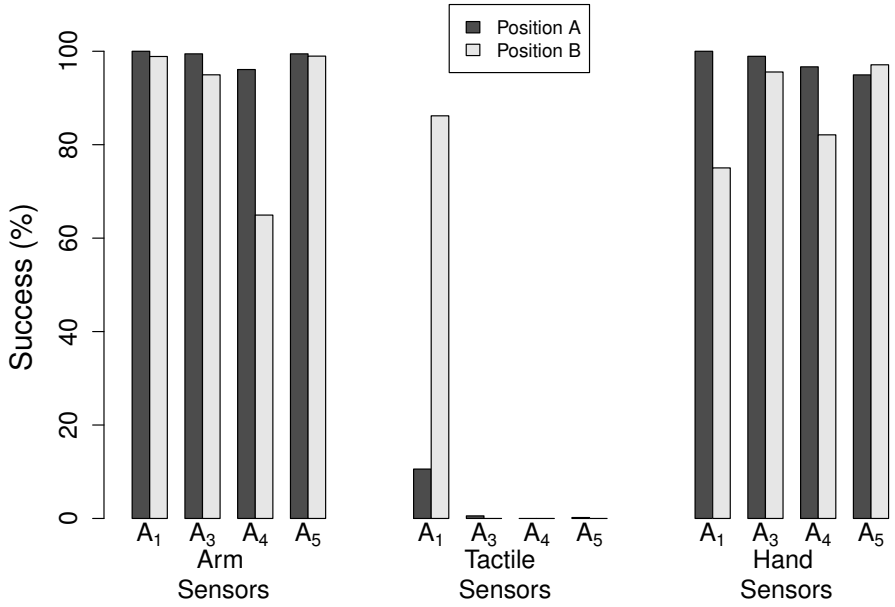


Figure 7.9: Results of substitution tests. Graphs showing, for agents A_1, A_2, A_3, A_4 and A_5, the results of substitution tests regarding the readings of the arm's proprioceptive sensors, tactile sensors, and the hand's proprioceptive sensors for position A (black columns) and for position B (grey columns).

categorisation process.

The results of this first series of substitution tests indicate that, for all the agents, the replacement of the sensory information originating from the arm's proprioceptive sensors and the hand's proprioceptive sensors in position A, only marginally interfered with their performance. That is, for position A, the agents underwent a substantial drop in performance only due to the replacement of tactile sensations (in Figure 7.9, see the black columns in correspondence with the tactile sensors). The clear drop in performance in these substitution tests concerning tactile sensation clearly indicates that for position A, the agents heavily relied upon tactile sensation to distinguish the ellipsoid from the sphere, and to correctly perform the categorisation task.

For position B, the results are slightly more heterogeneous. For agent A_1, the results of the substitution tests indicate that the replacement of both tactile sensations and of the hand's proprioceptive sensors produced about a 20% drop in performance

(In Figure 7.9, see the white columns in correspondence with the tactile and hand sensors). For the other agents, tactile sensation continued to be extremely important for the correct categorisation of the objects (In Figure 7.9, see the white columns in correspondence with the tactile sensors). However, for agent A_4, the replacement of the arm's and the hand's proprioceptive sensors produced a drop in performance of about 40% in the case of the arm, and 20% in the case of the hand sensors (In Figure 7.9, see the white columns in correspondence with the arm and the hand sensors). Hence, for agent A_1, the categorisation of the ellipsoid in position B was performed by exploiting information distributed over two sensory channels, that is, the tactile and the hand sensors. The information provided by these two sensory channels seems to be fused together in such a way that, for several orientations, the lack or the unreliability of information from one channel can be compensated for by the availability of reliable information from the other channel. The other agents seem to strongly rely upon tactile sensation, along with agent A_4, which also makes use of arm and hand sensations to discriminate among objects.

Given the above, we see that tactile sensation is the major source of discrimination cues in distinguishing spheres from ellipsoids in position A, for all the selected agents, and in position B for A_3 and A_5. Further investigations were then performed to see whether among the tactile sensors, there are any whose activation plays a predominant role in the categorisation task. In a series of further tests in which the substitution described above was applied only to single tactile sensors, the performance of all agents remained largely above 90%. Hence, the categorisation ability of the agents was not compromised by replacements that selectively affected the functioning of single tactile sensors.

Next, a different series of substitution tests was developed, in which all possible combinations of the two elements of the tactile sensors were replaced. Although this analysis was carried out for all the agents for position A, and for agents A_3, and A_5 for position B, this chapter reports only the results for agent A_1 (i.e. the best performing agent, see Table 7.1) for position A. The results are shown in Figure 7.10,

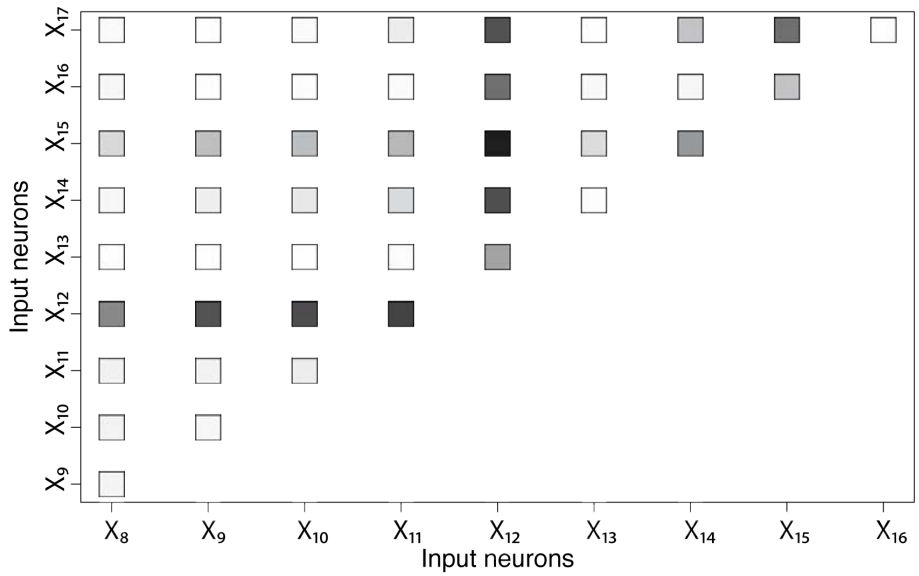


Figure 7.10: Result of *substitution tests* for combinations of two tactile sensors. Graph showing the results of *substitution tests* concerning the readings X_i with $i = 8, \cdots, 17$ of all the possible combinations of two elements of the tactile sensors for position A. Each square is coloured in a shade of grey. The grey scale is proportional to the percentage of success, with white indicating combinations in which the agent is 100% successful, and black combinations in which the agent is 100% unsuccessful.

in which each square is coloured in a shade of grey proportional to the percentage of success. The colour white indicates a combination in which the agent was 100% successful, and black a combination in which the agent was 100% unsuccessful.

These substitution tests did not produce clear-cut results. However, in Figure 7.10 we can note that there are specific sensors which, when disrupted in combination with any other sensor, produce a clear drop in performance. In particular, disruptions applied to the reading of the tactile sensors placed on the third phalanx of the middle finger (x_{12}), and in more minor terms, disruption applied to the reading of the tactile sensors placed on the first phalanx of the ring finger (x_{15}), caused the agent to mistake the ellipsoid for the sphere. Hence, agent A_1 heavily relied on the patterns of activation of the tactile sensors, in which the readings of x_{12} and x_{15} were particularly important in distinguishing the ellipsoid from the sphere.

With regard to the other agents, the performance of agent A_3 drops in position A when substitutions concerned the reading of x_{10} in combination with any other tactile sensor. In position B, a drop in performance was recorded when substitutions

concerned the reading of x_8 or x_{12} in combination with any other sensor. Agent A_4 in position A was particularly disrupted by substitutions concerning the reading of x_{11} or x_{12} in combination with any other sensor. Agent A_5 in position A was disrupted by substitutions concerning the reading of x_{12} with any other sensor, and of x_{12} or x_{17} with any other sensor in position B. In conclusion, in these circumstances, the agents tended to rely upon a combination of tactile sensors, with the tactile sensor on the third phalanx of the middle finger being more significant than the other sensors for all agents.

7.4.3 On the Dynamics of the Categorisation Process

This section presents a series of analyses whose aim is to reveal the dynamics of the categorisation process. More specifically, the analyses concern:

- to what extent the sensory stimuli experienced while the agents interact with the objects provide the regularities required to categorise the objects;
- to what extent the agents succeed in self-selecting discriminating stimuli (i.e. stimuli that can be unambiguously associated with either category);
- how long the agents need to interact with the object before being able to recognize whether they are touching a sphere or an ellipsoid;
- whether the categorisation process occurs instantaneously by exploiting the regularities provided by single unambiguous sensory patterns or whether it occurs over time by integrating the regularities provided by multiple stimuli.

Qualitative and quantitative tests were specifically designed to answer these questions. The former are simply composed of observations of the trajectories of the categorisation outputs in the two-dimensional categorisation space $C \in [0,1] \times [0,1]$, in single trials. The latter tests further explore the dynamics of the categorisation processes by taking advantage of the fact that in both positions, almost all the

best evolved agents exploit tactile sensation to carry out the task. The quantitative tests were carried out on all the agents for position A, and on agents A_3 and A_5 for position B. Here, we report only the details for the analysis of A_1 (i.e. the best performing agent, see Table 7.1) for position A. It turned out, however, that successful categorisation strategies are very similar from a behavioural point of view, as well as in terms of the mechanisms exploited to perform the task. Therefore, the operational description of A_1 is also representative of the categorisation strategies of A_3, A_4, and A_5 in position A, and of A_3 and A_5 in position B.

The aim of the first two tests was to establish to what extent the stimuli experienced by A_1 during its interactions with the objects provide the regularities required to categorise them. The analysis begins by computing a slightly modified version of the Geometric Separability Index (hereafter, referred to as GSI). The GSI, which was originally proposed by Thornton (1997), is an estimate of the degree to which tactile sensor readings associated with a sphere or with an ellipsoid are geometrically separated in sensory space. It is also related to the complexity of the categorisation task. In fact, if all tactile the sensors can be separated geometrically by means of a linear equation, the GSI reaches its maximum and the categorisation task is quite easy (i.e. there is no need for non-linear neurons and/or hidden neurons and/or recurrences). The test generates 800 data sets in total. 400 data sets concern each time step while the agent interacts with the ellipsoid, $\{\tilde{X}_k^E\}_{k=1}^{180}$, and 400 data sets concerning each time step while agent interacts with the sphere, $\{\tilde{X}_k^S\}_{k=1}^{180}$. Where, $\tilde{X}_k^E$ is the tactile sensors reading ($X = \langle x_8, \ldots, x_{17} \rangle$) experienced by the agent while interacting with the ellipsoid at time step t of trial k; and $\tilde{X}_k^S$ is the tactile sensor reading experienced by the agent while interacting with the sphere at time step t of trial k. Here, we should recall that trial after trial, the initial rotation of the ellipsoid around the z-axis was changed by 1°, from 0° in the first trial to 179° in the last trial. Each trial was differently seeded in order to guarantee random variations in the noise added to the sensor readings. At each time step t, the GSI was computed as follows:

$$
\begin{aligned}
GSI(t) &= \frac{1}{180}\sum_{k=1}^{180} z_k \\
z_k &= \begin{cases} 1 & if\, m^{EE} < m^{ES} \\ 0 & if\, m^{EE} > m^{ES} \\ \frac{u}{u+v} & otherwise \end{cases} \\
m^{EE} &= \min_{\forall j \neq k} H\left(\tilde{X}_k^E, \tilde{X}_j^E\right) \\
m^{ES} &= \min_{\forall j} H\left(\tilde{X}_k^E, \tilde{X}_j^S\right) \\
u &= \left|\left\{\tilde{X}_j^E : H\left(\tilde{X}_k^E, \tilde{X}_j^E\right) = m^{EE}\right\}_{\forall j \neq k}\right| \\
v &= \left|\left\{\tilde{X}_j^E : H\left(\tilde{X}_k^E, \tilde{X}_j^S\right) = m^{ES}\right\}_{\forall j}\right|
\end{aligned}
$$

where $H(x, y)$ is the Hamming distance between tactile sensors readings. $|A|$ denote the cardinality of the set A. m^{EE} is the minimum distance from the tactile pattern k for the data set concerning the ellipsoid. m^{ES} is the minimum distance from the tactile pattern k for the data set concerning the sphere. The terms u and v count the number of tactile patterns at distance m^{EE} and m^{ES}, respectively. $GSI(t)$ is equal to 1 indicates that at time step t, the closest neighbourhood of each $\tilde{X}_k^E$ is one or more elements of the set $\tilde{X}_k^E$. $GSI(t)$ equal to 0 indicates that at time step t, the closest neighbourhood of each $\tilde{X}_k^E$ is one or more elements from the set $\tilde{X}_k^S$.

As shown in Figure 7.11, for agent A_1 and position A, the $GSI(t)$ tends to increase from about 0.5 at time step 1 to about 0.9 at time step 200, and it remains around 0.9 until time step 400. This trend suggests that during the first 200 time steps, the agent acts in such a way as to bring forth those tactile sensor readings that facilitate the object identification and classification tasks. In other words, the behaviour exhibited by the agent allows it to experience two classes of sensory states that tend to become progressively more separated in the sensory space. However, the fact that the GSI does not reach a value of 1 indicates that the two groups of sensory patterns

belonging to the two objects are not fully separated in the sensory space. In other words, some of the sensory patterns experienced during the interactions with an ellipsoid are very similar or identical to those experienced during interactions with the sphere and vice versa.

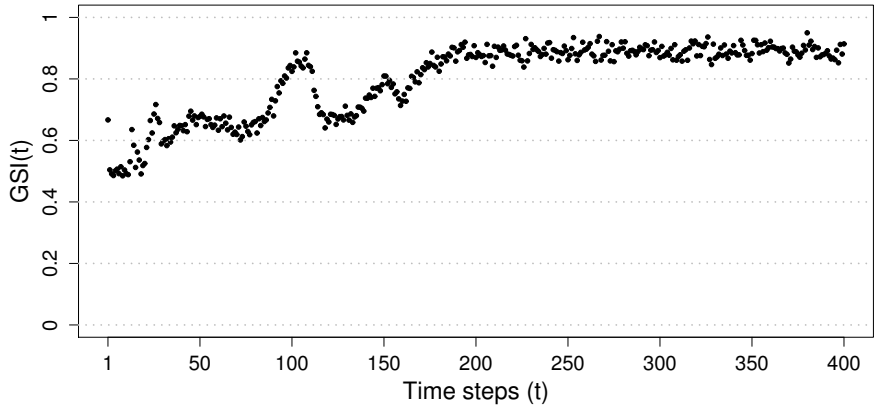


Figure 7.11: $GSI(t)$ for agent A_1. The values of $GSI(t)$ calculated for each time step for the 800 data sets generated.

To analyse in more detail to what extent the stimuli experienced by the agent could be associated with the correct or the wrong category, an index called *E-representativeness* (E-repr) was designed. The E-repr is computed from a set of 32,400 trials, which are produced by repeating 180 times each of the 180 trials corresponding to different ellipsoid initial orientations, from 0° to 179°. During these trials, for each single tactile sensor pattern, the number of times each pattern appears during interactions with the ellipsoid (N) and during interactions with the sphere (M) is recorded The E-repr of a single pattern is given by $\frac{N}{N+M}$. It is important to note that an E-repr of 1.0 or 0.0 corresponds to fully discriminating stimuli that can be unambiguously associated with the ellipsoid or the sphere category, respectively, while 0.5 corresponds to completely ambiguous stimuli.

The graph in Figure 7.12 presents the E-repr of the last 20 patterns (i.e. the patterns recorded from time step 380 to time step 400) of single successful trials of test P_i, which was described in Section 7.4.1. Each trial refers to a different initial orientation

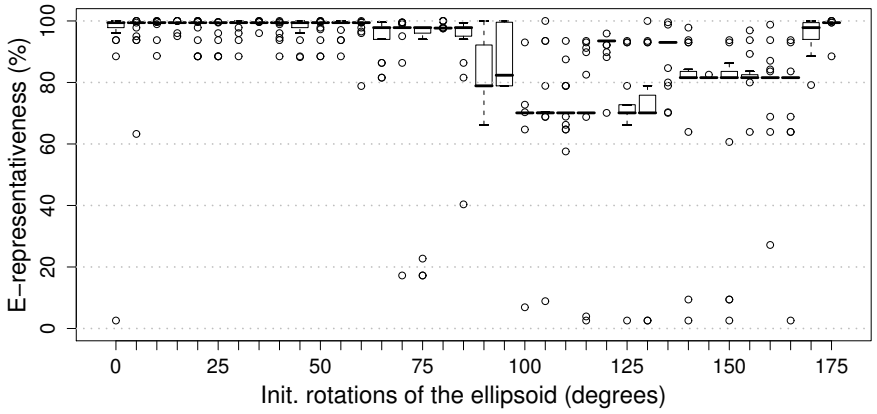


Figure 7.12: E-representativeness of the tactile sensors patterns. The graph shows the value of the E-representativeness of the tactile sensors patterns recorded in the last 20 time steps of 180 different trials with the ellipsoid. The x axis indicate the initial rotation of the ellipsoid in degrees. For each rotation, the corresponding boxplot on the graph shows the minimum, median, and maximum observation of E-representativeness over the 180 trials.

of the ellipsoid. A quick glance at Figure 7.12 shows that there are trials in which the agent had to deal with tactile sensor patterns that had a very low E-repr. That is, they were very weakly associated with the ellipsoid. Patterns with a very low E-repr tend to appear in trials in which the initial orientation of the ellipsoid is chosen in the interval $[75°, 175°]$. These patterns may have at least two origins that are not mutually exclusive:

1. They may be due to the fact that the agent is not able to effectively position the object in such a way as to unequivocally recognize whether there is a sphere or an ellipsoid; and

2. they may be determined by the noise injected into the system.

The fact that agent A_1 succeeds in correctly distinguishing the category of the objects, even during trials in which it does not experience fully discriminating stimuli, indicates that the problem was solved by integrating over time the partially conflicting evidence provided by sequences of stimuli. In fact, if the agent employed a reactive strategy (i.e. with no need for a memory structure), it would be deceived

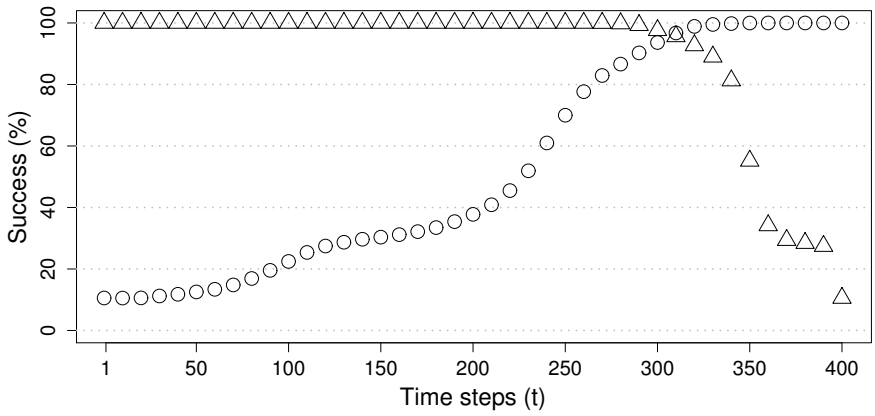


Figure 7.13: Performance on pre-substitution and post-substitution tests. The graph shows the percentage of success in pre-substitution tests (triangles) and post-substitution tests (circles). The points are at intervals of 10 time steps starting from 0.

by those sensor patterns that are quite strongly associated with the sphere, that appear in interactions with the ellipsoid. Under this circumstance, an agent that employs a reactive strategy would mistake the ellipsoid for a sphere. Since, in spite of the deceptive patterns, the agent was 100% successful, it appears that the agent employed a discrimination strategy that uses the dynamic properties of its controller (time-dependent neuron states and recurrent connections).

Other evidence that supports the integration-over-time hypothesis comes from additional analyses that were performed employing additional types of substitution tests. In one test in particular, for a certain time interval, the tactile sensor patterns experienced by A_1 in interactions with the ellipsoid were replaced by those experienced in interactions with the sphere. In a first series of tests, referred to as *pre-substitution tests*, substitutions were applied from the beginning of each trial up to time step t, where $t = 1, \ldots, 400$. In a second series of tests, referred to as *post-substitution tests*, substitutions were applied from time step t, where $t = 1, \ldots, 400$, to the end of a trial. Each test was repeated at intervals of 10 time steps. For agent A_1 and position A, the results of the pre-substitution and post-substitution tests are illustrated in Figure 7.13.

This graph shows that, regardless of the rotation of the ellipsoid, pre-substitutions that did not affect the last 100 time steps did not cause any drop in performance. For pre-substitution tests that involved more than 300 time steps, the degree to which the performance dropped was higher for longer substitution periods (see the triangles in Figure 7.13). Similarly, the agent did not incur any drop in performance if the post-substitutions affected less than 100 time steps. For post-substitution tests that affected more than the last 100 time steps, the degree to which the performance dropped was higher for longer substitution periods (see the empty circles in Figure 7.13).

The results of these pre/post-substitution tests suggest that the agent was integrating sensory states over time for a certain amount of time around time step 310. In particular, the results shown in Figure 7.13 seem to indicate that, as regards agent A_1 position A, the interactions between the agent and the objects can be divided into the following three temporal phases, which are qualitatively different, from the point of view of the categorisation process:

- an initial phase whose upper bound can be approximately fixed at time step 250, in which the categorisation process begins, but in which the categorisation answer produced by the agent is still reversible;
- an intermediate phase whose upper bound can be approximately fixed at time step 350, in which a categorisation decision is quite often taken on the basis of all previously experienced evidence; and
- a final phase, in which the previous decision (which is now irreversible) is maintained.

The fact that the categorisation decision formed by A_1 during the initial phase is not yet definitive is demonstrated by the fact that substitutions of the critical sensory stimuli performed during this phase did not cause any drop in performance (see Figure 7.13, triangles). The fact that the intermediate phase corresponds to a

critical period is demonstrated by the fact that the pre/post-substitution tests that affect this phase produced a significant drop in performance (see Figure 7.13). The fact that A_1 makes its ultimate decision during the intermediate phase is demonstrated by the fact that the post-substitution tests that affect the last 80 time steps, approximately, did not produce any drop in performance (see Figure 7.13, empty circles).

Further tests, namely *window-substitution tests*, were employed in order to estimate the existence and the dimensions of the hypothesised temporal phase in which it is supposed that the agent integrates tactile sensor states. In these tests, substitutions were applied before and after a temporal window centred around time step 310. The length of the temporal window with no substitutions could vary from 1 time step (i.e. no substitution at time step 310) to 69 time steps (i.e. no substitution from time step 276 to 344). As shown in Figure 7.14, the wider the window with no substitutions, the higher the performance of the agent, with a 100% success rate when no substitutions were applied to a temporal phase of about 50 time steps or longer. Although the graph in Figure 7.14 does not exclude the possibility that the agent employed an instantaneous categorisation process, it seems to suggest that the performance of the agent is in some way correlated to the amount of empirical evidence it manages to gather over time, starting from about time step 270, until time step 340.

Finally, additional evidence that supports the hypothesis of a dynamic categorisation process based on the integration of tactile sensation over time comes from a qualitative analysis of the trajectories of the categorisation outputs in the two-dimensional categorisation space $C \in [0,1] \times [0,1]$, in single trials. Figure 7.15-a shows the trajectory recorded by A_1 in a trial in which the initial orientation of the ellipsoid was 115°. As we can see, A_1 moves rather smoothly in the categorisation space by reaching the corresponding bounding-box in slightly less than $2\,s$ (200 time steps). If we then look at Figure 7.15-b, we see that during the interaction with the ellipsoid, A_1 experienced:

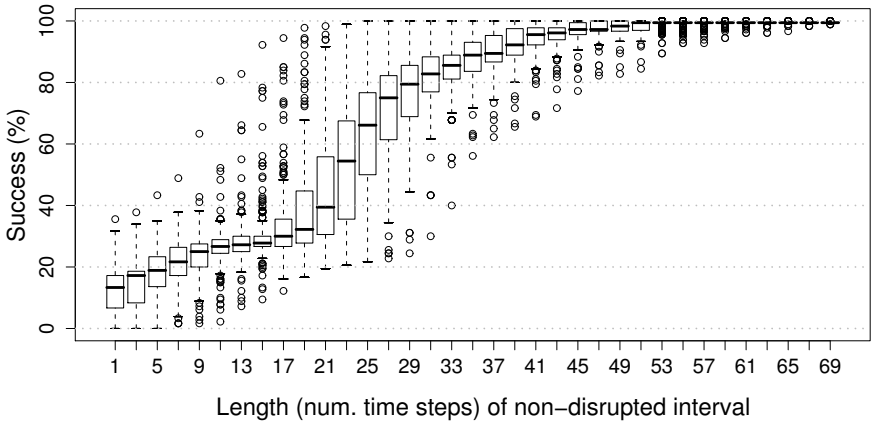
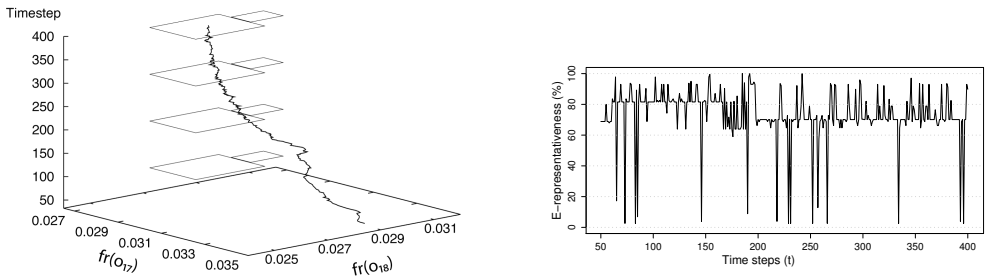


Figure 7.14: Performance on window-substitution tests. The graph shows the percentage of success of window-substitution tests. The x axis is the length of the temporal window with no substitutions, centred around time step 310.

- few stimuli with a high percentage of E-repr (i.e., stimuli that are experienced in interactions with an ellipsoid object most of the times);
- several stimuli with an intermediate level of E-repr (i.e., stimuli that are experienced in interactions with the ellipsoid and the sphere in about 3/4 and 1/4 of the cases, respectively); and
- few stimuli with a low percentage of E-repr (i.e., stimuli that are experienced in interaction with a spherical object most of the times).

If we visually compare Figure 7.15-a with 7.15-b, it is possible to note that the experienced sensory patterns with a different percentage of E-repr appear to drive the categorisation output in different regions of the categorisation space, which correspond to the ellipsoid and the sphere bounding-boxes, respectively. Therefore, the final position of the categorisation output (i.e. the categorisation decision) is not determined by a single or by a few selected patterns. Rather, it is the result of a process extended over time, in which partially conflicting evidence provided by the experienced tactile sensation is integrated over time. Similar dynamics were observed by inspecting all other trials. Given this evidence, it is likely that the performance of all the best evolved agents in position A, and of agent A_3 and A_5 in



a) **b)**

Figure 7.15: Comparison of categorisation outputs and E-representativeness of tactile patterns over time for the same trial. **a)** Trajectory of categorisation outputs from $t = 50$ to end of trial; the large and small rectangles at 100, 200, 300, and 400 time steps indicate the bounding box of the ellipsoid and sphere category, respectively. **b)** E-representativeness of the tactile sensory patterns recorded in the same trial as **a)** with the ellipsoid initially orientated at 115°

position B, is the result of a dynamic categorisation process based on the integration of tactile sensation over time.

7.5 Discussion

This chapter describes an experiment in which a simulated anthropomorphic robotic arm acquires the ability to categorise un-anchored spherical and ellipsoid objects placed in different positions and orientations over a planar surface. The agents' neural controller was trained through an evolutionary process in which the free parameters of the neural networks were varied randomly, and in which variations were retained or discarded on the basis of their impact on the overall ability of the robots to carry out their task. This implies that the robots were left free to determine:

- how to interact with the external environment (by eventually modifying the environment itself);
- how the experienced sensory stimuli are used to distinguish between the two categories; and
- how to represent each object category in the categorisation space.

The analysis of the obtained results indicates that the agents were indeed capable of developing the ability to effectively categorise the shapes of the two types of objects despite the high degree of similarity between them, the difficulty of effectively controlling a body with many DOFs, and the need to master the effects produced by gravity, inertia, collisions, etc. More specifically, the best individuals displayed an ability to correctly categorise the objects when they were located in different positions and orientations from those already experienced during evolution, as well as an ability to generalise their skill to objects, positions, and orientations they had never experienced during evolution. Moreover, the agents were robust enough to deal with categorisation tasks in which the longest radius of the ellipsoid was progressively increased. Other distortions of the dimensions of the original objects resulted in a greater degree of disruption. These results prove that the proposed method can be successfully applied to scenarios that appear to be more complex than those investigated in previous works based on similar methodologies.

The analysis of the best evolved agents indicates that one fundamental skill that enables them to solve the categorisation problem consists in the ability to interact with the external environment and to modify the environment itself, so as to experience sensory states that are progressively more different for different categorical contexts. This result represents a confirmation of the importance of sensory-motor coordination, and more specifically, of the active nature of situated categorisation, which has already been highlighted in previous studies (Scheier et al., 1998; Nolfi & Marocco, 2002).

On the other hand, the fact that sensory-motor coordination does not allow the agents to experience fully discriminating stimuli demonstrates how in some cases, sensory-motor coordination needs to be complemented by additional mechanisms. Such a mechanism, in the case of the best evolved individuals, consists in an ability to integrate the information provided by sequences of sensory stimuli over time. More specifically, the analyses performed suggest that agent A_1 categorised the current object as soon as it experienced useful regularities, and that the categorisation

process was realised during a significant period of time (i.e. about 50 time steps), during which the agent kept using the experienced evidence to either confirm and reinforce its current tentative decision, or to change it. Similar strategies were also observed in the other three best evolved agents. In this regard, see also (Townsend & Busemeyer, 1995; Platt, 2002; Beer, 2003).

The importance of the ability to integrate the regularities provided by sequences of stimuli was also confirmed by the results obtained in a control experiment, which was replicated 10 times, in which the agents were provided with reactive neural controllers (i.e. neural networks without recurrent connections, with simple logistic internal neurons, and in which all other parameters were kept the same as those described in this chapter). Indeed, the performance displayed by the best evolved individuals in this control experiment was significantly worse than that observed in the basic experiment, in which the agents were allowed to keep information about previously experienced sensory states. Although this does not exclude the possibility that different experimental scenarios (e.g. scenarios involving agents provided with different neural architectures and/or physical characteristics different from those of the agents) could lead to qualitatively different results, the analysis of the results obtained in this specific scenario, taken overall, indicate that the task does not admit of pure reactive solutions, or alternatively, that such solutions are hard to synthesise through an evolutionary process. This mixed conclusion may also be due to the functional constraints the limit the movement of the robotic arm (e.g. the fact that the fingers could not be extended/flexed separately, or that there was no adduction/abduction of the fingers), as well as other implementation details (e.g. the dimensions of the objects with respect to the hand).

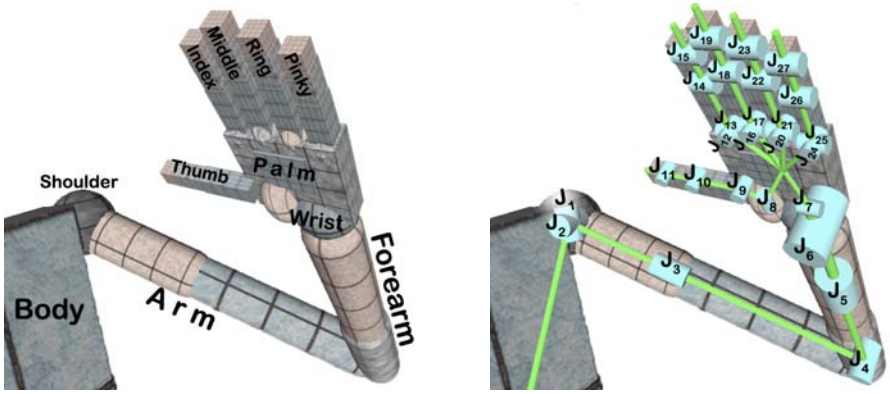
The analysis of the role played by different sensory channels indicates that the categorisation process in the best evolved individuals is primarily based on tactile sensors, and secondarily, on the hand and arm proprioceptive sensors (with the arm proprioceptive sensors playing a role only for agent A_4 position B; see Figure 7.9). It is interesting to note that at least one of the best evolved agents (i.e. A_1) not

only displayed the ability to exploit all relevant information, but also the ability to combine information coming from different sensory modalities in order to maximise the chance that it would make the appropriate categorisation decision (Waxman, 2003). More specifically, the ability to combine the information provided by the tactile and hand proprioceptive sensors, for objects located in position B, enables the robot to correctly categorise the shape of the object in the majority of cases, even when one of the two sources of information has been corrupted (see Figure 7.9).

8 Reaching, Grasping, Lifting: On the facilitatory role of 'linguistic' input

This chapter presents an anthropomorphic robotic arm with a five-fingered hand controlled by an artificial neural network that is evolved to have the ability to manipulate spherical objects located on a table by reaching for, grasping, and lifting them. The robot develops the sensory-motor coordination required to carry out this task in two different conditions, one in which it receives as input *linguistic instructions* (binary input vectors) that specify the type of elementary behaviour to be exhibited during a certain period of the task, and the other in which it receives no such instructions. The obtained results shown that linguistic instructions facilitate the development of the required behavioural skills. These instructions are binary input vectors associated with elementary behaviours that need to be displayed by the robot during the task. They are referred to as linguistic instructions because they are not related to any measurable property of or entity in the environment (i.e. distances, angles, positions, etc.), and also because they are not perceived in the same way as other inputs, but instead represent symbolic entities (the behaviour to be displayed) that resemble a very simple language.

The main objective of the study presented in this chapter is to investigate whether the use of linguistic instructions facilitates the acquisition of a sequence of complex behaviours. The long-term goal of this research is to verify whether the acquisition of elementary skills guided by linguistic instructions provides a scaffolding for more complex behaviours.



(a) **(b)**

Figure 8.1: The kinematic chain of the arm and the hand. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation. The links among the cylinders represent the rigid connections that make up the arm structure. The joints are named as indicated in **b).**

8.1 The Robot

The robot that is the subject of the experiment presented in this chapter is a variant of the full anthropomorphic manipulator used in the previous experiments. The details of the differences in the robot's structure are presented in Appendix C. In the following sections, only a broad view of the arm and hand structure is given.

8.1.1 Arm Structure

The arm (Figure 8.1) consists mainly of three elements (the arm, the forearm and the wrist) that are connected through articulations distributed in the shoulder, arm, elbow, forearm, and wrist. It is an enhancement of a previous 4-DOF model, to which is added a wrist that comprises another 3-DOF joint. The wrist adds the ability to produce pitch, yaw and roll of the five-fingered hand. In Figure 8.1, the cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation, and the links among the cylinders represent the rigid connections that make up the arm structure.

8.1.2 Arm Actuators

The joints of the arm are actuated by two simulated antagonist muscles that are implemented accordingly to Hill's muscle model (Sandercock et al., 2003; Shadmehr & Wise, 2005b). More precisely, the total force exerted by a muscle is the sum of three forces $T_A(\alpha, x) + T_P(x) + T_V(\dot{x})$ which depend on the activity of the corresponding motor neuron (α), the current elongation of the muscle (x) and the muscle contraction/elongation speed ($\dot{x}$), which are calculated on the basis of equations B.1 (for details see the appendix C).

The active force T_A depends on the activation of muscle α and on the current elongation/compression of the muscle. When the muscle is completely elongated/compressed, the active force is zero regardless of the activation α. When the muscle is at its t resting length, the active force reaches its maximum, which depends upon the activation α.

The passive force T_P depends only on the current elongation/compression of the muscle. T_P tends to elongate the muscle when it is compressed to less than its resting length, and tends to compress the muscle when it is elongated beyond its resting length. T_P differs from a linear spring in that it has an exponential trend that produces a strong opposition to muscle elongation and little opposition to muscle compression.

T_V is the viscosity force. It produces a force that is proportional to the velocity of the elongation/compression of the muscle.

8.1.3 Hand Structure

The hand is attached to the robotic arm just below the wrist (at joint J_7, as shown in Figure 8.1). One of the most important features of the hand is its compliance. In detail, this compliance was obtained by setting a maximum threshold of $300\,N$ to the force exerted by each joint. When an external force acting on a joint exceeds this

threshold, the joint either cannot move further, or it moves backward in response to the external force.

The robotic hand is composed of a palm and 15 phalanges that make up the digits (three phalanges for each finger) that are connected through 20 DOFs, $J_8, \ldots, J_{27}$ (see Figure 8.1 and Appendix C for details).

8.1.4 Hand Actuators

The joints are not controllable independent of each other, but rather, they are grouped. The same grouping principle that was used in developing the iCub hand (Sandini et al., 2004) was used here. Essentially, there are only 9 actuators that move all the joints of the hand. For details on which joints are moved by these 9 actuators, see Appendix C. These actuators are simple motors that are control joints in terms of their positions.

8.1.5 Hand Tactile Sensors

The hand is equipped with tactile sensors that are distributed over the wrist, the palm, and all five fingers. The tactile sensors are placed and that behave exactly as in the previous experiment; see Figure 6.2-b and Section 7.1.5 for details.

8.2 The Neural Controller

The architecture of the neural controllers varies slightly, depending upon the ecological conditions in which the robot develops its skills. In the case of the development supported by linguistic instructions, the robot is controlled by a neural network, as shown in Figure 8.2 which includes 29 sensory neurons ($x_1, \ldots, x_{29}$), 12 internal neurons ($h_1, \ldots, h_{12}$) with recurrent connections and 23 motor neurons ($o_1, \ldots, o_{23}$). In the case that no support is given by linguistic instructions, the neural network lacks

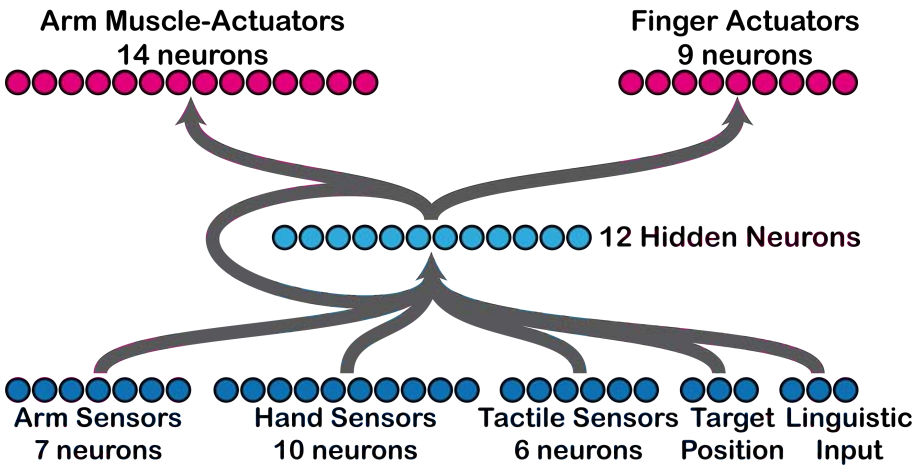


Figure 8.2: The architecture of the neural controllers. The arrows indicate blocks of fully connected neurons.

the sensory neurons that are dedicated to the linguistic instructions (x_{27}, x_{28}, x_{29}). Thus, it is then composed of 26 sensory neurons instead of 29. The neurons are divided into height blocks in order to facilitate the description of their functionality and connectivity:

- **Arm Propriosensors** $(x_1, \ldots, x_7)$: The activation values x_i of the arm's propriosensory neurons encode the current angles of the 7 corresponding DOFs located on the arm and wrist normalised in the range of $[0, 1]$.

- **Hand Propriosensors** $(x_{8,} \ldots, x_{17})$: The vector of the activation values $\langle x_8, \ldots, x_{17} \rangle$ of the hand's propriosensor neurons correspond to the following vector, which is computed on the basis of current angles of the hand's joints:

$$\Big\langle a\left(J_8\right), a\left(J_9\right), \tfrac{a(J_{10})+a(J_{11})}{2}, a\left(J_{13}\right), \tfrac{a(J_{14})+a(J_{15})}{2},$$
$$a\left(J_{17}\right), \tfrac{a(J_{18})+a(J_{19})}{2}, a\left(J_{21}\right), \tfrac{a(J_{22})+a(J_{23})}{2}, a\left(J_{12}\right)\Big\rangle$$

  where $a\left(J_i\right)$ is the angle of joint J_i normalised in the range of $[0, 1]$ with 0 meaning fully extended, and 1 fully flexed. This way of representing the hand posture mirrors the way in which the hand joints are actuated (see Section 8.1.4).

- **Tactile Sensors** $(x_{18}, \ldots, x_{23})$: These measure whether the 5 fingers and the unit constituted by the palm and wrist are in physical contact with another

object. More precisely, the output of each tactile sensor is calculated with the following equation:

$$\begin{aligned} x_{18}(t) &= map_{[0,1]}(T_P + T_W) \\ x_i(t) &= map_{[0,1]}\left(\sum_{s=1}^{3} T_{s+3(i-19)}\right) \; for\, i = 19, \ldots, 23 \end{aligned}$$

where T_i is the value of the corresponding tactile sensor, and $map_{[0,1]}$ normalises the number of contacts in the range of $[0, 1]$. Normalisation is performed using a map function that becomes saturated to 1 when more than 20 contacts take place.

- **Target Position** (x_{24}, x_{25}, x_{26}): These neurons receive the output of a vision system (which was not simulated) that computes the relative distance in *cm* of the object with respect to the hand over three orthogonal axes. These values are fed into the networks, since they are without any normalisation. In detail, if $P_{target} = \langle x, y, z \rangle$ are the Cartesian coordinates of the target object with respect to a common fixed frame, and $P_{hand} = \langle x, y, z \rangle$ are the Cartesian coordinates of the centre of the palm with respect to the same fixed frame, then the values of the target position neurons are: $\langle x_{24}, x_{25}, x_{26} \rangle = P_{hand} - P_{target}$.

- **Linguistic Input** $(x_{27,}x_{28}, x_{29})$: This is a block of three neurons, each of which represents one of the three commands: *reach*, *grasp* and *lift*. Specifically, the vector $\langle 50, 0, 0 \rangle$ corresponds to the linguistic instruction "*reach for the object*", $\langle 0, 50, 0 \rangle$ corresponds to the linguistic instruction "*grasp the object*" and $\langle 0, 0, 50 \rangle$ corresponds to the linguistic instruction "*lift the object*". The way in which the state of these sensors is set is determined by equation 8.1, as explained below.

- **Internal Neurons** $(h_1, \ldots, h_{12})$: They are fully recurrent

- **Arm Actuators** $(o_1, \ldots, o_{14})$: The values o_i directly indicate the activation

status of the 14 motor neurons that control the corresponding muscles of the arm.

- **Hand Actuators** $(o_{15}, \ldots, o_{23})$: The values o_i correspond to the desired extension/flexion positions of the nine hand actuators, as described in Section 8.1.4. For more details, also see Appendix C.

Note that the state of the **Linguistic Input** and **Target Position** varies at a larger interval than the other sensors in order to increase the relative impact of these neurons. Indeed, control experiments in which all sensory neurons were normalised within a $[0, 1]$ interval led to significantly lower performance (results not shown).

The state of the sensors, the desired state of the actuators, and the internal neurons are updated every $10\,ms$, accordingly to the following equations:

$$\begin{aligned} h_i(t) &= \delta_i \left(\sum_{j=1}^{29} w_{ji} \sigma_{0.2}\left(x_j(t)\right) + \beta_i \right) + (1-\delta_i)\, h_i(t-1) \\ o_i &= \sum_{j=1}^{12} w_{ji} \sigma_{0.2}\left(h_j\right) \end{aligned}$$

where $\sigma_\lambda(x) = (1+e^{-x})^{-\lambda}$. x_i are the output of sensory neurons as described above. h_i and o_i are the output of the internal and actuator neurons, respectively. w_{ji} is the synaptic weight from neuron j to neuron i. β_i is the bias of the i-th neuron. Also, δ_i is the coefficient for implementing leaky neurons. as proposed in (Nolfi & Marocco, 2001) for the internal neurons. With respect to the hidden neurons, the output neurons do not have any bias or decay-factor.

This particular type of neural network architecture was chosen in order to minimise the number of assumptions and to reduce, as much as possible, the number of free parameters. Also, this particular sensory system was chosen in order to be able to study situations in which the visual and tactile sensory channels need to be integrated.

8.3 The Evolutionary Process

The free parameters of the neural controller (i.e. the connection weights, the biases of the internal neurons, and the time-constant of leaky-integrator neurons) were set using an evolutionary algorithm (Nolfi & Floreano, 2000; Yao & Islam, 2008).

The initial population consisted of 100 randomly generated genotypes, which encode the free parameters of 100 corresponding neural controllers. In the conditions in which **Linguistic Inputs** are employed (hereafter, referred to as Exp. A), the neural controller has 792 free parameters. In the other condition, without **Linguistic Inputs** (hereafter, referred to as Exp. B), there are 756 free parameters. Each parameter is encoded into a binary string (i.e. a gene) of 16 bits. In total, a genotype is composed of $792 \cdot 16 = 12672$ bits in Exp. A, and $756 \cdot 16 = 12096$ bits in Exp. B. In both experiments, each gene encodes a real value in the range of $[-6, +6]$, but for genes encoding the decay-factors δ_i, the encoded value is mapped in the range of $[0, 1]$.

The 20 best genotypes of each generation were allowed to reproduce by generating five copies each. Four out of five copies were subject to mutation, and one copy did not mutate. During mutation, each bit of the genotype had a 1.5% probability of being replaced by a new, randomly selected value. The evolutionary process was repeated for 1,000 generations.

The agents were rewarded for reaching, grasping and lifting a spherical object with a radius of $2.5\,cm$ that was placed on the table in exactly the same way as in both Exp. A and Exp. B. Each agent of the population was tested 4 times, and each time, the initial position of the arm and the sphere were changed. Figure 8.3 shows the four initial positions of the arm and of the sphere superimposed on one another. The four initial postures of the arm corresponded to the following angles of joints $J_1, \ldots J_4$: $\langle -73, -30, -40, -56 \rangle$, $\langle -73, -30, -40, -113 \rangle$, $\langle -6, +30, -10, -56 \rangle$ and $\langle -73, -30, +45, -113 \rangle$. In addition, the initial sphere positions were: $\langle -18, +10 \rangle$, $\langle -26, +18 \rangle$, $\langle -18, +26 \rangle$ and $\langle -10, +18 \rangle$. Also, for each initial arm/object config-

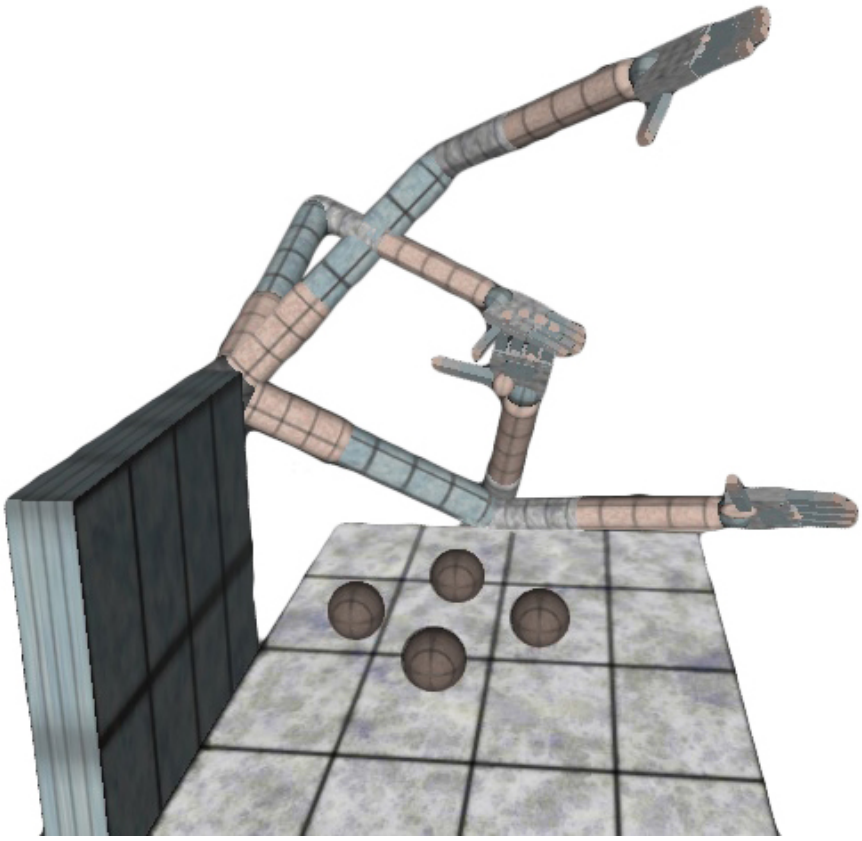


Figure 8.3: Initial positions of the arm and the sphere superimposed on one another. The four initial postures of the arm correspond to the following angles, given in degrees of joints J_1, J_2, J_3, J_4: $\langle -73, -30, -40, -56 \rangle$, $\langle -73, -30, -40, -113 \rangle$, $\langle -6, +30, -10, -56 \rangle$ and $\langle -73, -30, +45, -113 \rangle$. Also, the initial sphere positions, in cm, are $\langle -18, +10 \rangle$, $\langle -26, +18 \rangle$, $\langle -18, +26 \rangle$ and $\langle -10, +18 \rangle$.

uration, a random displacement of $\pm 1^o$ was added to each joint of the arm, and a random displacement of $\pm 1.5\, cm$ was added to the x and the y coordinates of the sphere position. Each trial lasted $6\, s$, which corresponds to 600 simulation steps. The sphere was able to move freely and could eventually fall off the table, in which case the trial was stopped prematurely.

The fitness function was made up of three components: FR for reaching, FG for grasping, and FL for lifting the object. Each trial was divided into 3 phases, in each of which only a single fitness component was updated. The conditions that defined the current phase at each time-step, and consequently, determined which component had to be updated, were as follows:

$$
\begin{aligned}
r(t) &= 1 - e^{(-0.1 \cdot ds(t))} \\
g(t) &= e^{(-0.2 \cdot graspQ(t))} \\
l(t) &= 1 - e^{(-0.3 \cdot contacts(t))} \\
Phase(t) &= \begin{cases} reach & r(t) > g(t) \vee g(t) < 0.5 \\ grasp & otherwise \\ lift & g(t) > 0.7 \wedge l(t) > 0.6 \end{cases}
\end{aligned}
\tag{8.1}
$$

where $ds(t)$ is the distance from the centre of the palm to a point located $5\,cm$ above the centre of the sphere. The term $graspQ(t)$ is the distance between the centroid of the fingertips-palm polygon and the centre of the sphere. The term $contacts(t)$ is the number of contacts between the fingers and the sphere.

The shifts between the three phases were irreversible (i.e. the reach phase was always followed by the reach or grasp phases, and the grasp phase was always followed by the grasp or lift phases).

Essentially, the current phase is determined by the values $r(t)$, $g(t)$ and $l(t)$. When $r(t)$ is high (i.e. when the hand is far from the object), the robot must reach for the object. When $r(t)$ decreases and $g(t)$ increases (i.e. when the hand approaches the object from above), the robot needs to grasp the object. Finally, when $l(t)$ increases (i.e. when the number of activated contact sensors is large enough) the robot is able to lift the object. The rules and the thresholds included in equation 8.1 were set manually on the basis of our intuition and were adjusted through a trial-and-error process. In Exp. A, the phases were used to define which linguistic instruction the robot perceives.

The three fitness components were calculated in the following way:

$$
\begin{aligned}
FR &= \sum_{t \in T_{Reach}} \left(\frac{0.5}{1 + ds(t)/4} + \frac{0.25}{1 + ds(t)} \left(fingersOpen(t) + palmRot(t) \right) \right) \\
FG &= \sum_{t \in T_{Wrap}} \left(\frac{0.4}{1 + graspQ(t)} + \frac{0.2}{1 + contacts(t)/4} \right) \\
FL &= \sum_{t \in T_{Lift}} objLifted(t)
\end{aligned}
$$

where T_{Reach}, T_{Wrap} and T_{Lift} are the time ranges determined by equation 8.1. $fingersOpen(t)$ corresponds to the average degree of extension of the fingers, where 1 denote that all fingers are extended and 0 that all fingers are closed. $palmRot(t)$ is the dot product between the normal of the palm and the table, with 1 denoting the condition in which the palm is parallel to the table, and 0 to the condition in which it is orthogonal to the table. $objLifted(t)$ is 1 only if the sphere is not touching the table and is in contact with the fingers, otherwise it is 0.

The total fitness was calculated at the end of four trials as: $F = min\,(500, FR) + min\,(720, FW) + min\,(1600, FL) + bonus$, where $bonus$ adds 300 for each trial in which the agent switches from the reach phase to the grasp phase only, and 600 for each trial in which the agent switches from the reach to grasp phase and from the grasp to lift phase.

During the reach phase the agent is rewarded for approaching a point located $5\,cm$ above the centre of the object with the palm parallel to the table and the hand open. Note that the rewards for the hand opening and the rotation of the palm are relevant only when the hand is near the object (due to $0.25/(1 + ds(t))$ factor). In this way, the agent is free to rotate the palm when the hand is away from the sphere, thus allowing any reaching trajectory.

During the grasp phase, the centroid of the fingertips-palm polygon can reach the centre of the sphere only when the hand wraps the sphere within the fingers, producing a potential power grasp.

During the lift phase, the reward is given when the agent effectively moves the sphere

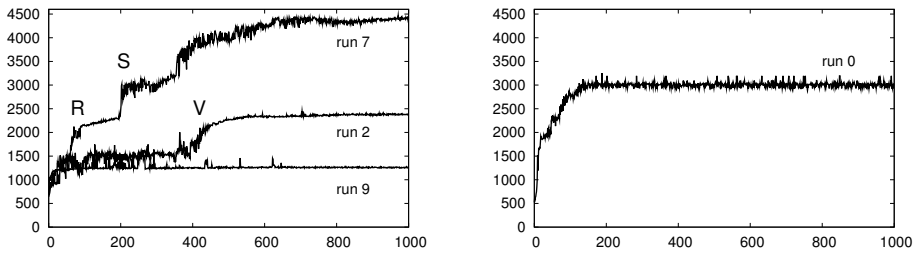


Figure 8.4: Fitness curves of the best agents at each generation of **a)** run 2, run 7, and run 9 of Exp. A, and **b)** run 0 of Exp. B.

up, off the table.

8.4 Results

For both Exp. A (with linguistic instructions) and Exp. B (without linguistic instructions), 10 evolutionary simulations for 1,000 generations were run, each using a different random initialisation. Looking at the fitness curves of the best agents of each generation of each evolutionary run, we noticed that for Exp. A, there are three distinct evolutionary paths (see Figure 8.4-a). The most promising is run 7, in which the last generation's agents have the highest fitness. The curve corresponding to run 2 is representative of a group of seven evolutionary paths which, after a short phase of fitness growth, reach a plateau at $F = 2000$. The curve corresponding to run 9 is representative of a group of two evolutionary paths that are characterised by a long plateau slightly above $F = 1000$. Generally speaking, these curves progressively increase by going through short evolutionary intervals in which the fitness grows quite rapidly, and is then followed by a long plateau.

In Exp. B, all the runs show a very similar trend, reaching and constantly remaining on a plateau at about $F = 3000$ (see Figure 8.4-b).

Due to the nature of the task and of the fitness function, it is quite hard to infer from these fitness curves what the behaviour of the agents might be during each evolutionary phase. However, based on the characteristics of the task, and by visual

inspection of the behaviour exhibited by the agents, it is possible to figure out how the agents behaved in different generations of each evolutionary run. In Exp. A, the phases of rapid fitness growth are determined by the *bonus* factor, which substantially rewards those agents that successfully accomplish single parts of the task. The first jump in fitness is due to the *bonus* factor associated with the execution of a successful reaching behaviour. This jump corresponds to the phase of fitness growth observed in run 7 in correspondence with label R in Figure 8.4-a, and in run 2 in correspondence with label V in Figure 8.4-a. The agents generated after these jumps in fitness jumps are able to systematically reach the object. Run 9 does not produce this first jump in fitness, and the agents of this run lack the ability to systematically carry out a successful reaching behaviour.

The second jump in fitness is due to the *bonus* factor associated with the execution of a successful grasping behaviour. Only in run 7 is it possible to observe a phase of rapid fitness growth corresponding to a second jump in fitness (see label S in Figure 8.4-a). The agents generated after this jump are able to successfully carry out reaching and grasping. Note also that in run 7, the fitness curve keeps on growing until the end of the evolution. This growth is determined by the evolution of the capability to lift the object. Thus, in run 7, the best agents following generation 400 are capable of reaching, grasping, and lifting the object. The constant increment of fitness is determined by the fact that the agents become progressively more effective in lifting the object. Run 2 does not produce a second jump in fitness jump. The agents of this run lack the ability to systematically carry out a successful grasping behaviour.

In summary, only run 7 generated agents (i.e. those best agents generated after generation 400) capable of successfully accomplishing reaching, grasping, and lifting. The best agents of run 2, and of the other six runs that show a similar evolutionary trend, are able to systematically reach but not grasp the object, and completely lack the ability to lift it. The best agents of run 9, and of the other runs that show a similar evolutionary trend, are not even able to systematically reach the object. In

Exp. B, they are able to successfully reach and grasp the object, but not lift it.

8.4.1 Robustness & Generalisation

The effectiveness and robustness of the best agents' behavioural strategies was evaluated in a series of post-evaluation tests. In these tests, the agents, from generation 900 to generation 1,000 of each run, were subjected to a series of trials in which the position of the object as well as the initial position of the arm were systematically varied. For the position of the object, a rectangular area ($28\,cm \times 21\,cm$) divided into 11×11 cells defined the possible displacements of the object. The agents were evaluated for reaching, grasping and lifting the object, which was positioned in the centre of each cell of the rectangular area. From the four initial positions employed during evolution (see Figure 8.3), 100 slightly different initial positions were generated with the addition of a $\pm 10^{\circ}$ random displacement to joints J_1, J_2, J_3, and J_4. Thus, this test was comprised of 48,400 trials, given by 400 initial positions ($4 \cdot 100$) for each cell, which were repeated for 121 cells corresponding to the different initial positions of the object during the test. In each trial, reaching was considered successful if an agent met the conditions and was able to switch from the reach phase to the grasp phase (see equation 8.1). Grasping was considered successful if an agent met the conditions and was able to switch from the grasp phase to the lift phase (see equation 8.1). Lifting was considered successful if an agent managed to keep the object more than $1\,cm$ above the table until the end of the trial.

The results shown here concern a single agent for each run. However, agents belonging to the same run produced very similar performance. Thus, the reader should consider the results of each agent as being representative of all the other agents in the same evolutionary run.

All the graphs in Figure 8.5 show the relative positions of the rectangular area and the cells with respect to the agent/table system. Moreover, each cell in this area is coloured in a shade of grey, with black indicating a 0% success rate, and white

indicating a 100% success rate. As expected from the results in the previous section, the agent chosen from run 7 Exp. A proved to be the only one capable of successfully accomplishing all three phases of the task. This agent proved capable of successfully reaching the object when it was placed almost anywhere within the rectangular area.

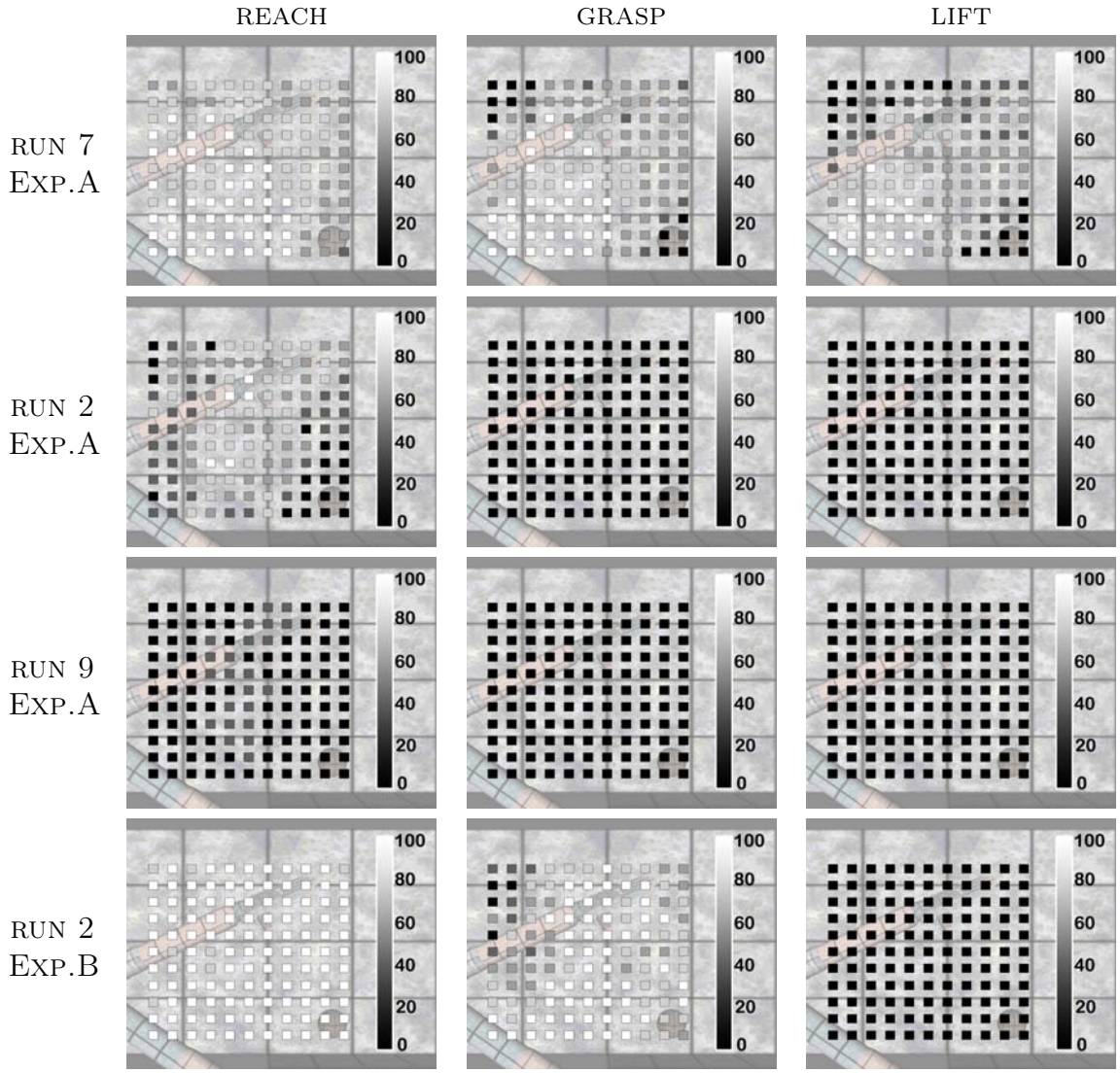


Figure 8.5: Performance on robustness tests. Performance of the best agent at generation 1,000, indicated by the row on post-evaluation tests regarding the robustness of reaching, grasping and lifting behaviours. The coloured cells indicate the initial positions of the object, and the background shows the positions of the cells with respect to the table and the robotic arm. Each cell is coloured depending upon the average performance obtained when testing the robot over 400 trials in which the initial posture of the arm was varied. A white cell corresponds to a 100% rate of success, and black to a 0% rate of success. (See the text for details as to how success/failure was computed).

Its grasping and lifting behaviour were less robust than its reaching behaviour.

Indeed, its grasping and lifting performance was quite good everywhere, except in two small zones located in the top-left and bottom-right of the rectangular area in which the cells are coloured black. The agent chosen from run 2 Exp. A proved to be capable of successfully performing reaching behaviour for a broad range of initial positions of objects, but was completely unable to perform grasping and lifting behaviours. The agent chosen from run 9 Exp. A did not even manage to systematically bring the hand close to the object, regardless of the object's initial position. The agent chosen from run 0 Exp. B proved capable of successfully performing reaching and grasping behaviours, but not lifting behaviour.

8.5 Discussion

This chapter has shown how a simulated humanoid robot controlled by an artificial neural network can acquire the ability to manipulate spherical objects set out on a table by reaching for, grasping, and lifting them. The agent was trained through an adaptive process in which the free parameters encoded the control rules that regulate the fine-grained interaction between the agent and the environment, and the variations of these free parameters were retained or discarded on the basis of their effects at the level of the behaviour exhibited by the agent. This means that the agents developed their skills autonomously through interaction with the environment. This further means that the agents are left free to determine the ways in which they solve the task, within the limits imposed by i) their body/control architecture, ii) the characteristics of the environment, and iii) the constraints imposed by the utility function, which rewards the agents for their ability to reach an area located above the object, wrap the fingers around the object, and lift the object.

An analysis of the best individuals generated by the adaptive process shows that the agents of a single evolutionary run managed to reach for, grasp, and lift the object reliably and effectively. Moreover, when tested in new conditions with respect to those experienced during the adaptive process, these agents proved capable of gen-

eralising their skills with respect to new object positions they had never experienced before.

A comparison of two experimental conditions (i.e. with and without the use of linguistic instructions that specify the behaviours that the agents are required to exhibit during the task) indicates that the agents succeeded in solving the entire problem only with the support of linguistic instructions (i.e. in Exp. A). This result confirms the hypothesis that access to linguistic instructions that represent the category of the behaviour to be exhibited in the current phase of the task, might be a crucial pre-requisite for the development of the corresponding behavioural skills, and for the ability to trigger the right behaviour at the right time. More specifically, the fact that the best agents in Exp. B succeeded in exhibiting reaching and then grasping behaviour, but not lifting behaviour, suggests that linguistic instructions represent a crucial pre-requisite in situations in which the agent has to develop the ability to produce different behaviours in similar sensory-motor circumstances. The transitions from reaching to grasping were marked by well differentiated sensory-motor states, which were probably sufficient to induce the agents to stop the reaching phase and to start the grasping phase, even without the support of a linguistic instruction. The grasping-to-lifting transition was not characterised by well differentiated sensory-motor states. Thus, in Exp. A, it is likely that it was the valuable support of the linguistic instruction that induced the successful agents to move on to the lifting phase.

9 Conclusions

The main research aim of this PhD thesis was to use evolutionary robotics methodologies to synthesise neural controllers for anthropomorphic robots in order for them to be able to manipulate objects. Looking at this problem, we find that various abilities are necessary in order for an anthropomorphic robot to manipulate objects. The skill of reaching enables the hand to approach the object. The skill of grasping enables it to take and hold the object. The ability to discriminate among different objects allows the robot to trigger different actions.

During the author's PhD studies, these aspects were studied using ER and by using neural networks as controllers. The first experiment, presented in Chapter 5, concerns the development of reaching behaviour for a simple robotic arm without a hand. Next, a full model of an anthropomorphic robotic arm with a five-fingered hand was implemented, and the second experiment, reported in Chapter 6, employed this robotic arm in order to develop a grasping behaviour. The experiments that followed, in Chapters 7 and8, addressed two different problems beyond that of grasping ability: the discrimination among different objects on the basis of tactile information, and the ability to perform a sequence of actions that consists of reaching, grasping, and lifting behaviours.

Overall, an analysis of the obtained results indicates that robots equipped with neural network controllers can successfully achieve these tasks and can demonstrate good performance in two ways:

- with respect to the fitness function used and the robustness tests that were performed, and
- by generalising their skills when tested in new conditions (with different initial positions of the arm/hand, and with objects located in different positions and/or orientations, with objects of different shapes).

An analysis of the behaviours displayed by evolved individuals and a comparison of the different experimental conditions indicate that the methodology used tends to produce solutions that are parsimonious from the point of view of the control system, due to the fact that they exploit properties emerging from the interaction between the robot and the environment. In fact, the agents solve their adaptive tasks while also evolving an ability to interact with the external environment and to modify the environment itself in order to self-select favourable conditions.

The obtained results also demonstrate how the robots can successfully develop the ability to display multiple behaviours (i.e. reaching, grasping, and lifting) and to arbitrate among them. The co-development of different behaviours leads to strongly integrated solutions in which behaviours are realised in a way that maximises the chances of success of the other behaviours. For example, reaching is realised in a way that maximises the chances that the successive execution of the grasping behaviour will be successful. In the case of the evolved behaviours reported in Chapter 6, the robots reach for and make contact with objects on one side, and then move the arm and the palm in a way that ensures that the objects will move toward the inner part of the hand, thus facilitating the grasp behaviour. Hence, the ability to exploit the interactions between behaviours leads to solutions that are effective and parsimonious, from the point of view of the control system.

This tight integration, however, also tends to prevent the robot from having the possibility to develop relatively independent behavioural skills that can be recombined in different sequences and reused to achieve different functions. The experiment reported in Chapter 8, in which the robots were rewarded both for the ability to display each elementary behaviour in isolation, and to combine different elementary behaviours in sequence, represents a way to overcome the disadvantages of the tight integration of behaviours.

9.1 Contribution to Knowledge

From the point of view of the author, the results of the experiments reported in this thesis make a number of contributions to knowledge. Following the order in which the experiments were reported, these contributions are as follows.

- In chapter 5, it was demonstrated that it is possible to develop a reliable and efficient solution of the inverse kinematic problem using a simple neural network topology and with a simple neuron model, with respect to (Kokera et al., 2004; Manocha & Zhu, 1994; Toal & Flanagan, 2002; Williamson, 1998; Li & Leong, 2003; Oyama et al., 2001; Martìn & del R Del Milla, 1998; Rathbone & Sharkey, 1999; Krose & der smagt, 1993a; Bekey, 2006), and that this could be achieved without constraining the kinematic structure of the robotic arm in order to satisfy the requirements for a solution to the IK problem (Siciliano & Khatib, 2008). This approach to solving the IK problem, which was also used in all successive experiments, makes it possible to evolve reliable and efficient solutions for reaching with a full anthropomorphic arm, which are essential to the achievements of all the other experiments.

- Chapter 6 demonstrated the feasibility of applying the ER approach to complex robots in order to evolve complex behaviours, such as grasping objects. In fact, the experimental setup presented is significantly more advanced than that of previous works that are based on similar adaptive techniques (Bianco & Nolfi, 2004; Buehrmann & Paolo, 2004; Gomez et al., 2005; Massera et al., 2006; Bongard, 2010). The morphology of the anthropomorphic arm and hand with 27 DOFs is rather more complex than that of the arm models cited. Hence, the size of the neural controller and the dimensions of the corresponding search space are greater. Also, the task involves the ability to reach and grasp freely moving objects with different shapes placed on a table. The results of the experiments presented not only in this chapter, but also in 7 and 8, demonstrate that the ER approach can be scaled up to this level of complexity.

- Chapter 7 presents a new mechanism to categorise objects that was designed by the author of this thesis. In the large majority of cases, researchers have focused their attention on categorisation processes that are passive and instantaneous (see (Cohen & Lefebvre, 2005) for a comprehensive review), in which the neural networks usually must represent categories in a such way as to (i) form areas related to a certain category that are as concentrated as possible around a centre (which is often used to represent the prototype of that category), and (ii) to keep these areas as far as possible from one another. Taking a different approach, the experiment in Chapter 7 focuses upon categorisation processes that are active and that are eventually distributed over time (Beer, 2000; Nolfi, 2009). Hence, the actions and behaviours exhibited by the agent later influence the stimuli that it senses, which results in evolved robots acting so as to experience the regularities that enable them to categorise appropriately. In addition, the relaxation of the two aforementioned constraints and the use of er make it possible to avoid the imposition of a representation scheme in which different categories are associated with a state/s of the categorisation neurons that is/are determined a priori. This results in at least two advantages of the proposed categorisation mechanism: (i) the points inside a category's area can be structured in such a way as to represent some feature of the category itself, and (ii) the category's area can be arranged in a way that facilitates the discrimination task.

- Chapter 8 highlighted the facilitatory role of external inputs in the development of the skill of manipulating objects. In the particular setup presented in Chapter 8, neural networks are evolved to produce the ability to manipulate spherical objects distributed on a table by reaching for, grasping, and lifting them under two different conditions: one in which they receive as input a linguistic instruction that specifies the type of behaviour to be exhibited during the current phase, and the other in which they receive no such instruction. The obtained results show that the linguistic instructions facilitated the

development of the required behavioural skills. The results suggest that the acquisition of elementary skills with the guidance of linguistic instructions may provide a scaffolding for more complex behaviours. Although the experimental setup is quite simplified as compared to real social interactions, the linguistic instructions may seen as a very simple form of interaction between a robot and a teacher. The results can thus be considered supportive of existing evidence about the importance of social interactions in the development of complex manipulation behaviours (Cangelosi et al., 2010, 2007; Cappa & Perani, 2003; Glenberg & Kaschak, 2002; Hauk et al., 2004; Pulvermuller, 2002; Rizzolatti & Arbib, 1998).

9.2 Future Ideas

A variety of future works are planned that have the following aims:

- port some of the results obtained through simulations of real robots;
- integrate aspects of this research that have thus far been studied in isolation; and
- carry on additional research by combining the development of behavioural and linguistic skills.

As regards the first point, future research will focus on porting the most significant results on the iCub robot by exploiting the compliant system recently developed by Mohan et al. (2009). Such a compliant system, which consists of force sensors placed on the arm and a software library, makes possible the implementation of muscle-like actuators on the iCub.

As regards the integration of the experiments described in this thesis, new experiments will be designed in which the robots use their ability to categorise the shapes

of objects in order to appropriately modify the way in which the objects are manipulated (i.e. in order to trigger force or precision grasping behaviour, depending on the type of object).

Finally, as regards the third point, the experiment presented in Chapter 8 will be extended by training the robot to also self-generate linguistic instructions and use them to trigger the corresponding behaviours autonomously (i.e. without the need to rely on external instructions), and by verifying whether the role played by linguistic instructions can later be internalised in the agents' cognitive abilities (Vygotsky, 1962, 1978; Mirolli & Parisi, 2009).

10 References

Angeles, J. (2003). *Fundamentals of Robotic Mechanical Systems: Theory, Methods, and Algorithms*. Springer, 2nd edition edition. 3.1, 3.1, 3.2

Arbib, M. A. (2003). *The Handbook of Brain Theory and Neural Networks*. MIT Press, 2nd edition edition. 2

Arbib, M. A., Iberall, T., & Lyons, D. (1985). Coordinated control programs for movements of the hand. *Experimental Brain Research*, 10, 111–129. 3

Bar-Yam, Y. (1997a). *Dynamics of Complex Systems*, chapter 3 Neural Networks II: Models of Mind, (pp. 371–419). Addison-Wesley. 4, 4.2

Bar-Yam, Y. (1997b). *Dynamics of Complex Systems*. Addison-Wesley. 4.1

Bar-Yam, Y. (1997c). *Dynamics of Complex Systems*, chapter 2 Neural Networks I: Subdivision and Hierarchy, (pp. 295–370). Addison-Wesley. 4.2

Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adaptive Behavior*, 3(4), 469–509. 6.2

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3), 91–99. 4.3, 9.1

Beer, R. D. (2003). The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4), 209–243. 4.3, 7.5

Bekey, G. A. (2006). *Autonomous Robots: from biological inspiration to implementation and control*. MIT Press. 2, 3.1, 9.1

Bianco, R. & 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. 4.2, 4.3, 6.5, 9.1

Bicchi, A. & Kumar, V. (2000). Robotic grasping and contact: A review. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 1 (pp. 348–353). 3.1

Bongard, J. (2010). The utility of evolving simulated robot morphology increases with task complexity for object manipulation. *Artificial Life*, 16(3), 201–223. 4.2, 6.5, 9.1

Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*. The MIT Press. 4.1

Buehrmann, T. & Paolo, E. A. D. (2004). Closing the loop: Evolving a model-free visually-guided robot arm. In *Artificial life IX: Proceedings of the Ninth International Conference on the Synthesis of Living Systems* (pp. 63–68). 4.2, 4.3, 6.5, 9.1

Buehrmann, T. & Paolo, E. A. D. (2006). Biological actuators are not just springs. In *From Animals to Animats 9*, volume 4095 of *Lecture Notes in Computer Science* (pp. 89–100). Springer. 5.5

Butterfass, J., Grebenstein, M., Liu, H., & Hirzinger, G. (2001). Dlr hand ii: Next generation of a dextrous robot hand. *Proceeding of IEEE International Conference on Robotics and Automation*, (pp. 109–114). 3

Cangelosi, A., Bugmann, G., & Borisyuk, R., Eds. (2005). *Modeling Language, Cognition and Action: Proceedings of the 9th Neural Computation and Psychology Workshop*, Singapore. World Scientific. 1

Cangelosi, A., Metta, G., Sagerer, G., Nolfi, S., Nehaniv, C. L., Fischer, K., Tani, J., Sandini, G., Fadiga, L., Wrede, B., Rohlfing, K., Tuci, E., Dautenhahn, K., Saunders, J., & Zeschel, A. (2010). Integration of action and language knowledge: A roadmap for developmental robotics. *IEEE Transactions on Autonomous Mental Development*, 2(3), 167–195. 4.4, 9.1

Cangelosi, A., Tikhanoff, V., Fontanari, J. F., & Hourdakis, E. (2007). Integrating language and cognition: A cognitive robotics approach. *IEEE Computational Intelligence Magazine*, 2(3), 65–70. 4.4, 9.1

Cappa, S. F. & Perani, D. (2003). The neural correlates of noun and verb processing. *Journal of Neurolinguistics*, 16(2-3), 183–189. 4.4, 9.1

Cliff, D., Husbands, P., & Harvey, I. (1993). Explorations in evolutionary robotics. *Adaptive Behavior*, 2(1), 73–110. 4.1

Cohen, H. & Lefebvre, C. (2005). *Handbook of Categorisation in Cognitive Science*. Elsevier. 4.3, 9.1

Corballis, M. C. (2003). *From Hand to Mouth: the Origins of Language*. Princeton University Press. 1

Dario, P., Laschi, C., Carrozza, M., Guglielmelli, E., Teti, G., Massa, B., Zecca, M., Taddeucci, D., & Leoni, F. (2000). An integrated approach for the design and development of a grasping and manipulation system in humanoid robotics. In *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1 (pp. 1–7). 4.3

Darwin, C. (1859). *On The Origin of Species*. John Murray. 4.1

Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*. John Wiley. 4.2

Feix, T., Pawlik, R., Schmiedmayer, H.-B., Romero, J., & Kragić, D. (2009). The generation of a comprehensive grasp taxonomy. http://www.csc.kth.se/grasp/taxonomyGRASP.pdf. 3

Felip, J. & Morales, A. (2009). Robust sensor-based grasp primitive for a three-finger robot hand. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1811–1816). 3.2, 4

Floreano, D., Husband, P., & Nolfi, S. (2008). Evolutionary robotics. In *Handbook of Robotics* (pp. 1423–1451). Springer. 4.1

Floreano, D. & Mondada. . . , F. (1994). Automatic creation of an autonomous agent: Genetic evolution of a neural-network driven robot. In *From Animals to Animat 3: Proceedings of third Conference on Simulation of Adaptive Behavior*. 4.1

Gallese, V. & Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, 22(3-4), 455–479. 1

Gialias, N. & Matsuoka, Y. (2004). Muscle actuator design for the act hand. In *Proceedings of the IEEE International Conference on Robotics and Automation* (pp. 3380–3385). 4

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, Acting and Knowing. Toward an Ecological Psychology* (pp. 67–82). John Wiley. 1

Gienger, M., Toussaint, M., Jetchev, N., Bendig, A., & Goerick, C. (2008). Optimization of fluent approach and grasp motions. In *Proceedings of 8th IEEE/RAS International Conference on Humanoid Robots* (pp. 111–117). 1

Gigliotta, O. & Nolfi, S. (2008). On the coupling between agent internal and agent/environmental dynamics: Development of spatial representations in evolving autonomous robots. *Adaptive Behavior*, 16(2-3), 148–165. 4.3

Glenberg, A. & Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, 9(3), 558–565. 4.4, 9.1

Gomez, G., Hernandez, A., Hotz, P. E., & Pfeifer, R. (2005). An adaptive learning mechanism for teaching a robotic hand to grasp. In *Proceedings of International Symposium on Adaptive Motion of Animals and Machines*. 6.5, 9.1

Greenhill, R. (2010). Shadow hand. http://www.shadowrobot.com/hand. 3

Harnad, S. R. (1987). *Categorical Perception: The Groundwork of Cognition.* Cambridge University Press. 4.3

Harvey, I., Paolo, E., Wood, R., Quinn, M., & Tuci, E. (2005). Evolutionary robotics: A new scientific tool for studying cognition. *Artificial Life*, 11(1-2), 79 – 98. 4.1

Hauk, O., Johnsrude, I., & Pulvermuller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, 41(2), 301–307. 4.4, 9.1

Haykin, S. (1999). *Neural Networks: a comprehensive foundation.* Pearson Prentice Hall, 2nd edition edition. 4

Hofsten, C. V. (1982). Eye-hand coordination in the newborn. *Developmental Psychology*, 18(3), 450–461. 4

Hofsten, C. V. (1984). Developmental changes in the organization of prereaching movements. *Developmental Psychology*, 20(3), 378–388. 4

Hofsten, C. V. (1991). Structuring of early reaching movements: a longitudinal study. *Journal of Motor behavior*, 23, 280–292. 4

Iberall, T. & Arbib, M. A. (1990). Schemas for the control of hand movements: an essay on cortical localization. In M. A. Goodale (Ed.), *Vision and action: the control of grasping.* Ablex Publishing Corporation. 3

Iossifidis, I. & Schoner, G. (2006). Dynamical systems approach for the autonomous avoidance of obstacles and joint-limits for an redundant robot arm. In *Proceedings of the IEEE 2006 International Conference on Intelligent Robots and Systems (IROS).* 3

Jerez, J. & Suero, A. (2004). Newton game dynamics. http://www.newtondynamics.com. A.2, B, C

Johnsson, M. & Balkenius, C. (2006). A robot hand with t-mpsom neural networks in a model of the human haptic system. In *Proceedings of TAROS* (pp. 80–87). 4.3

Johnsson, M. & Balkenius, C. (2007a). Experiments with proprioception in a self-organizing system for haptic perception. In *Proceedings of TAROS 2007* (pp. 239–245). 4.3

Johnsson, M. & Balkenius, C. (2007b). Neural network models of haptic shape perception. *Journal of Robotics and Autonomous Systems*, 55(9), 720–727. 4.3

Johnsson, M., Pallbo, R., & Balkenius, C. (2005). Experiments with haptic perception in a robotic hand. In *Advances in Artificial Intelligence* (pp. 81–86). Malardalen University. 4.3

Jones, L. A. & Lederman, S. J. (2006). *Human Hand Function*. Oxford University Press. 2, 2, 4, 6.2, 6.5, 7.2

Kawato, M. (2003). Cerebellum and motor control. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd Edition* (pp. 190–195). MIT Press. 1, 6.5

Kokera, R., Oz, C., Cakar, T., & Ekiz, H. (2004). A study of neural network based inverse kinematics solution for a three-joint robot. *Robotics and Autonomous Systems*, 49, 227–234. 3.1, 9.1

Krose, B. J. A. & der smagt, P. P. V. (1993a). *An Introduction to Neural Networks*. University of Amsterdam, 5th edition edition. 3.1, 9.1

Krose, B. J. A. & der smagt, P. P. V. (1993b). *An Introduction to Neural Networks*, chapter Robot Control, (pp. 79–92). University of Amsterdam. 4.2

Li, Y. & Leong, S. H. (2003). A cmac neural network approach to redundant manipulator kinematics control. *Journal of Mechanical Engineering*, 54(2), 65–81. 3.1, 9.1

Lungarella, M. & Metta, G. (2003). Beyond gazing, pointing, and reaching: A survey of developmental robotics. In *Proceedings of International Conference on Epigenetic Robotics '03* (pp. 81–89). 6.5

Manocha, D. & Zhu, Y. (1994). A fast algorithm and system for the inverse kinematics of general serial manipulators. In *Proceedings of IEEE International Conference on Robotics and Automation*, volume 4 (pp. 3348–3353). 3.1, 9.1

Marocco, D., Cangelosi, A., & Nolfi, S. (2003). The role of social and cognitive abilities in the emergence of communication: Experiments in evolutionary robotics. In *EPSRC/BBSRC International Workshop Biologically-Inspired Robotics* (pp. 174–181). 4.2, 4.3

Martìn, P. & del R Del Milla, J. (1998). Learning reaching strategies through reinforcement for a sensor-based manipulator. *Neural Networks*, 11, 359–376. 3.1, 9.1

Massera, G., Cangelosi, A., & Nolfi, S. (2006). Developing a reaching behaviour in a simulated anthromorphic robotic arm through an evolutionary technique. In L. M. Rocha (Ed.), *Artificial Life X: Proceeding of the Tenth International Conference on the simulation and synthesis of living systems.* 4, 4.2, 6.5, 9.1

Massera, G., Cangelosi, A., & Nolfi, S. (2007). Evolution of prehension ability in an anthropomorphic neurorobotic arm. *Frontiers in neurorobotics*, 1, 1–9. 4, 4.1

Massera, G., Nolfi, S., & Cangelosi, A. (2005). Evolving a simulated robotic arm able to grasp objects. In *Modeling Languange, Cognition and Action: Proceeding of the Ninth Neural Computation and Psychology Workshop*, volume 16 of *Progress in Neural Processing*. Singapore World Scientific. 4.2

Massera, G., Tuci, E., Ferrauto, T., & Nolfi, S. (2010). The facilitatory role of linguistic instructions on developing manipulation skills. *IEEE Computational Intelligence Magazine*, 5(3), 33–42. 4, 4.1

McCarty, M. K., Clifton, R. K., Ashmead, D. H., Lee, P., & Goulet, N. (2001). How infants use vision for grasping objects. *Child Development*, 72(4), 973–987. 4

Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., & Wright, C. E. (1990). Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod (Ed.), *Motor Representation and Control*, volume 13 of *Attention and Performance* (pp. 173–226). Collection. 2

Mial, R. C. (2003). Motor control, biological and theoretical. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Network, 2nd Edition* (pp. 686–689). MIT Press. 3

Mirolli, M. & Parisi, D. (2009). Towards a vygotskyan cognitive robotics: The role of language as a cognitive tool. *New Ideas in Psychology*. 9.2

Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. MIT Press. 7.3

Mohan, V., Zenzeri, J., Morasso, P., & Metta, G. (2009). Equilibrium point hypothesis revisited: Advances in the computational framework of passive motion paradigm. In *Advanced Computational Motor Control Conference*. 9.2

Natale, L. & Torres-Jara, E. (2006). A sensitive approach to grasping. In *Proceedings of the sixth international workshop on Epigenetic Robotics* (pp. 87–94). 4.3

Noe, A. (2004). *Action in Perception*. MIT Press. 1

Nolfi, S. (2002). Power and limits of reactive agents. *Neurocomputing*, 42(1-4), 119–145. 4.3

Nolfi, S. (2005a). Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour. *ComplexUS*, 2(3-4), 195–203. 1, 4.1, 4.4

Nolfi, S. (2005b). Categories formation in self-organizing embodied agents. In H. Cohen & C. Lefebvre (Eds.), *Handbook of Categorization in Cognitive Science* (pp. 869–889). Elsevier. 4.4

Nolfi, S. (2009). Behavior and cognition as a complex adaptive system: Insights from robotic experiments. In C. Hooker (Ed.), *Handbook of the Philosophy of Science*, volume 10 of *Philosophy of Complex Systems*. 4.3, 9.1

Nolfi, S. & Floreano, D. (2000). *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press. 4.1, 5, 5.3, 6.3, 8.3

Nolfi, S., Floreano, D., Miglino, O., & Mondada, F. (1994). How to evolve autonomous robots: Different approaches in evolutionary robotics. In R. Brooks & P. Maes (Eds.), *Proceedings of the International Conference Artificial Life IV* (pp. 190–197). 4.1

Nolfi, S. & Marocco, D. (2001). Evolving robots able to integrate sensory-motor information over time. *Theory in Biosciences*, 120(3-4), 287–310. 6.2, 8.2

Nolfi, S. & Marocco, D. (2002). Active perception: A sensorimotor account of object categorization. In B. Hallam, D. Floreano, J. Hallam, G. Hayes, & J.-A. Meyer (Eds.), *From Animals to Animats 7: Proceedings of the Seventh International Conference on Simulation of Adaptive Behavior* (pp. 266–271). 4.2, 4.3, 7.5

Oyama, E., Agah, A., MacDorman, K. F., Maeda, T., & Tachi, S. (2001). A modular neural network architecture for inverse kinematics model learning. *Neurocomputing*, 38-40, 797–805. 3.1, 9.1

Oztop, E., Bradley, N. S., & Arbib, M. A. (2004). Infant grasp learning: a computational model. *Experimental Brain Research*, 158(4), 480–503. 3, 4

Page, R. E. (1998). The structure of the hand. In K. J. Connolly (Ed.), *The Psychobiology of the Hand* (pp. 1–15). Mac Keith Press. 2, 6.2, 7.2

Pfeifer, R. & Scheier, C. (1999). *Understanding Intelligence*. MIT Press. 2

Platt, M. (2002). Neural correlates of decisions. *Current Opinion in Neurobiology*, 12(2), 141–148. 7.5

Pulvermuller, F. (2002). *The Neuroscience of Language. On Brain Circuits of Words and Serial Order*. Cambridge University Press. 4.4, 9.1

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nature review Neuroscience*, 6, 576–582. 1

Rathbone, K. & Sharkey, N. (1999). Evolving robot arm controllers for continued adaptation. In *Proceedings of IEEE International Symposium on Computational Intelligence in Robotics and Automation* (pp. 345–350). 3.1, 9.1

Rizzolatti, G. & Arbib, M. (1998). Language within our grasp. *Trends in neurosciences*, 21, 188–194. 1, 4.4, 9.1

Rizzolatti, G. & Craighero, L. (2004). The mirror-neuron system. In *Annual Review of Neuroscience*, volume 27 (pp. 169–192). Harvard University. 1

Rochat, P. (1998). Self-perception and action in infancy. *Experimental Brain Research*, 123(1-2), 102–109. 4

Salisbury, J. K. (1982). *Kinematic and Force Analysis of Articulated Hands*. thesis, Stanford University. 3

Sandercock, T. G., Lin, D. C., & Rymer, W. Z. (2003). Muscle models. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd Edition* (pp. 711–715). MIT Press. 1, 6.1.2, 7.1.2, 8.1.2, B.2

Sandini, G., Metta, G., & Vernon, D. (2004). Robotcub: An open framework for research in embodied cognition. In *Proceedings of IEEE/RAS International Conference on Humanoid Robots* (pp. 13–32). 4, 8.1.4, C.4

Schaal, S. (2003). Arm and hand movement control. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Network, 2nd Edition* (pp. 110–113). MIT Press. 1, 2

Schaal, S., Peters, J., Nakanishi, J., & Ijspeert, A. (2005). Learning movement primitives. *Robotics Research*, 15, 561–572. 4

Scheier, C. & Lambrinos, D. (1996). Categorization in a real-world agent using haptic exploration and active perception. In *Proceedings of International Conference SAB '96* (pp. 65–75). 4.3

Scheier, C., Pfeifer, R., & Kunyioshi, Y. (1998). Embedded neural networks: exploiting constraints. *Neural Networks*, 11, 1551–1569. 4.3, 7.5

Schoner, G. & Santos, C. (2001). Control of movement time and sequential action through attractor dynamics: A simulation study demonstrating object interception and coordination. In *Proceedings of the 9th Intelligent Symp. On Intelligent Robotic Systems* (pp. 18–20). 3

Shadmehr, R. (2003). Equilibrium point hypothesis. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Network, 2nd Edition* (pp. 409–412). MIT Press. 1, 3, 5.5, 6.2, 6.5

Shadmehr, R. (2004). *The Computational Neurobiology of Reaching and Pointing*, chapter Supplement: A Mathematical Muscle Model. Computational Neurobiology of Reaching and Pointing. 4

Shadmehr, R. & Wise, S. P. (2005a). *The Computational Neurobiology of Reaching and Pointing*, chapter What generates force and feedback, (pp. 93–118). MIT Press. 1, 5.5

Shadmehr, R. & Wise, S. P. (2005b). *The Computational Neurobiology of Reaching and Pointing*. MIT Press. 2, 2, 4, 6.1.2, 6.5, 7.1.2, 8.1.2, B.2

Shadmehr, R. & Wise, S. P. (2005c). *The Computational Neurobiology of Reaching and Pointing*, chapter What maintains limb stability, (pp. 119–140). MIT Press. 4, 5.5

Siciliano, B. & Khatib, O. (2008). *Handbook of Robotics*. Springer. 3, 3, 3.1, 3.1, 3.2, 9.1

Smith, R. (2004). Open dynamics engine. http://www.ode.org. A

Stansfield, S. (1991). A haptic system for a multifingered hand. In *Proceeding of IEEE International Conference on Robotics and Automation* (pp. 658–664). 4.3

Sternad, D. & Schaal, S. (1999). Segmentation of endpoint trajectories does not imply segmented control. *Experimental Brain Research*, 124, 118–136. 5.5

Suárez, R., Roa, M., & Cornella, J. (2006). *Grasp quality measures.* Technical report, Robòtica Industrial i de Sistemes. 3, 3.1

Takamuku, S., Gomez, G., Hosoda, K., & Pfeifer, R. (2007). Haptic discrimination of material properties by a robotic hand. In *Proceedings of 6th IEEE International Conference on Development and Learning* (pp. 1–6). 4.3

Thelen, E., Schoner, G., Scheier, C., & Smith, L. B. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, (24), 1–86. 3

Thornton, C. (1997). Separability is a learner's best friend. In *Proceedings of the 4th Neural Computation and Psychology Workshop*. 7.4.3

Toal, D. & Flanagan, C. (2002). 'pull to position', a different approach to the control of robot arms for mobile robots. *Journal of Materials Processing Technology*, 123, 393–398. 3.1, 9.1

Torras, C. (2003). Robot arm control. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks, 2nd Edition* (pp. 979–983). MIT Press. 3, 3.2, 4.2, 5.5

Townsend, B. (2010). Barrett hand. http://www.barrett.com/robot/products-hand.htm. 3

Townsend, J. T. & Busemeyer, J. (1995). Dynamic representation of decision-making. In R. F. Port & T. V. Gelder (Eds.), *Mind as motion: explorations in the dynamics of cognition* (pp. 101–120). MIT Press. 7.5

Tuci, E., Massera, G., & Nolfi, S. (2010). Active categorical perception of object shapes in a simulated anthropomorphic robotic arm. *IEEE Transaction on Evolutionary Computation*, 14(6), 1–15. 4, 4.1

Tuci, E., Trianni, V., & Dorigo, M. (2004). Feeling the flow of time through sensory-motor coordination. *Connection Science*, 16(4), 301–324. 4.3

Veber, M., Dolanc, M., & Bajd, T. (2005). Optimal grasping in humans. *Journal of Automatic Control*, 15(supplement), 15–18. 3.1

Vygotsky, L. S. (1962). *Thought and language*. MIT Press. 9.2

Vygotsky, L. S. (1978). *Mind in society*. Harvard College. 9.2

Waxman, A. (2003). Sensor fusion. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd Edition* (pp. 1014–1016). MIT Press. 7.5

Weng, J. (2004). Developmental robotics: Theory and experiments. *International Journal of Humanoid Robotics*, 1(2), 199–236. 4.4

Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., & Thelen, E. (2001). Autonomous mental development by robots and animals. *Science*, 291(5504), 599–600. 4.4

Williamson, M. M. (1998). Neural control of rhythmic arm movements. *Neural Networks*, 11, 1379–1394. 3.1, 9.1

Wolpert, D. M. & Flanagan, J. R. (2003). Sensorimotor learning. In M. A. Arbib (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd Edition* (pp. 1020–1023). MIT Press. 1

Yade, T. (2004). Yet another dynamic engine. http://developer.berlios.de/projects/yade. A.2

Yao, X. & Islam, M. M. (2008). Evolving artificial neural network ensembles. *IEEE Computational Intelligence Magazine*, 3(1), 31–42. 8.3

Yasumuro, Y., Chen, Q., & Chihara, K. (1999). Three-dimensional modeling of the human hand with motion constraints. *Image and Vision Computing*, 17, 149–156. 6.2, 7.2

Appendices

A Robotic Arm Version A

The first version of the anthropomorphic robotic arm is provided with only 4 DOFs. The arm and the arm/environmental interaction were simulated using ODE (Open Dynamics Engine, Smith, 2004), a library for the accurate simulation of rigid body dynamics and collisions.

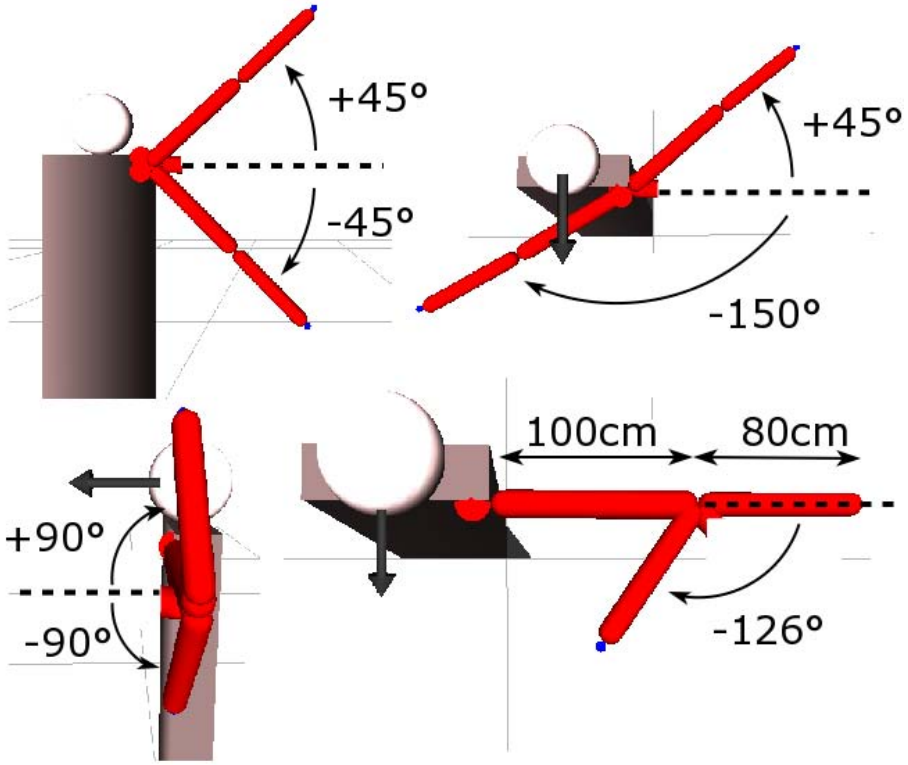


Figure A.1: Structure of the 4-DOF robotic arm. The four DOFs of the simulated robotic arm. The two illustrations at the top of the figure indicate the abduction/adduction (left) and extension/flexion of the shoulder joint (right). The bottom figures indicate the rotation of the shoulder (left) and the extension/flexion of the elbow (right). In all illustrations, the arrows indicate the frontal direction of the robot.

A.1 Arm Structure and Actuators

The simulated robot consists of cylindrical segments articulated by revolute joints, as illustrated in Figure A.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 elbow joints). The arm and the forearm have lengths of $100\,cm$ and $80\,cm$, diameters of $8\,cm$ and $7\,cm$, and weights of $13\,kg$ and $8\,kg$ respectively. The shoulder has three DOFs that allow abduction/adduction of $[-45°, +45°]$, extension/flexion of $[-150°, +45°]$ and rotation of $[-90°, +90°]$. The elbow has one DOF that allows extension/flexion of $[-126°, +0°]$. Since the robot is only asked to reach a given target position with the endpoint of its arm, we did not model the wrist and the wrist joints. Therefore the arm has four motorised joints and four DOFs, see Figure A.1. The joints are moved by directly setting the desired angular velocity specified by the neural network. The maximum velocity at which each joint of the arm can be set is $890\,rpm$. The acceleration of gravity was set to $9.8\,m/s^2$.

A.2 The Issue of Physics Engines

This section explains the reasons behind the move from Open Dynamics Engine to Newton Game Dynamics. In fact, the latter was employed to implement the full anthropomorphic robotic arms described in appendices B and C.

One of the most important issues in being able to obtain results relevant to robotics, is being able to develop an accurate model of the rigid-body dynamics. In this regard there are many commercial libraries that work quite well and serve most needs, but the high cost of these products makes them quite difficult to use in academic research. The alternative is to use open-source or free libraries, or commercial packages such as MatLab.

The most important factor in the evolution of agents in a physical environment using the er approach is the speed of the engine. In fact, rigid-body dynamics libraries such as YADE (Yade, 2004) or MatLab simulate interactions quite accurately, but using them, it is practically impossible to evolve agents due to the long time required to complete an evolutionary process. Hence, the engine must have the ca-

pacity to simulate the world at a rate that is faster than real-time. This feature is accomplished by an engine fitted for video games.

At the beginning of the author's Ph.D. studies, the ODE library was chosen due to its free GPL license. Although the ODE project began in 2004, it is still at the stage of beta release. The results presented in Chapter 5 were achieved using ODE. This was possible because the simulation was simple enough that no problem arose on account of the library. In the implementation of muscle actuators, however, numerous problems arose that were related to the bugs and the lack of features in the available version of ODE.

Hence, another physical engine, Newton Game Dynamics (NGD), was chosen, and the simulator was adapted to this new library (Jerez & Suero, 2004). All the new models developed to that point were replicated, and all evolutionary processes were re-run in order to verify the correct porting of the simulator to NGD.

The choice of NGD was made taking into account the simulation of grasping. When the hand touches an object in order to grasp it, numerous forces are generated, and the correct simulation of these forces is fundamental to obtaining valid results after the evolutionary process. Friction and gravity are critical from the point of view of grasping, and neither of these forces are not simulated as well in ODE as they are in NGD.

B Robotic Arm Version B

The second version of the robot arm that was implemented is a full anthropomorphic manipulator with a five-fingered hand attached to a 7-DOF arm. The actuation of the arm's joints is performed by muscle-like actuators, while the fingers are controlled by their velocity, which is performed by a simple position controller, in order to simplify the complexity. The arm and the arm/environmental interaction were simulated using NGD (Newton Game Dynamics, Jerez & Suero, 2004), a library for the accurate simulation of rigid body dynamics and collisions.

B.1 Arm Structure

The arm (Figure B.1) consists mainly of three elements (the arm, the forearm and the wrist) that are connected through articulations displaced into the shoulder, the arm, the elbow, the forearm and wrist. It is an enhancement of a previous 4-DOF model to which a wrist comprised of another 3-DOF joint was added. The wrist adds the ability to produce pitch, yaw and roll of the end-effector of the arm (i.e. the hand that will be added in a further step). This is the first step toward the addition of a hand for the purpose of studying grasping behaviours.

The shoulder is composed of a sphere with a radius 2.8 cm. The lengths of the arm and forearm are 23 and 18 cm, respectively. The wrist consists of an ellipsoid with a radius of 1.45, 1.2 and 1.45 cm along the x-, y- and z-axis, respectively.

The joints J_1, J_2 and J_3 (Figure B.1) provide *abduction/adduction*, *extension/flexion* and *supination/pronation* of the arm in the range of $[-140°, +60°]$, $[-90°, +90°]$ and $[-60°, +90°]$, respectively. These three DOFs act like a ball-and-socket joint moving the arm in a way analogous to the human shoulder joint. Joint J_4, which is located in the elbow, consists of a hinge joint that provides *extension/flexion* within a range

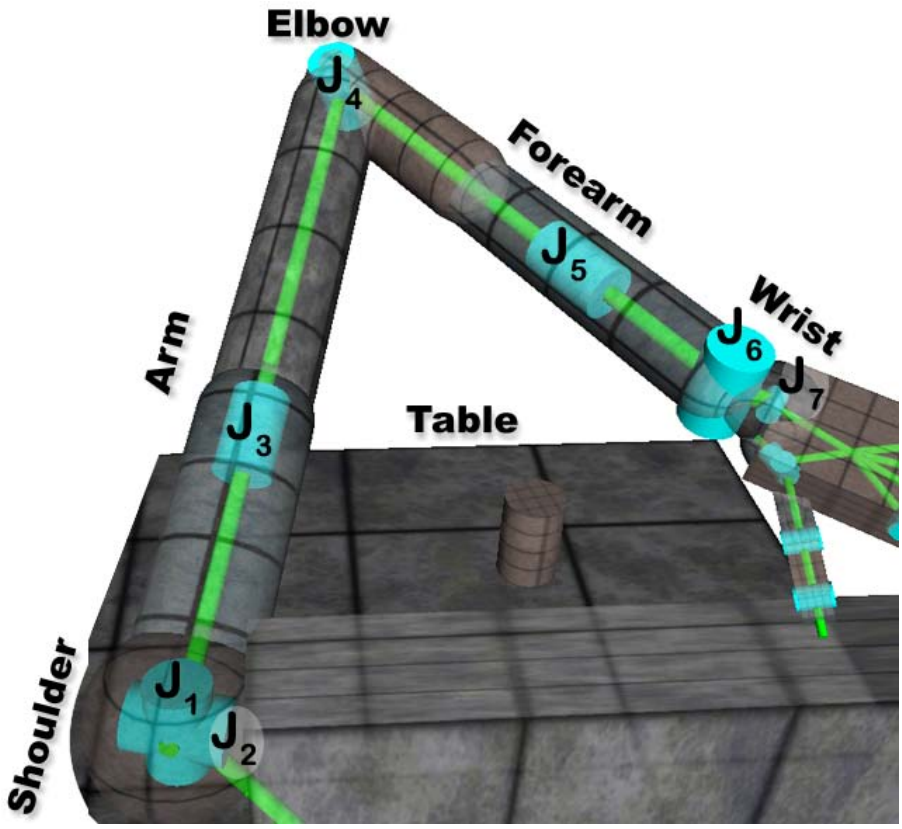


Figure B.1: The kinematic chain of the arm. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation. The links among the cylinders represent the rigid connections that make up the arm structure.

of $[-170°, +0°]$ (the radius and ulna bones). Joint J_5 rotate the forearm providing *pronation/supination* of the wrist (and the palm) within a range of $[-90°, +90°]$. Joints J_6 and J_7 on the wrist provide *flexion/extension* and *abduction/adduction* of the hand within a range of $[-30°, +30°]$ and $[-90°, +90°]$, respectively.

B.2 Arm Actuators

The arm joints $(J_1, \ldots, J_7)$ are actuated by two simulated antagonist muscles that were implemented accordingly to Hill's muscle model (Sandercock et al., 2003; Shadmehr & Wise, 2005b). More precisely, the total force exerted by a muscle (Figure B.2) is the sum of three forces $T_A(\alpha, x) + T_P(x) + T_V(\dot{x})$ which depend on the activity of the corresponding motor neuron (α) on the current elongation of the muscle

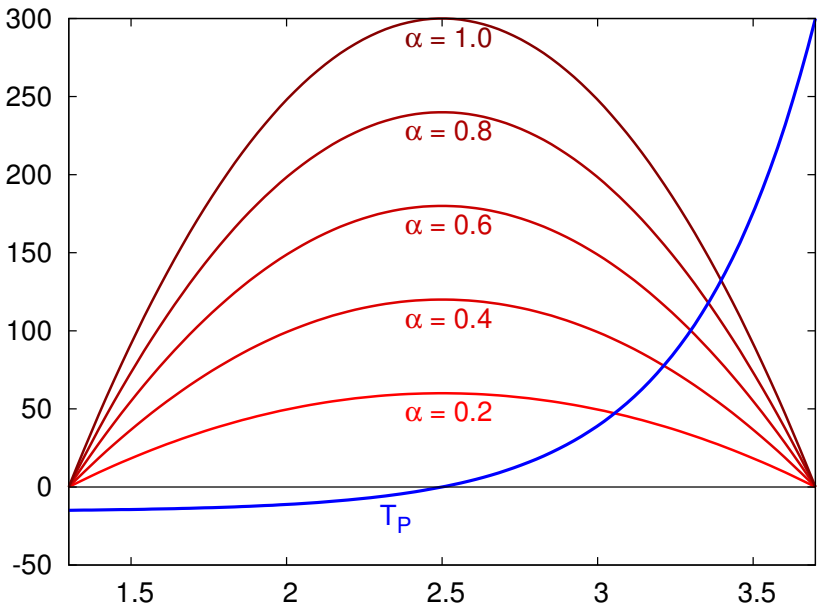


Figure B.2: An example of the force exerted by a muscle. The graph shows how the force exerted by a muscle varies as a function of the activity of the corresponding motor neuron and of the elongation of the muscle for a joint in which T_{max} is set to $300\,N$.

(x) and on the muscle contraction/elongation speed $(\dot{x})$ which are calculated on the basis of the following equations:

$$\begin{aligned} T_A &= \alpha\left(-\frac{A_{sh}T_{max}(x-R_L)^2}{R_L^2}+T_{max}\right) \\ A_{sh} &= \frac{R_L^2}{(L_{max}-R_L)^2} \\ T_P &= T_{max}\frac{exp\left\{K_{sh}\left(\frac{x-R_L}{L_{max}-R_L}\right)\right\}-1}{exp\{K_{sh}\}-1} \\ T_V &= b\cdot\dot{x} \end{aligned} \tag{B.1}$$

where L_{max} and R_L are the maximum and the resting length of the muscle, T_{max} is the maximum force that can be generated, K_{sh} is the passive shape factor, and b is the viscosity coefficient.

The active force T_A depends on the activation of muscle α and on the current elongation/compression of the muscle. When the muscle is completely elongated/compressed, the active force is zero, regardless of activation α. At the resting length R_L, the active force reaches its maximum, which depends on activation α. The red curves in Figure B.2 show how the active force T_A changes with respect to the elongation

of the muscle for some possible values of α. The passive force T_P depends only on the current elongation/compression of the muscle (see the blue curve in figure B.2). T_P tends to elongate the muscle when it is compressed to a degree less than R_L, and tends to compress the muscle when it is elongated beyond R_L. T_P differs from a linear spring in that it has an exponential trend that produces a strong opposition to muscle elongation and little opposition to muscle compression. T_V is the viscosity force. It produces a force that is proportional to the velocity of the elongation/compression of the muscle.

The parameters of the equation are identical for all 14 muscles that control the seven dofs of the arm, and they were set to the following values: $K_{sh} = 3.0$, $R_L = 2.5$, $L_{max} = 3.7$, $b = 0.9$, $A_{sh} = 4.34$ with the exception of parameter T_{max} which was set to $3000N$ for joint J_2, $300N$ for joints J_1, J_3, J_4, and J_5, and $200N$ for J_6 and J_7.

Muscle elongation is simulated on the basis of the actual angular position of each DOF, which is mapped linearly within the allowable angular range of each DOF. For instance, in the case of the elbow where the limits are $[-170^o, +0^o]$, this range is mapped onto $[+1.3, +3.7]$ for the agonist muscle, and inversely, onto $[+3.7, +1.3]$ for the antagonist muscle. Hence, when elbow is completely extended (angle 0), the agonist muscle is completely elongated (3.7) and the antagonist is completely compressed (1.3), and vice versa when the elbow is flexed.

B.3 Hand Structure

The hand was added to the robotic arm just below the wrist (at joint J_7, as shown in Figure B.1).

The robotic hand (Figure B.3) is composed of a palm and 14 phalange segments (Table B.1) that make up the digits (two for the thumb, and three for each of the other four fingers), which are connected through 15 joints with 20 DOFs. The palm

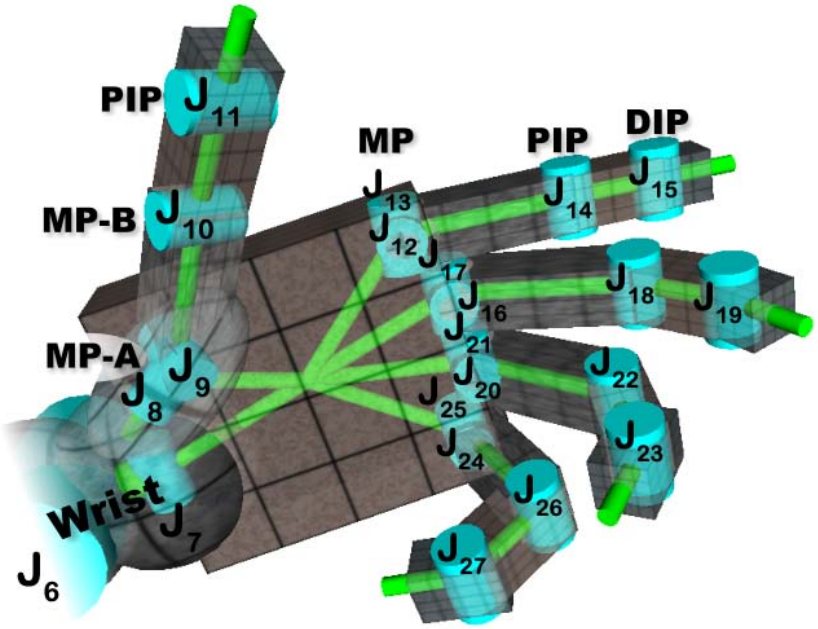


Figure B.3: The kinematic chain of the hand. The cylinders represent rotational DOFs. The axes of the cylinders indicate the corresponding axis of rotation, and the labels on the cylinders are the names of the joints. The links among the cylinders represent the rigid connections that make up the hand structure.

consists of a box with dimensions of $4.6 \times 1.2 \times 4.2\,cm$. The thumb is composed of four connected objects:

1. an ellipsoid with a radius of $1.5 \times 0.6 \times 0.8\,cm$ that is half-sunk into the palm;
2. a box with dimensions of $2.4 \times 0.8 \times 0.9\,cm$, which corresponds to the metacarpal bones of the human thumb;
3. a box with dimensions of $1.6 \times 0.75 \times 0.85\,cm$, which corresponds to the first phalanx; and
4. a box with dimensions of $1.12 \times 0.75 \times 0.8\,cm$, which corresponds to the second phalanx.

The other fingers are connected to the palm through the knuckles, which are represented by an ellipsoid with radius dimensions of $0.65 \times 0.65 \times 0.5\,cm$. The three phalanges that compose each finger are boxes that are jointed serially.

Finger	First phalanx	Second phalanx	Third phalanx
Index	2.40×0.80×0.80	1.50×0.75×0.75	1.12×0.70×0.70
Middle	2.62×0.80×0.80	1.50×0.75×0.75	1.12×0.70×0.70
Ring	2.40×0.80×0.80	1.50×0.75×0.75	1.12×0.70×0.70
Pinky	2.25×0.80×0.80	1.50×0.75×0.75	1.12×0.70×0.70

Table B.1: Size of the segments forming the hand (in cm).

The joints in the hand are grouped into *metacarpophalangeal* ($MP = \{J_{12}, J_{13}, J_{16}, J_{17}, J_{20}, J_{21}, J_{24}, J_{25}\}$, $MP\text{-}A = \{J_8, J_9\}$ and $MP\text{-}B = \{J_{10}\}$), *proximal interphalangeal* ($PIP = \{J_{11}, J_{14}, J_{18}, J_{22}, J_{26}\}$), and *distal interphalangeal* ($DIP = \{J_{15}, J_{19}, J_{23}, J_{27}\}$) types. Each finger has two hinge joints, PIP and DIP (see Figure B.3), that *extend/flex* the phalanges within the range of $[-90°, +0°]$. The MP group is composed of two joints that allow both *extension/flexion* and *abduction/adduction* of the first phalanx of each finger. The *extension/flexion* of MP is in the range of $[-90°, +0°]$ for all fingers, but the range of *abduction/adduction* varies for different fingers, and corresponds to $[-7°, +0°]$, $[-2°, +2°]$, $[-2°, +5°]$ and $[+0°, +7°]$ for the index, the middle, the ring and the little fingers, respectively. The thumb does not have a DIP joint, and the MP provides three DOFs, which are located in the MP-A and MP-B joints. The former joint has two DOFs, which provide *supination/pronation* and *abduction/adduction* of the metacarpal part of the thumb in ranges of$[-120°, +0°]$ and $[-15°, +90°]$, respectively, which allows for good opposition of the thumb to the fingers. The MP-B and PIP joints consists of hinge joints that *extend/flex* the first and second phalanx of the thumb in the same range $[-90°, +0°]$ as the PIP and DIP joints of the other four fingers.

B.4 Hand Actuators

The joints are controllable independent of one another by specifying the desired position. One of the most important features of the hand's joints is their compliance, which facilitates the grasping of objects. This was obtained using elastic actuators. In detail, the compliance was obtained by setting a maximum threshold of $300\,N$

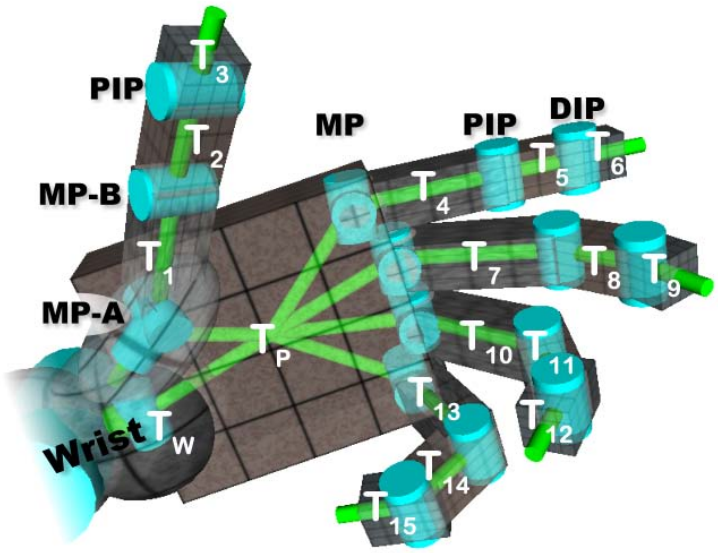


Figure B.4: Distribution of tactile sensors on the hand. The links among the cylinders represents the rigid connections that constitute the hand structure, and the white labels on the links indicate the names and the positions of the tactile sensors.

for the force exerted by each joint. When an external force acting on a joint exceeds this threshold, the joint either cannot move further, or it moves backward due to the external force. The joints are moved by a proportional controller that sets the angular velocity of a joint in order for it to reach the position specified by the neural network.

B.5 Hand Tactile Sensors

The hand is equipped with tactile sensors that are distributed over the wrist, the palm, and all five fingers. Figure B.4 shows where the tactile sensors are placed. The white labels indicate the names of the tactile sensors. Each tactile sensor simply counts the number of contacts that take place on the part corresponding to the one it is placed on. The contacts coming from the humanoid parts are not counted. For example, in the case of T_P, the sensor reports all contacts between the palm and another object(s), but not the contacts between the palm and the fingers.

C Robotic Arm Version C

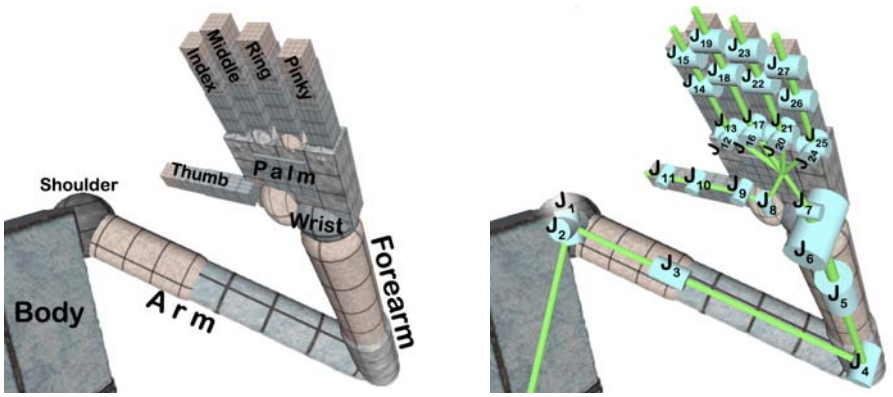
The third version of the robot arm that was implemented is a modified version of the previous one. It is a full anthropomorphic manipulator with a five-fingered hand attached to a 7-DOF arm. The actuation of the arm's joints is performed by muscle-like actuators (as in Version B), while the fingers are controlled as a unit, using the same grouping principle that was used in the development of the iCub hand. The arm and the arm/environmental interactions were simulated using ngd (Newton Game Dynamics, Jerez & Suero, 2004), a library for the accurate simulation of rigid body dynamics and collisions.

C.1 Arm Structure

The arm consists mainly of three elements (the arm, the forearm, and the wrist), which are connected through articulations placed in the shoulder, the arm, the elbow, the forearm, and wrist (see Figure C.1-a). The dimensions of the elements and the distribution of the joints are the same as in Version B; see Appendix B for details. However, the angles of the joints vary, and move in slightly different ranges than they do in Version B. The angles of joints $J_1, \ldots J_7$ vary, respectively, having ranges of $[-140°, +100°]$, $[-110°, +90°]$, $[-110°, +90°]$, $[-170°, +0°]$, $[-100°, +100°]$, $[-40°, +40°]$ and $[-100°, +100°]$ (see Figure C.1-b).

C.2 Arm Actuators

Arm joints $(J_1, \ldots, J_7)$ are actuated by two simulated antagonist muscles, exactly as in version B. For details on how the muscles are implemented, see Section B.2 of Appendix B.



(a) **(b)**

Figure C.1: The kinematic chain of the arm and the hand. Cylinders represent rotational DOFs. The axes of cylinders indicate the corresponding axis of rotation. The links amongst cylinders represents the rigid connections that make up the arm structure. The joints are named as indicated in **b)**

C.3 Hand Structure

The robotic hand of version C differs from that of version B in terms of the distribution and the angle limits of the thumb joints. Joint J_8 allows the opposition of the thumb to the other fingers, and it varies within the range of $[-120°, +0°]$, where the lower limit corresponds to the thumb-little finger opposition. All the other thumb joints (J_9, J_{10} and J_{11}) are for the *extension/flexion* of the phalanges, and vary within a range of $[-90°, +0°]$. where the lower limit corresponds to complete flexion of the phalanx (i.e. the thumb closed). For all others details, see Section B.3 of Appendix B.

C.4 Hand Actuators

The joints cannot be controlled independent of each other, but rather, they are grouped according to the same grouping principle as was used in the development of the iCub hand (Sandini et al., 2004). More precisely, the two distal phalanges of the thumb move together, as do the two distal phalanges of the index and the middle fingers. Also, all the *extension/flexion* joints of the ring and little fingers are linked, as are all the joints of *abduction/adduction* of the fingers. Hence, only 9 actuators

move all the joints of the hand, with one actuator for each of the following groups of joints: $\langle J_8 \rangle$, $\langle J_9 \rangle$, $\langle J_{10}, J_{11} \rangle$, $\langle J_{13} \rangle$, $\langle J_{14}, J_{15} \rangle$, $\langle J_{17} \rangle$, $\langle J_{18}, J_{19} \rangle$, $\langle J_{12}, J_{16}, J_{20}, J_{24} \rangle$ and $\langle J_{21}, J_{22}, J_{23}, J_{25}, J_{26}, J_{27} \rangle$. These actuators are simple motors that control the joints according to their positions.

C.5 Hand Tactile Sensors

The hand is equipped with tactile sensors exactly as in version B; see Section B.5 of Appendix B.

D Bound in copies of publications

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 trajec-

tories 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).

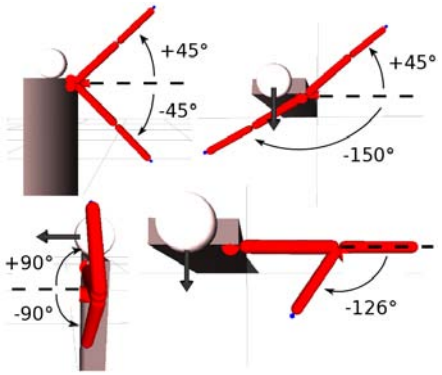


Figure 1: The four DOF of the simulated robotic arm. The two pictures on the top part of the figure indicate the abduction/adduction, extension/flexion of the shoulder joint, respectively from left to right. The bottom figure indicates the rotation of shoulder and the extension/flexion DOF of the elbow.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

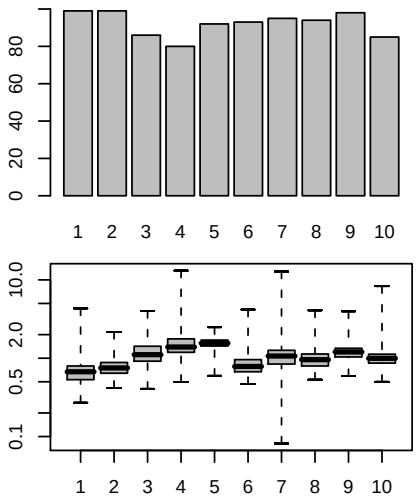


Figure 2: Performance on reaching a fixed target; **Top:** Percentage of trials in which the distance between the endpoint of the arm and the target is below 1*cm*, at the end of the trial. **Bottom:** Average distance between the endpoint of the arm and the target at the end of trials. Each column represents 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

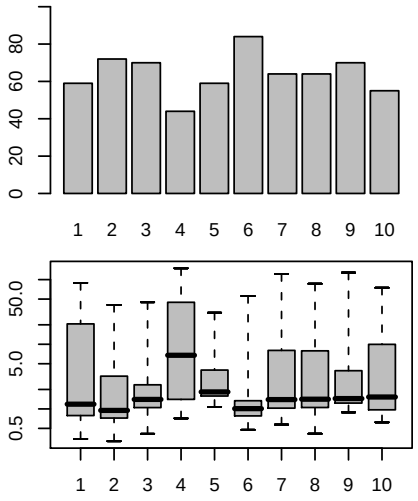


Figure 3: Performance on reaching a random positioned target; **Top:** Percentage of trials in which the distance between the endpoint of the arm and the target is below 1*cm*, at the end of the trial. **Bottom:** Average distance between the endpoint of the arm and the target at the end of trials. Columns,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

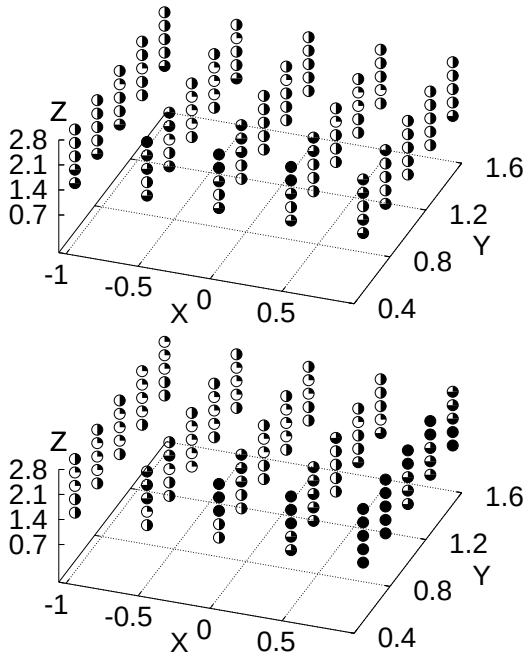


Figure 4: Performance obtained by testing with 125 targets points evenly distributed in front of robot on a $5 \times 5 \times 5$ grid area. The top and bottom graphs report the result obtained by testing two typical evolved individuals. The filled area of each bullet indicates the average distance between 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.

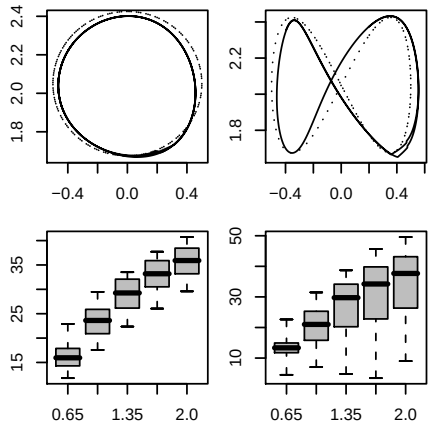


Figure 5: **Top:** trajectory produced by the endpoint of the arm and by a moving target (solid and dotted lines, respectively). Results obtained in two tests in which the target move by displaying a circular and an eight-shape trajectory (left and right picture, respectively). The vertical and horizontal 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.

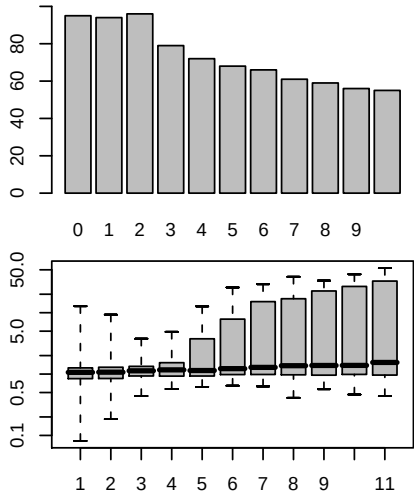


Figure 6: Performances obtained by testing robots evolved in a normal condition in a test condition in which the update of the sensory neurons is delayed. **Top:** Percentage of trials in which the distance between the endpoint of the arm and the target is below $1cm$, at the end of the trial. **Bottom:** 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.

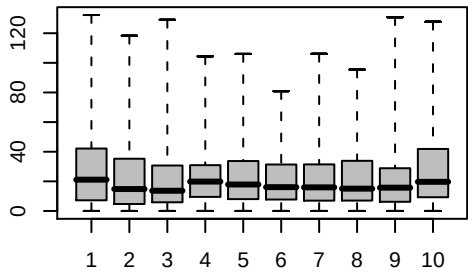


Figure 7: Average distance in cm between the trajectories produced by an evolved neural controller and the trajectories produced by manually setting the desired position of the joints on the basis of the final postures produced by the evolved neural controller. Each column indicates the result obtained 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.

frontiers
IN NEUROROBOTICS

Evolution of prehension ability in an anthropomorphic neurorobotic arm

Gianluca Massera [1,2], Angelo Cangelosi [2,*] and Stefano Nolfi [1]

1. Institute of Cognitive Science and Technologies, National Research Council (CNR), Italy
2. School of Computing, Communications and Electronics, University of Plymouth, UK

Edited by: Frederic Kaplan, Ecole Polytechnique Federale De Lausanne, Switzerland

Reviewed by: Jun Tani, RIKEN Brain Science Institute, Saitama, Japan
Simon Bovet, University of Zurich, Switzerland

In this paper, we show how a simulated anthropomorphic robotic arm controlled by an artificial neural network can develop effective reaching and grasping behaviour through a trial and error process in which the free parameters encode the control rules which regulate the fine-grained interaction between the robot and the environment and variations of the free parameters are retained or discarded on the basis of their effects at the level of the global behaviour exhibited by the robot situated in the environment. The obtained results demonstrate how the proposed methodology allows the robot to produce effective behaviours thanks to its ability to exploit the morphological properties of the robot's body (i.e. its anthropomorphic shape, the elastic properties of its muscle-like actuators and the compliance of its actuated joints) and the properties which arise from the physical interaction between the robot and the environment mediated by appropriate control rules.

Keywords: robotic arm, reaching and grasping, adaptation, evolutionary robotics

INTRODUCTION

The control of arm and hand movements in human and nonhuman primates is a fundamental research topic in cognitive sciences, neurosciences and robotics. Within arm and hand control, reaching and grasping behaviours represent key abilities as they constitute a prerequisite for complex object manipulation and use. In cognitive sciences, experimental and modelling studies have demonstrated the strict interdependence between action control and other cognitive functions such as language (Cangelosi et al., 2005; Pulvermuller, 2005). For example, some theories of language evolution have focused on the relationship between hand use, tool making and language evolution (Corballis, 2003). In neuroscience, numerous studies have demonstrated the fundamental role of the mirror neuron systems for motor control and in general for cognitive processing (Gallese and Lakoff, 2005; Rizzolatti and Arbib, 1998). In robotics, the motor control of arm and hand is a paradigmatic example of the difficulties that arise in the reverse engineering problem and the use of bio-inspired techniques in intelligent systems design (Schaal, 2002).

Despite the importance of the topic, the large body of available behavioural and neuroscientific data, and the vast number of studies done, the issues of how primates and humans learn to display reaching and grasping behaviour still remains highly controversial (Schaal, 2002; Shadmehr, 2002). Moreover, whilst many of the aspects that make these problems difficult have been identified, experimental research based on different techniques does not seem to converge towards the identification of a general methodology for developing robots able to display effective reaching and grasping abilities.

In this respect, one of the most controversial contraposition is between internal models (Kawato, 2002; Wolpert and Flanagan, 2002) and equilibrium point approaches (Shadmehr, 2002). The former approach is based on the assumption that our brain possess an internal model which allow us to: (a) predict how our limb will move and the sensations which will arise given the current sensory state and given a certain motor command which is going to be executed (direct mapping, and (b) transform a desired sensory state into the corresponding motor command which will achieve it (inverse mapping). In contrast, the latter approach is based on the assumption that muscles and associated spinal reflex circuitry confer to our limbs the ability to passively settle into stable position (i.e. equilibrium points) independently from their previous position. According to this hypothesis, the role of the central nervous system simply consists in the modification of the current equilibrium point.

In this paper, we will show how a simulated anthropomorphic arm can develop reaching and grasping skills through an adaptive evolutionary process (Nolfi and Floreano, 2000a) in which the free parameters regulate the fine-grained interactions between the robot and the environment and in which variations of free parameters are retained or discarded on the basis of their effects on the overall ability of the robot to reach and grasp objects. The analysis of the obtained results confirms the importance of dynamics resulting robot/environmental interactions and from the use of muscle-like actuators. Moreover, the results obtained demonstrate that effective reaching and grasping skills can be developed without relying on internal models performing direct and inverse mappings.

We will first review current work on reaching, with a brief discussion of the main research issues in this field and a review of current literature on the adaptive design of arm control behaviour in cognitive robots. The robotic model experimental setup will be described in section Materials and Methods. Subsequently in section Results we describe the results obtained. Finally, in Discussion the significance of the results obtained and our plans for the future.

* Correspondence: Angelo Cangelosi, School of Computing, Communications and Electronics, University of Plymouth, Drake Circus, Plymouth, UK.
e-mail: acangelosi@plymouth.ac.uk

Received: 06 Sep. 2007; paper pending published: 08 Oct. 2007; accepted: 12 Oct. 2007; published online: 02 Nov. 2007

Full citation: Frontiers in Neurorobotics (2007) **1**:4 doi: 10.3389/neuro.12/004.2007

Research issues in reaching and grasping in primates and humans

The primate arms consist of three main segments: arm, forearm and hand. These are attached to previous segments (the shoulder) through three actuated joints: 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 (Jones and Lederman, 2006).

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.

Various issues have been identified in the study of reaching behaviour in primates and humans. The main research questions related to robotics research include (i) the role of the redundancy of DOFs; (ii) the nonlinear relationship between joint movement and hand/target position; (iii) the role of gravity and inertia in suspended arms; (iv) the effects of speed and noise in motor control signals. First, we need to consider that 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, making learning very difficult.

The second issue regards the fact that anthropomorphic arms are highly nonlinear systems. 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. In addition, 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.

Gravity and physics dynamics also have a fundamental role in arm control. In articulated and suspended structures such as an anthropomorphic arm, gravity and inertia play a key role. In primate arms, muscles and associated spinal reflex circuitry appear 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 correct, the contribution of the central nervous system would simply consists in the modification of the current equilibrium point (Shadmehr, 2002).

Finally, the fact that sensors and actuators might be slow and noisy greatly affects the development of robotic arm. For instance, in humans the visual and proprioceptive information encoding changes of joints positions is available with a delay up to 100 ms. Motor commands issued by the central nervous system may take up to 50 ms 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).

Evolutionary robotics and neural network models of arm control

Evolutionary robotics consists on the autonomous design of the controller of robots through the use of evolutionary computation methods such as genetic algorithms (Nolfi and Floreano, 2000b). Typically in evolutionary robotic experiments the researcher defines the body of a robot (joints, limbs, sensors) and the surrounding environment (objects, obstacles, physics dynamics) with which it will interact. The robot control system consists of artificial neural network which has to learn to map input signals into motor responses. Learning is achieved through the evolution of the neural network parameters (connection weights and/or network topology).

Artificial neural networks have been typically used in robotics research to learn the correct mapping between two different spaces (e.g. joints, actuators, workspace; see (Torras, 2002)) as in inverse dynamics methods based on internal model approaches. Instead in the evolutionary approach the neural controller is seen as an internal dynamical system that interacts with the environments via the agent's body. There are not explicit mappings between spaces, but this emerges from minute continuous controller-body–environment interactions. From this point of view, the agent's behaviour is an emergent property of those tiny interactions. The evolutionary process is able to exploit the potential of simple architectures via dynamical interaction and is likely to lead to complex adaptive behaviour starting from minimal agents.

Notwithstanding the fact that evolution process leads to the selection and design of neural networks able to accomplish some task, this process is not a correlation-learning procedure, neither error-miniziming learning, nor a reinforcement-based procedure. Evolutionary robotics directly deals with some of the weakness of inverse dynamics approaches. In particular, it has been shown that an accurate mapping for inverse kinematics using feed-forward neural networks, due to global effect of weights, is extremely difficult to achieve (Krose and Van der smagt, 1993; Torras, 2002). In the evolutionary approach, as the controller is a dynamical system that acts directly and continuosly onto the dynamic of agent/environment interaction, it is possible to exploit simple architecture, such as multi-layer perceptrons, to learn inverse kinematic-dinamic solutions (Bianco and Nolfi, 2004; Massera et al., 2006). Furthermore, with the evolutionary paradigm there is no need to specify exactly the desired output, as in error-minimizing learning. This allows us to tackle the inverse dynamic problem for redundant anthropomorphic arm. In fact, in supervised approach for a given input the controller has to generate a sequence of forces to apply that are difficult to calculate 'a-priori' and to learn by error-minimizing procedures or reinforcement learning.

There are also situations where the same sensory pattern requires different responses of the robotic agent such as sensory aliasing (Nolfi and Marocco, 2001). This is a major issue in neural network correlation-based learning such as Hebbian rule or self-organizing maps. Instead, in evolutionary robotics the neural controller does not have to learn to follow a predefined pathway and can explore different solutions to achieve the same target points in space.

There have been few previous attempts to use evolutionary techniques to develop the controller for a robotic arm. Bianco and Nolfi (2004) used a standard evolutionary robotics approach for the autonomous design of the neural 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. These robots, however, were not able to deal with larger variations of the objects positions.

Buehrmann and Di Paolo (2004) evolved the control system for a 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 end-point 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 evolved separately for the ability to produce distinct 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 towards the object, approaching the object by controlling the second and the third joint, etc.).

Marocco et al. (2003) use a 6 DOF arm model to evolve the ability to touch or avoid objects according to their shape. In addition to the capability of discriminating objects, the robots are also evolved for their ability to 'name' the object (or the action) with which they are interacting. This permitted the analysis of different social interaction protocols to investigate social and cognitive factors that support the evolutionary

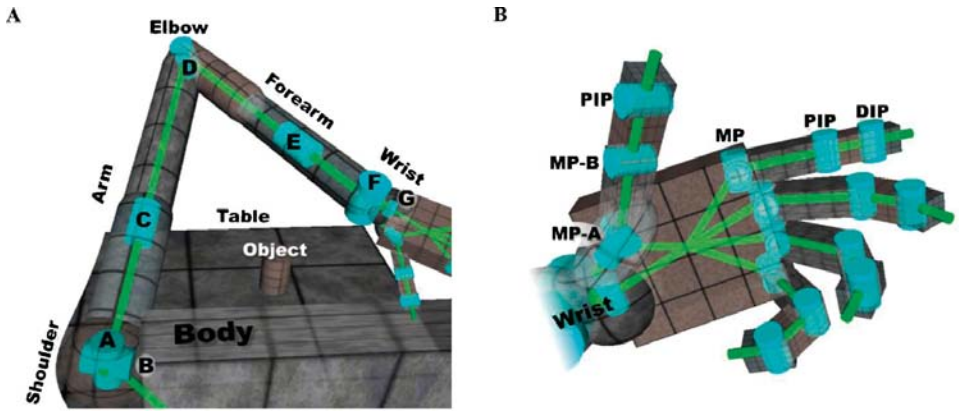


Figure 1. The kinematic chain of the arm and of the hand. *Cylinders represent rotational DOFs. The axes of cylinders indicate the corresponding axis of rotation. The links amongst cylinders represents the rigid connections that make up the arm structure.*

emergence of shared lexicons. Although this model used a very simplified arm model and limited object set and location, it provided a first attempt to use a neurorobotic model to study the link between action and linguistic representations from an evolutionary perspective.

Finally, in our previous evolutionary robotic model of reaching (Massera et al., 2006) we developed a realistic anthropomorphic arm with four DOFs. This is quite a different system from industrial arm robots 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. We successfully employed a method by which individuals were selected 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. This is in opposition to incremental approaches in which elementary components of the required final behaviour are identified by the experimenter and gradually included in the fitness evolutionary criteria.

MATERIALS AND METHODS

In this section, we describe the simulated robot, the robots actuators and sensors, the architecture of the neural controller and the adaptive process used to train the robot to grasp objects of different shapes.

The robot

The robot used in the experiments reported in this paper is a simulated humanoid robot provided with anthropomorphic robotic arm with 7 actuated DOFs, a robotic hand with 20 actuated DOFs, proprioceptive and touch sensors distributed within the arm and the hand and a vision system located in the robot's head.

The arm (Figure 1A) consists mainly of three elements (the arm, the forearm and the wrist) connected through articulations displaced into the shoulder, the arm, the elbow, the forearm and wrist. The shoulder is composed of a sphere with a radius 2.8 cm. The length of arm and forearm is 23 and 18 cm, respectively. The wrist consists of an ellipsoide with a radius of 1.45, 1.2 and 1.45 cm along x-, y- and z-axis, respectively. The joints **A**, **B** and **C** (Figure 1A) provide *abduction/adduction*, *extension/flexion* and *supination/pronation* of the arm in the range $[-140^\circ, +60^\circ]$, $[-90^\circ, +90^\circ]$ and $[-60^\circ, +90^\circ]$, respectively. These three DOFs acts like a ball-and-socket joint moving the arm in a way analogous to the human shoulder joint. The fourth DOF (**D**) located in the elbow is constituted by a hinge joint which provides *extension/flexion* within the $[-170^\circ, +0^\circ]$ range (radius–ulna bones). The fifth DOF (**E**) twists forearm providing *pronation/supination* of the wrist–hand in the range $[-90^\circ, +90^\circ]$. The sixth and seventh DOFs (**F** and **G**) on the wrist provide *flexion/extension* and *abduction/adduction* of the hand within $[-30^\circ, +30^\circ]$ and $[-90^\circ, +90^\circ]$ ranges, respectively.

Table 1. The size (in cm) of the segments forming the hand.

Finger	First phalanx	Second phalanx	Third phalanx
Index	$2.40 \times 0.80 \times 0.80$	$1.50 \times 0.75 \times 0.75$	$1.12 \times 0.70 \times 0.70$
Middle	$2.62 \times 0.80 \times 0.80$	$1.50 \times 0.75 \times 0.75$	$1.12 \times 0.70 \times 0.70$
Ring	$2.40 \times 0.80 \times 0.80$	$1.50 \times 0.75 \times 0.75$	$1.12 \times 0.70 \times 0.70$
Pinky	$2.25 \times 0.80 \times 0.80$	$1.50 \times 0.75 \times 0.75$	$1.12 \times 0.70 \times 0.70$

The robotic hand (Figure 1B) is composed of a palm and 14 phalangeal segments (see Table 1) that make up the digits (two for the thumb and three for each of the other four fingers) connected through 15 joints with 20 DOFs. The palm consists of a box of $4.6 \times 1.2 \times 4.2\,\text{cm}^3$. The thumb is composed of four connected objects: (i) an ellipsoide with a radiuns of $1.5 \times 0.6 \times 0.8\,\text{cm}^3$ which is half-sunked into the palm, (ii) a box of $2.40 \times 0.80 \times 0.90\,\text{cm}^3$ (corresponding to the metacarpal bones of the human thumb), (iii) a box of $1.60 \times 0.75 \times 0.85\,\text{cm}^3$ (corresponding to the first phalanx) and (iv) a box of $1.12 \times 0.75 \times 0.80\,\text{cm}^3$ (corresponding to the second phalanx). The other fingers are connected to the palm through knuckles represented by an ellipsoide of $0.65 \times 0.65 \times 0.5\,\text{cm}^3$ of radius. The three phalanges composing a finger are boxes jointed serially.

The joints in the hand are grouped in the metacarpophalangeal (MP), proximal interphalangeal (PIP) and distal interphalangeal (DIP) types. Each finger has two hinge joints, PIP and DIP (see Figure 1, right), that *extend/flex* phalanges within the range $[-90^\circ, +0^\circ]$. The MP group is composed of two joints that allow both *extension/flexion* and *abduction/adduction* of the first phalanx of each finger. The *extension/flexion* of MP is in the range $[-90^\circ, +0]$ for all fingers but the *abduction/adduction* movement range varies for different fingers and corresponds to $[-7^\circ, +0^\circ]$, $[-2^\circ, +2^\circ]$, $[-2^\circ, +5^\circ]$ and $[+0^\circ, +7^\circ]$ for the index, the middle, the ring and the pinky fingers, respectively. The thumb does not have the DIP joint, and the MP provides three DOF located in the MP-A and MP-B joints. The former joint has two DOFs providing *supination/pronation* and *abduction/adduction* of metacarpal part of the thumb in $[-120^\circ, +0^\circ]$ and $[-15^\circ, +90^\circ]$ ranges, respectively, which allow a good opposition of the thumb with the fingers. The MP-B and PIP joints consists of hinge

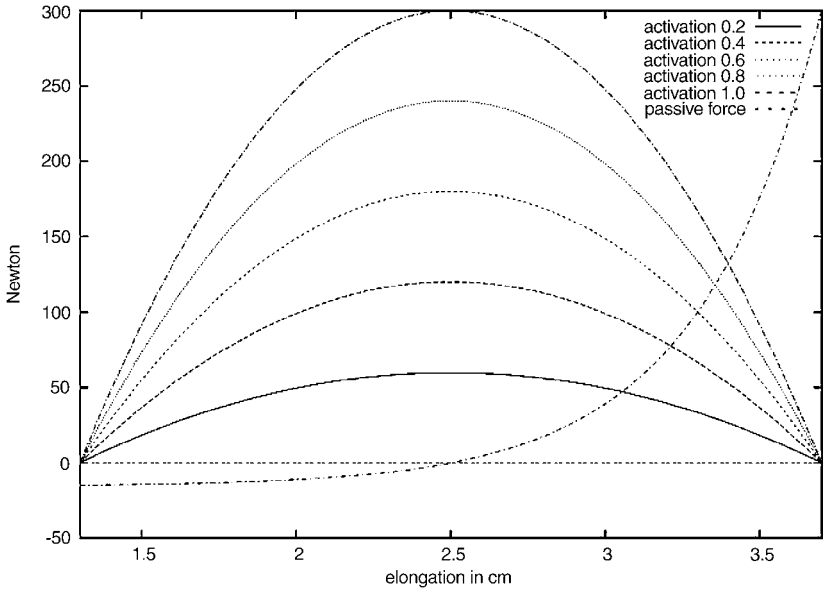


Figure 2. An exemplification of how the force exhorted by a muscle. *The graph shows how the force exhorted by a muscle varies as a function of the activity of the corresponding motor neuron and of the elongation of the muscle for a joint in which T_{max} is set to 300 N.*

joints that *extend/flex* the first and second phalanx of thumb in the same range of PIP and DIP joints of the other four fingers: $[-90^\circ, +0^\circ]$.

The actuators

The joints of the arm are actuated by two simulated antagonist muscles implemented accordingly to the Hill's muscle model (Sandercock et al., 2002; Shadmehr and Wise, 2005). More precisely, the total force exerted by a muscle (**Figure 2**) is the sum of three forces $T_A(\alpha, x) + T_P(x) + T_V(x')$ which depend on the activity of the corresponding motor neuron (α) and on the current elongation of the muscle (x) and which are calculated on the basis of the following equations:

$$\begin{aligned} T_A &= \alpha\left(-\frac{A_{sh}T_{max}(x-R_L)^2}{R_L^2}+T_{max}\right) \\ A_{sh} &= \frac{R_L^2}{(L_{max}-R_L)^2} \\ T_P &= T_{max}\frac{\exp\left\{K_{sh}\dfrac{x-R_L}{L_{max}-R_L}\right\}-1}{\exp\{K_{sh}\}-1} \\ T_V &= b\cdot\dot{x} \end{aligned} \quad (1)$$

where L_{max} and R_L are the maximum and the resting length of the muscle, T_{max} is the maximum force that could be generated, K_{sh} is the passive shape factor, b is the viscosity coefficient. The parameters of the equation are identical for all 14 muscles controlling the seven DOFs of the arm and have been set to the following values: $K_{sh} = 3.0$, $R_L = 2.5$, $L_{max} = 3.7$, $b = 0.9$, $A_{sh} = 4.34$ with the exception of parameter T_{max} which is set to 3000 N for joint B, to 300 N for joints A, C, D and E, and to 200 N for joints F and G.

Muscle elongation is simulated on the basis of actual angular position of each DOF, which is mapped linearly within the allowable angular range of each DOF. For instance, in the case of elbow where the limits are $[-170^\circ, +0^\circ]$, this range is mapped onto $[+1.3, +3.7]$ for the agonist muscle and inversely $[+3.7, +1.3]$ for antagonist muscle. Hence, when elbow is completely extended (angle 0), the agonist muscle is completely elongated (3.7) and antagonist completely compressed (1.3), and vice versa when elbow is flexed.

In the case of the hand, the positions of the joints are controlled by a limited number of variables (i.e. they are interdependent as in the case of human hands) through a velocity-proportional controller (joint maximum velocity is set to 0.30 rad/second). More precisely, the force exerted by the MP, PIP and DIP joints (MP-A, MP-B and PIP in the case of the thumb) which determine the *extension/flexion* of the corresponding finger are controlled by a single variable *theta* ranging between $[-90^\circ, +0^\circ]$. The desired position of the three joints is set to *theta*, *theta* and (2.0/3.0)**theta*, respectively. In the case of the thumb, the *supination/pronation* is also controlled by *theta* by setting the desired angle to $-(2.0/3.0)/theta$. The DOF which determine the *abduction/adduction* of the first phalanx of each finger is controlled by a second variable which has been set to the constant value of 0.0 rad.

The total weight of the arm and of the hand is 520.47 g. The robot and the robot/environmental interactions have been simulated by using Newton Game Dynamics (NGD, see: www.newtongamedynamics.com), a library for accurately simulating rigid body dynamics and collisions.

The sensors

The robot is provided with proprioceptive sensors which encode the current position of the DOFs of the arm and the hand, tactile sensors distributed over the hand, and of the vision system located on the robot head.

Seven arm propriosensors encode the current angles of the seven corresponding DOFs located on the arm and on the wrist normalized in the range $[-1, +1]$. Five hand propriosensors encode the current *extension/flexion* state of the five corresponding fingers in the range [0, 1] where 0 means fully extended and 1 means fully flexed. The hand propriosensors report the actual value of the MP-B joint for the thumb and the PIP joints of fingers. Due to compliance of finger's joints when the hand hit with an object and to the fact that the state of the three corresponding DOFs is summarized in a single variable, the same sensory state might correspond to a different states of the MP and DIP joints.

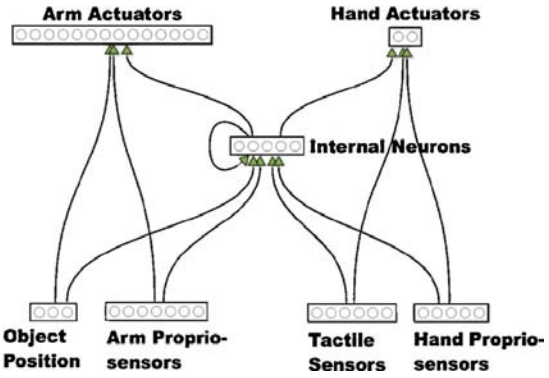


Figure 3. The architecture of the neural controllers. *Arrows indicated blocks of fully connected neurons. Internal and hand actuators neurons are also provided with a bias.*

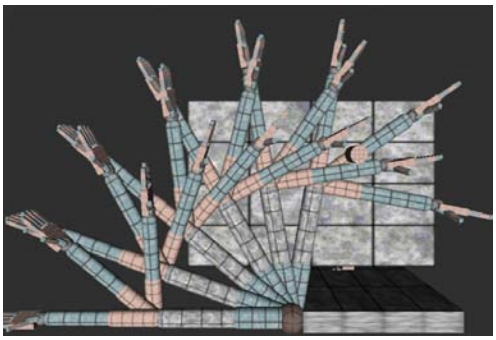


Figure 4. The 18 initial postures of the arm and of the hand used during the 18 corresponding trials.

The six tactile sensors measure whether the five fingers and the part constituted by the palm and wrist are in physical contact with another object. More precisely, each sensor encodes the number of contacts occurring in the corresponding body part normalized in the range [0, 1] through a logistic function with 0.2 as slope coefficient. The three vision sensors encode the output of a vision system (which has not been simulated) that computes the relative distance of the object with respect to the hand up to a distance of 80 cm normalized in the range [−1, +1] over three orthogonal axes.

The reason behind the choice of this particular sensory system configuration is that to study situations in which the vision and tactile sensory channels need to be integrated. In isolation, each of the two types of sensor does not provide enough information to perform the task.

The neural controller

The robot is provided with the neural controllers shown in **Figure 3** which include 21 sensory neurons, five internal neurons with recurrent connections and 16 motor neurons.

The object position sensors, arm and hand propriosensors and tactile sensors encode the state of the corresponding sensors described above. The actuators of the arm encode the activity of the 14 motor neurons controlling the corresponding muscles of the arm. The two actuators of the hand encode the desired *extension/flexion* state of the thumb and of the other four fingers, respectively (i.e. the four fingers are not controlled independently).

The state of the sensors, the desired state of the actuators and the internal neurons are updated every 0.010 second. The activity of the internal and motor neurons is calculated on the basis of a standard logistic function (with a slope coefficient of 0.5 in the case of the internal neurons and of 1.0 in the case of the motor neurons). In the case of the arm actuators and of the internal neurons, the output of the neuron corresponds to the neurons' activity. In the case of the hand actuators and the tactile sensors, instead, the output of the neurons is also depends from the neurons previous activation. More precisely, these neurons consist of leaky integrators in which the output is calculated on the basis of the following equation (Nolfi and Marocco, 2001):

$$O(t) = \delta \cdot \mathrm{Act}(t) + (1 - \delta) \cdot \mathrm{Act}(t - 1) \quad (2)$$

where Act is the activity of the neuron calculated on the basis of the logistic function (with slope coefficient 0.2 for tactile sensors and 1.0 for hand actuators) and δ is a time constant parameter ranging between [0, +1] (for alternative ways to implement leaky neurons see, for example, Beer, 1995).

The main criteria behind the choice of this particular neural network architecture have been to reduce the number of assumptions to the minimum and to reduce the number of free parameter as much as possible. A systematic analysis of the role of the architecture will be made in future work. For the moment, the analysis of the results obtained by varying some of the aspects of the architecture (results not shown) did not lead to qualitatively different results.

The adaptive process

The free parameters of the neural controller, i.e. the connection weights, the biases of internal neurons and hand actuators and the time constant of leaky-integrator neurons, have been adapted through an evolutionary robotics method (Nolfi and Floreano, 2000a).

The initial population consisted of 100 randomly generated genotypes, which encode the free parameters of 100 corresponding neural controllers. Each parameter is encoded with 16 bits. Each genotype contains 6096 bits corresponding to 381 free parameters: 366 connection weights and 7 biases normalized in the range [−10, +10] and 8 time constant normalized in the range [0.0, 1.0]. The 20 best genotypes of each generation were allowed to reproduce by generating five copies each. Four out of the five copies are subjected to mutations and one copy is left intact. During mutation each bit of the genotype has a 1.5% probability to be replaced with a new randomly selected value. The evolutionary process is continued for 400 generations (i.e. the process of testing, selecting and reproducing robots is iterated 400 times). The experiment was replicated 10 times.

The robot is adapted for the ability to grasp spherical and cylindrical objects placed on a table located in front of the robot. The objects can move freely by eventually falling off the table (**Figure 1A**). During the adaptive process, each genotype is translated into a corresponding neural controller, embodied in the simulated robot and tested for 18 trials. Each trial lasts 4 second corresponding to 400 steps. At the beginning of each trial the arm is set in the ith of the 18 corresponding predefined postures shown in **Figure 4**. The target object is placed in a fixed position in the central portion of the table. Spherical objects have a radius of 2.5 cm and a weight of 32.72 g, cylindrical objects have a radius of 2.0 cm, a height of 6.0 cm and a weight of 37.70 g.

Evolving robots are evaluated on the basis of the following two components' fitness function which reward reaching and grasping behaviours,

respectively

$$\frac{1}{1803600}\sum_{18}^{t=1}\sum_{400}^{s=200}\left(\frac{1}{1+0.25\cdot \mathrm{dist}}+500\cdot \mathrm{grasp}\right) \quad (3)$$

where *dist* encodes the distance between the barycentre of the hand and the object, *grasp* encode whether an object has been successfully grasped (i.e. grasp is 1 when the target object is elevated with respect to the table and is in physical contact with the robot hand and is 0 otherwise), *t* is the current trial and *s* is the current time step. To allow the robot to reach and grasp the object, the fitness is calculated only in the second-half of each trial (i.e. from time step 200 to time step 400). The constant at the beginning of the function, which corresponds to the maximum fitness that can be gathered by grasping each object during the first phase of each trial and by holding the object above the plane for the rest of the trial, is used to normalize the fitness value in the range [0, 1].

RESULTS

By analysing the behaviour of the evolved robots throughout generations, we observed that in 8 out of 10 replications of the experiment evolving robots develop an ability to reach and grasp objects which allows them to display optimal or close to optimal performance (see Figure 5).

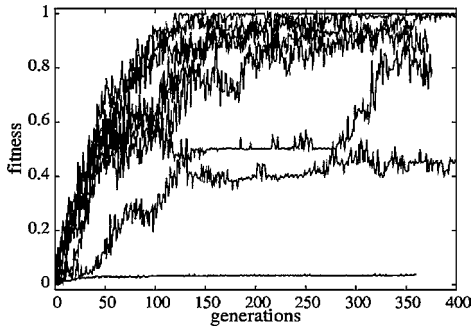


Figure 5. The fitness of the best individual throughout generations for 10 replications of the experiment.

By analysing the behaviour of the best evolved individual of one of the most successful replication, we observed that it successfully grasp the two types of objects from any of the 18 initial postures described above. As shown in Figure 6, the behaviour displayed by this individual

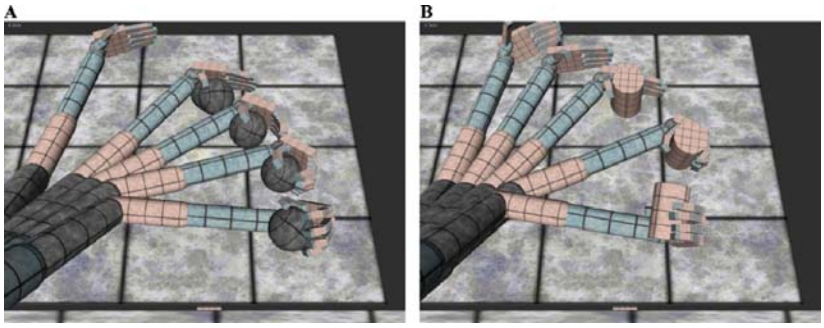


Figure 6. Five superimposed snapshots of the behaviour displayed by one of the best evolved robots. *(**A**) The evolved robot grasping a sphere; (**B**) The same evolved robot grasping a cylinder.*

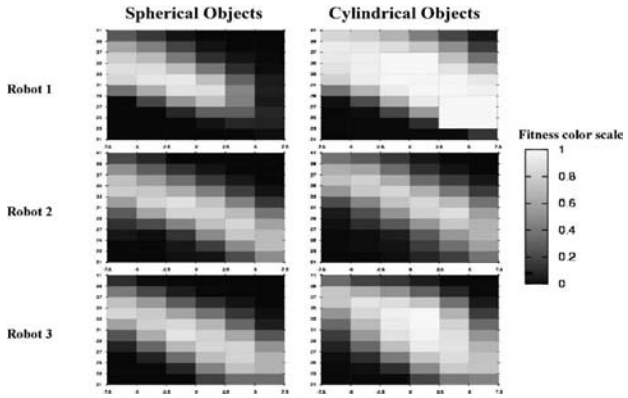


Figure 7. Performance of the best evolved robots of the three best replications of the experiment.

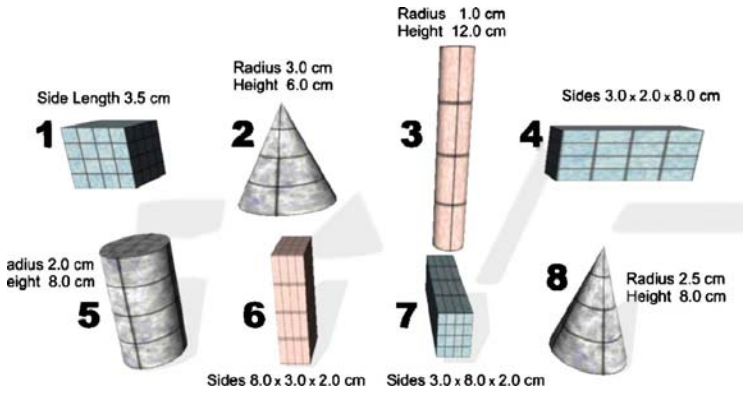


Figure 8. Objects used for testing robots' generalization ability with respect to object shape and size.

can be divided into three phases: (1) an initial phase in which the arm moves towards the object by first increasing and then decreasing the movement speed and in which the hand initiates to flex, (2) a second phase in which the tactile sensors start to be activated, the arm stays still or almost still and in which the wrist and the fingers flex around the object, (3) a final phase in which the arm rotates and moves the wrist so to lift the object from the table and so to reduce the risk that the object fall down from the hand. A set of video showing the behaviour of evolved robots in detail can be accessed from the following Web page: http://laral.istc.cnr.it/esm/arm-grasping/.

By testing evolved robots in different conditions with respect to the condition in which they have been evolved, we observed that they display remarkable generalization abilities with respect to the position of the object on the table and with respect to the shape of the object.

Figure 7 shows the average performance of the best evolved robots of three of the best replications of the experiment observed by systematically varying the position of the objects on the table. As can be seen, although different individuals vary with respect to their generalization capabilities, they all display rather good performance on the central diagonal area which corresponds to the preferential trajectory followed by the arm in normal conditions (i.e. when the objects are placed on the central position of the table). The decrease in performance on the top-right and bottom-left part of the table can be explained by considering that grasping objects located in these positions require postures which differ significantly from those assumed by the robots to grasp objects in the central area of the table.

Each robot has been tested in 120 different conditions corresponding to 60 different position of the object on the table and to two types of objects (spherical and cylindrical objects). For each testing condition, the robot has been tested for 18 trials corresponding to the 18 different starting position of the arm. The colours of the rectangles indicate the performance. For each picture, the left and right areas correspond to the left and right area of the table with respect to the robot, respectively. The top and bottom areas correspond to the proximal and distant areas of the table with respect to the robot, respectively.

By testing evolved robots in environment containing the objects shown in Figure 8, we also observed that evolved robots display remarkable generalization abilities with respect to the shape and size of the objects (see Figure 9).

The difference in performance amongst the individual robots of different replications of the experiment are due to the different behavioural

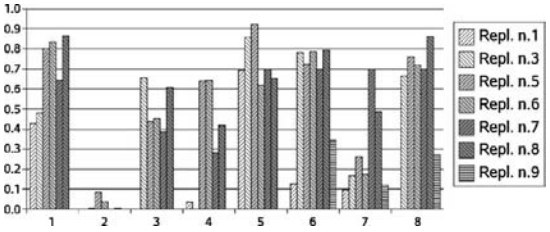


Figure 9. Performances of the evolved robots of the seven best replications of the experiment observed by testing the robots with the eight objects shown in Figure 8.

strategies displayed by evolved individuals with particular reference to the second and third phases of the behaviour in which the robots grasp and lift the objects (for more information, see the video available from http://laral.istc.cnr.it/esm/arm-grasping/). For example, the fact that best individual of replication 1 displays poor performance with objects 2, 4, 6 and 7 with respect to other evolved individuals is due to the fact that it flexes its fingers very quickly. This type of strategy, in fact, prevents this robot from the possibility of exploiting the adjustments of the relative position of the fingers with respect to the objects which arise spontaneously in time as a result of the effects of the forces exhorted by the hand, the collisions between the fingers and the object and the compliance of the hand The poor performance of the best individuals of replication 7 on objects 2, 3 and 4 can be explained by considering that the way in which this individual lifts the objects after the grasping phase tends to produce collisions with the plane in the case of big objects which might cause the falling down of the object from the hand. Finally, the good performance of replication 8 can be explained by the ability of this robot in controlling the thumb, which is crucial for grasping difficult slippery objects, and by the fact that this robot produces a limited rotation of the arm and of the wrist during the lifting phase which minimize the risk of collisions with the plane after the objects have been grasped.

Overall these results suggest that certain behavioural strategies might be effective for a large variety of objects and that the limited differences in terms of shape and size of the objects to be grasped should not necessarily have an impact on the rules that regulate the robot/environmental

interactions.

Although a systematic analysis which can allow us to identify the factors that lead to such good generalization ability will be carried out in future work, preliminary analysis (not shown) suggest that the muscle-like properties of the actuators of the arm and the compliance of the actuators of the fingers combined with the adaptation process, which manages to exploit these properties, play an important role. With respect to the compliance of the fingers, in particular, it greatly simplifies the problem of adapting the postures of the fingers to the shape of the object. Regarding the generalization ability with respect to the position of the object, an important factor is constituted by the fact that the position of the object extracted by the vision system is encoded in relation to the position of the hand.

DISCUSSION

In this paper, we showed how effective reaching and grasping behaviour can be developed through a trial and error process in which the free parameters encode the control rules which regulate the fine-grained interaction between the robot and the environment and variation of the free parameters are retained or discarded on the basis of their effects at the level of the global behaviour exhibited by an anthropomorphic robotic arm situated in the environment and provided with muscle-like actuators. The robots are let free to choose the way in which the problem can be solved during the adaptation process, since they are rewarded only with respect to their ability to approach and lift objects irrespectively of the particular trajectory with which they approach the objects, the posture of the arm and of the hand that they assume, and the way in which different motor actions produced by the robot in interaction with the environment are distributed over time.

The experimental setup presented in the paper is significantly more advanced with respect to previous works based on similar adaptive techniques (Bianco and Nolfi, 2004; Buehrmann and Di Paolo, 2004; Gomez et al., 2005; Massera et al., 2006) with respect to the morphology of the robot (which an anthropomorphic robotic arm and hand with 27 DOFs), with respect to the size of the neural controller and to the dimensionality of the corresponding search space and with respect to the task which involved the ability to reach and grasp freely moving objects with different shapes placed on a table which constraints the movements of the robot.

The obtained results demonstrate how the proposed methodology and the exploitation of the properties which arise from the physical interaction between the robot and the environment allow the robot to produce effective behaviours on the basis of a parsimonious control system. For example, the effects of the collisions between the fingers of the robotic hand and the objects being grasped combined with the compliance of the finger's joints allow the spontaneous conformation of the robot hand to the shape of the object which in turn allows the robot to effectively grasp objects with different shapes and orientations without the need of control mechanisms able to regulate the movement of the arm and of the hand on the basis of the characteristics of the objects.

This line of research is also consistent with recent cognitive robotics approaches such as in the field of developmental robotics (Lungarella and Metta, 2003). Developmental robotics, also known as epigenetic robotics, in an interdisciplinary approach to robot design. Developmental robots are characterized by a prolonged developmental process through which varied and complex cognitive and perceptual structures emerge as a result of the interaction of an embodied system with a physical and social environment. Lungarella and Metta show that although most of the current developmental robotic investigations have focussed on sensorimotor control (e.g. reaching) and social interaction (e.g. gaze control), future cognitive robotics research should go beyond gazing, pointing and reaching. In order to design truly autonomous behaviour, future robotics research should integrated motor control with better sensory and motor apparata, more refined value-based learning mechanisms and means of exploiting neural and body dynamics.

This neurorobotic approach also has a potential relevance to computational neuroscience research on motor control (Shadmehr and Wise, 2005). The current architecture of the robot's neural controller has not been constrained on any specific brain region known to be involved in limb control. Therefore, the current model and simulation results cannot be used to make any speculation on the relevance to neuroscience research. However, future extensions of the model might focus specifically on investigating the role and structure of the neural network controller and its mapping onto brain regions and circuitries (e.g. cerebellum, motor areas) known to be involved in prehension ability (Jones and Lederman, 2006; Kawato, 2002). This would also make possible the testing of current theories of minimization criteria such as energy minimum, jerk minimum and stability maximization, for generating voluntary movements and the comparison between robotic model results and limb neurophysiology literature (Shadmehr, 2002). For example, recent evolutionary robotic models on the development and integration of action and language capabilities have demonstrated that neural network architectures can be constrained to reflect known neurophysiological phenomena (Arbib et al., 2000; Cangelosi and Parisi, 2004). For example, Cangelosi and Parisi (2004) used synthetic brain imaging techniques to demonstrate that the region of the robot's neural network that specializes for sensorimotor integration is also involved in the processing of the names of actions (verbs), whilst the network region specialized in the representation and categorization of visual information only is also involved in the processing of the names of objects (nouns).

In future work, we plan to extend the variability of the objects to be grasped in order to investigate problems which require an ability to display a variety of qualitatively different approaching and grasping strategies. Within this future research line, we would like to study how neurocontrollers, developed through the methodology described in this paper, can be complemented with additional mechanisms which, on one hand might favour the development of different behavioural strategies and, on the other hand might allow the robot to select the approaching and grasping strategy which is appropriate to the current robot/environmental circumstances. To achieve this goal, we plan to implement and to compare different mechanisms such as continuous time recurrent neural networks including neurons varying at tuneable time scales (Beer, 2005; Nolfi and Marocco, 2001) and internal models operating at the level of elementary behaviour rather than at the level of the fine-grained robot/environmental interactions (i.e. which allow the robot to select the behavioural strategy which produces a desired effect by exploiting the ability of forecasting the global effects of the execution of a given behavioural strategy in a given robot/environmental situation, see (Tani et al., 2004) and Nishimoto et al., in press).

Moreover, to address the relevance of such a simulation model to research with physical robotic platform, we are currently involved in a collaborative project to test the evolved controllers on the RobotCub physical robot (Sandini et al., 2004 www.robotcub.org). This will allow us to verify the accuracy of the simulator and to revise the experiments performed in simulation so to progressively reduce the gap between the simulated and the real robot/environmental systems.

SUPPLEMENTAL DATA

Supplemental data for this article including movies of the behaviours displayed by evolved robots of different replications of the experiment can be found at the following address: http://laral.istc.cnr.it/esm/arm-grasping.

CONFLICT OF INTEREST STATEMENT

We declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ACKNOWLEDGEMENTS

The research has been supported by the ECAGENTS project funded by the Future and Emerging Technologies programme (IST-FET) of the European Community under EU R&D contract IST-1940.

REFERENCES

Arbib, M. A., Billard, A., Iacoboni, M., and Oztop, E. (2000). Synthetic brain imaging: grasping, mirror neurons and imitation. *Neural Netw.* 13, 975–997.

Beer, R. D. (1995). On the dynamics of small continuous-time recurrent neural networks. *Adapt. Behav.* 3, 471–511.

Beer, R. D. (2005). A dynamical systems perspective on agent-environment interaction. *Artif. Intell.* 72, 173–215.

Bianco, R., and Nolfi, S. (2004). Evolving the neural controller for a robotic arm able to grasp objects on the basis of tactile sensors. *Adapt. Behav.* 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. (2005). Modeling Language, Cognition and Action: Proceedings of the 9th Neural Computation and Psychology Workshop (Singapore, World Scientific).

Cangelosi, A., and Parisi, D. (2004). The processing of verbs and nouns in neural networks: insights from synthetic brain imaging. *Brain Lang.* 89, 401–408.

Corballis, M. C. (2003). From hand to mouth: the origins of language.

Gallese, V., and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 21.

Gomez, G., Hernandez, A., Eggenberger Hotz, P., and Pfeifer, R. (2005). An adaptive learning mechanism for teaching a robotic hand to grasp. In International Symposium on Adaptive Motion of Animals and Machines.

Jones, L. A., and Lederman, S. J. (2006). Human hand function.

Kawato, M. (2002). Cerebellum and motor control. In Handbook of Brain Theory and Neural Networks, 2nd edn, M. A. Arbib, ed. (Cambridge, MA, MIT Press), pp. 190–195.

Krose, B. J. A., and Van der smagt, P. P. (1993). An introduction to neural networks.

Lungarella, M., and Metta, G. (2003). Beyond gazing, pointing, and reaching: A survey of developmental robotics. Paper presented at: 3rd International Workshop on Epigenetic Robotics (Boston, USA).

Marocco, D., Cangelosi, A., and Nolfi, S. (2003). Evolutionary robotics experiments on the evolution of language. *Philos. Trans. R. Soc. Lond. A* 361, 2397–2421.

Massera, G., Cangelosi, A., and Nolfi, S. (2006). Developing a reaching behaviour in a simulated anthromorphic robotic arm through an evolutionary technique. In Artificial Life X: Proceeding of the Tenth International Conference on the simulation and synthesis of living systems (Cambridge, MA, MIT Press).

Mial, R. C. (2002). Motor control, biological and theoretical. In Handbook of Brain Theory and Neural Networks, 2nd edn (Cambridge, MA, MIT Press), pp. 110–113.

Nolfi, S., and Floreano, D. (2000a). Evolutionary robotics: the biology, intelligence, and technology of self-organizing machines (Cambridge, MA, MIT Press/Bradford Books).

Nolfi, S., and Floreano, D. (2000b). Evolutionary robotics: the biology, intelligence, and technology of self-organizing machines.

Nolfi, S., and Marocco, D. (2001). Evolving robots able to integrate sensory-motor information over time. *Theory Biosci.* 120, 287–310.

Pulvermuller, F. (2005). Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582.

Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.*.

Sandercock, T. G., Lin, D. C., and Rymer, W. Z. (2002). Muscle models. In Handbook of Brain Theory and Neural Networks, 2nd edn, M. A. Arbib, ed. (Cambridge, MA, MIT Press), pp. 711–715.

Sandini, G., Metta, G., and Vernon, D. (2004). RobotCub: an open framework for research in embodied cognition. Paper presented at: Fourth International Conference on Humanoid Robots.

Schaal, S. (2002). Arm and hand movement control. In Handbook of Brain Theory and Neural Networks, 2nd edn, M. A. Arbib, ed. (Cambridge, MA, MIT Press), pp. 110–113.

Shadmehr, R. (2002). Equilibrium point hypothesis. In Handbook of Brain Theory and Neural Networks, 2nd edn (Cambridge, MA, MIT Press), pp. 409–412.

Shadmehr, R., and Wise, S. P. (2005). The computational neurobiology of reaching and pointing (Cambridge, MA, MIT Press).

Tani, J., Ito, M., and Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: reviews of robot experiments using RNNPB. *Neural Netw.* 17, 1273–1289.

Torras, C. (2002). Robot arm control. In Handbook of Brain Theory and Neural Networks, 2nd edn (Cambridge, MA, MIT Press), pp. 979–983.

Wolpert, D. M., and Flanagan, J. R. (2002). Sensorimotor learning. In Handbook of Brain Theory and Neural Networks, 2nd edn, M. A. Arbib, ed. (Cambridge, MA, MIT Press), pp. 1020–1023.

Active categorical perception of object shapes in a simulated anthropomorphic robotic arm

Elio Tuci, Gianluca Massera, and Stefano Nolfi

***Abstract*—Active perception refers to a theoretical approach to the study of perception grounded on the idea that perceiving is a way of acting, rather than a process whereby the brain constructs an internal representation of the world. The operational principles of active perception can be effectively tested by building robot-based models in which the relationship between perceptual categories and the body-environment interactions can be experimentally manipulated. In this paper, we study the mechanisms of tactile perception in a task in which a neuro-controlled anthropomorphic robotic arm, equipped with coarse-grained tactile sensors, is required to perceptually categorise spherical and ellipsoid objects. We show that best individuals, synthesised by artificial evolution techniques, develop a close to optimal ability to discriminate the shape of the objects as well as an ability to generalise their skill in new circumstances. The results prove that the agents solve the categorisation task in an effective and robust way by self-selecting the required information through action and by integrating experienced sensory-motor states over time.**

***Index Terms*—Categorical perception, evolutionary robotics, artificial neural networks.**

I. Introduction

Categorical perception can be considered the ability to divide continuous signals received by sense organs into discrete categories whose members resemble more one another than members of other categories. Categorical perception represents one of the most fundamental cognitive capacities displayed by natural organisms, and it is an important prerequisite for the exhibition of several other cognitive skills [see 1]. Not surprisingly, categorical perception has been extensively studied both in natural sciences such as Psychology, Philosophy, Ethology, Linguistics, and Neuroscience, and in artificial sciences such as Artificial Intelligence, Neural Networks, and Robotics [see 2, for a comprehensive review of this research field]. However, in the large majority of the cases, researchers have focused their attention on categorisation processes that are passive and instantaneous. Passive categorisation processes take place in those experimental setups in which the agents can not influence the experienced sensory states through their actions. Instantaneous categorisation processes are those in which the agents are demanded to categorise the current experienced sensory state rather than a sequence of sensory states distributed over a certain time period.

In this paper, instead, we study categorisation processes that are active and eventually distributed over time [3, 4]. This task is achieved by exploiting the properties of autonomous embodied and situated agents. An important consequence of being situated in an environment consists in the fact that the sensory stimuli experienced by an agent are co-determined by the action performed by the agent itself. That is, the actions and the behaviour exhibited by the agent later influence the stimuli it senses, their duration in time, and the sequence with which they are experienced. This implies that: (i) categorical perception is strongly influenced by an agent's action [see also 5, 6, on this issue]; and (ii) sensory-motor coordination (i.e., the ability to act in order to sense stimuli or sequence of stimuli that allow an agent to perform its task) is a crucial aspect of perception and more generally of situated intelligence [see 7].

Although the significance of embodiment and situatedness for the study of the underlying mechanisms of behaviour and cognition is widely recognised, building artificial systems that are able to actively perceive and categorise sensory experiences is a challenging task. This can be explained by considering that, from the point of view of the designer, identifying the way in which an agent should interact with the environment in order to sense the favourable sensory states is extremely difficult. One promising approach, in this respect, is constituted by evolutionary methods in which the agents are left free to determine how they interact with the environment (i.e., how they behave, in order to solve their task). With these methods, free parameters (i.e., those that are modified during the evolutionary process) encode features that regulate the fine-grained interactions between the agent and the environment. The evolutionary process consists in retaining or discarding the free parameters on the basis of their effects at the level of the overall behaviour exhibited by the agent [see 8, 9, 10, for a detailed illustration of the methodological approach employed].

In this paper, we describe an experiment in which evolutionary methods are used to investigate the perceptual skills of an autonomous agent demanded to actively categorise unanchored spherical and ellipsoid objects placed in different positions and orientations over a planar surface. The agent is a simulated anthropomorphic robotic arm with 27 actuated Degrees of Freedom (hereafter, DoFs). The arm is equipped with coarse-grained tactile sensors and with proprioceptive sensors encoding the position of the joints of the arm and of the hand. The task requires the agent to produce different categorisation outputs for objects with different shapes and similar categorisation outputs for objects with the same shape. The aim of this study is to prove that, in spite of the complexity of the experimental scenario, the evolutionary approach can be successfully employed to design neural mechanisms to allow the robotic arm to perform the perceptual categorisation task. Moreover, we unveil the operational principles of successful

E. Tuci, G. Massera, and S. Nolfi are with the ISTC-CNR, Via San Martino della Battaglia, n. 44, 00185 Rome, Italy, e-mail: {elio.tuci, gianluca.massera, stefano.nolfi}@istc.cnr.it (see http://laral.istc.cnr.it/elio.tuci/).

agents. In particular, we look at (i) how the robot acts in order to bring fourth the sensory stimuli which provide the regularities necessary for categorising the objects in spite of the fact that sensation itself may be extremely ambiguous, incomplete, and noisy; (ii) the dynamical nature of sensory flow (i.e., how sensory stimulation varies over time and the time rate at which significant variations occur); (iii) the dynamical nature of the categorisation process (i.e., whether the categorisation process occur over time while the robot interacts with the environment); (iv) the role of qualitatively different sensation originated by different sensory channels in the accomplishment of the categorisation task.

We prove that a further elaboration of evolutionary methods proposed in related studies can be successfully applied to problems that are non-trivial and significantly more complex with respect to the state of the art reviewed in Section II. In particular, we show that the best evolved robots develop a close to optimal ability to discriminate the shape of the objects as well as an ability to generalise their skill in new circumstances. These results prove that the problem can be solved in an effective and robust way by self-selecting the required information through action and by integrating experienced sensory-motor states over time.

II. State of the art

There is a growing body of literature in robotics which is devoting increasingly more efforts in obtaining discrimination of material properties (e.g., hardness, texture) and object shape using touch in artificial arms. Many of these works, like the one described in [11], draw inspiration from human perceptual capability to develop highly elaborated touch sensors. In [11], the authors describe a tendon driven robotic hand covered with artificial skin made of strain gauges sensors and polyvinylidene films. The strain gauges sensors mimic the functional properties of Merkel cells in human skin and detect the strain. Polyvinylidene films mimic the functional properties of the Meissner corpuscles and detect the velocity of the strain. The artificial hand, through the execution of squeezing and tapping procedures, manages to discriminate objects based on their hardness. In a similar vein, the research group at the Lund University has developed three progressively complex versions of a robotic hand (LUCS Haptic Hand I, II, and III) designed for haptic perception tasks [12, 13, 14]. The perceptual capabilities of the three version of LUCS, which differ in their morphology and in their sensory capabilities, have been tested during the execution of a grasping procedure on objects made of different material (e.g., plastic and wood). The authors showed that the sensory patterns generated in interactions with the objects are rich enough to be used as a basis for haptic object categorisation [15]. Other robotics systems combine visual and tactile perception to carry out fairly complex object discrimination tasks [see 16, 17, 18].

Generally speaking, we can say that, in spite of the heterogeneity in hardware and control design, the research works mentioned above focus on the characteristics of the tactile sensory apparatus and/or on the categorisation algorithms. In these works, the way in which the sensory feedback affects the movement of the hand is determined by the experimenter on the basis of her intuition. Moreover, the discrimination phase follows the exploration phase and it is performed by elaborating sensory data gathered during manipulation of the objects (i.e., the data collected during the exploration phase cannot influence the agents successive behaviour).

The work described in this paper differs significantly from the above mentioned literature since the way in which the agent interacts with the environment is not designed by the experimenter but is adapted in order to facilitate the categorisation task and since the agent is left free to shape its motor behaviour on the basis of previously experienced sensory states. Rather than studying the performances of particularly effective tactile sensors or of specific categorisation algorithms, we focus on the development of autonomous actions for the discrimination of objects shape through coarse-grained binary tactile sensors and proprioceptive sensors. The issue of how a robot can actively develop categorisation skills has been already investigated in few recent research works. In general terms, these works demonstrate how adapted robots exploit their action to self-select stimuli which enable and/or simplify the categorisation process and how this leads to solutions which are parsimonious and robust [see 19, 20, 21, 22].

Particularly relevant for this study is the work described in [23]. The authors studied the case of a simulated robotic "finger" which has been evolved for discriminating the shape of spherical versus cubic objects (anchored to a fixed point) of different sizes and orientations. The robotic finger is constituted by an articulated structure made by three segments connected through motorised joints with six DoFs, six corresponding actuators, six proprioceptive sensors encoding the current position of the joints, and three tactile sensors placed on the three corresponding segments of the finger. The authors observed that the adapted robots solve their problem through simple control rules that makes the robot scan for the object by moving horizontally from the left to the right side and by moving slightly up as a result of collisions between the finger and the object. These simple control rules lead to the exhibition of two different behaviours. With spherical objects, the robotic finger fully extends itself on the left side of the object after following the object surface. With cubic objects, the robotic finger remains fully bended close to one of the corners of the cube. These two behaviours corresponds to well differentiated activations of the proprioceptive sensors. These differences are used by the finger to distinguish the two types of object. Note that, although the discriminating cue necessary to categorise is available in each single sensory pattern experienced after the exhibition of the appropriate behaviour, this cue results from the dynamical process arising as a result of several robot/environmental interactions. In [24], the authors show that a visually guided robot arm whose neuro-controller is evolved for reaching and tracking, can exploit its actions to self-select stimuli which facilitates the accomplishment of spatial and temporal coordination.

Unlike in the experiments described in [23, 24], sensory-motor coordination does not always guarantee the perception of well differentiated sensory states in different contexts corresponding to different categories. Under these circumstances,

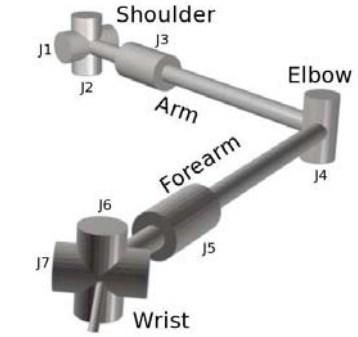
the agent can actively categorise their perceptual experiences by integrating ambiguous sensory information over time. Few studies have already shown that evolved wheeled robots compensate for unreliable sensory patterns due to coarse sensory apparatus by acting and re-acting to temporally distributed sensory experiences, in a way to bring forth the necessary regularities that allow them to associate a stimulus with its category [see 25, 26].

The experiment presented in this paper focuses on a non-trivial task that is significantly more complex to that investigated in previous studies due to the high similarity between the objects to be discriminated, the difficulty of controlling a system with many degree of freedom, and the need to master the effects produced by gravity, inertia, collisions, etc. As shown in Section VII, the analysis of the strategy displayed by best evolved robots demonstrates that, also in this case, sensory-motor coordination plays a crucial role, as in [23, 24]. Indeed, the best robots manipulate the objects so to experience the regularities which allow them to appropriately categorise the shape of the objects. However, sensory-motor coordination does not seem to guarantee the perception of fully differentiated sensory states corresponding to different categories. The problem caused by the lack of clear categorical evidences is solved through the development of an ability to integrate ambiguous information over time through a process of evidences accumulation.

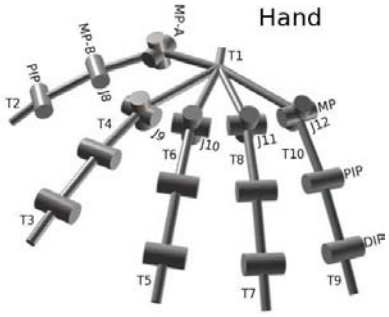
III. The Robot's Structure

The simulated robot consists of an anthropomorphic robotic arm with 7 actuated DoFs and a hand with 20 actuated DoFs. Proprioceptive and tactile sensors are distributed on the arm and the hand. The robot and the robot/environmental interactions are simulated using Newton Game Dynamics (NGD), a library for accurately simulating rigid body dynamics and collisions (more details at www.newtondynamics.com). The arm consists mainly of three elements: the arm, the forearm, and the wrist. These elements are connected through articulations displaced into the shoulder (joint J_1 for the extension/flexion, J_2 for the abduction/adduction, and J_3 for the supination/pronation movements), the elbow (joint J_4 for the extension/flexion movements), and the wrist (joints J_5, J_6, J_7 for the roll/pitch/yaw movements, see Figure 1a).

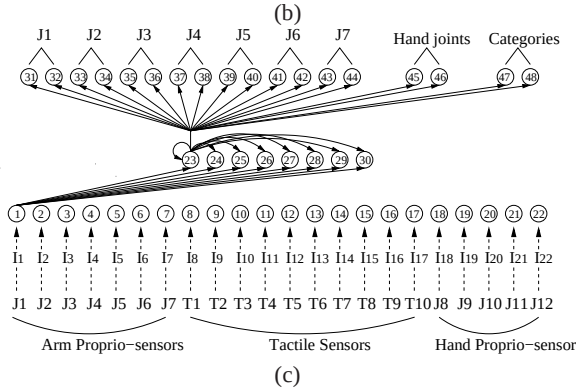
The robotic hand is composed of a palm and fourteen phalangeal segments that make up the digits (two for the thumb and three for each of the other four fingers) connected through 15 joints with 20 DoFs (see Figure 1b). There are three different types of hand joints: metacarpophalangeal (MP), proximal interphalangeal (PIP), and distal interphalangeal (DIP). All of them bring forth the extension/flexion movements of each finger while only the metacarpophalangeal joints are for the abduction/adduction movements. The thumb has an extra DoF in metacarpophalangeal joints which is for the axial rotation. This rotation makes possible to move the thumb towards the other fingers [see 27, for a detailed description of the structural properties of the arm]. The active joints of the robotic arm are actuated by two simulated antagonist muscles implemented accordingly to the Hill's muscle model, as detailed in the next Section.



(a)



(b)



(c)

Fig. 1: The kinematic chain (a) of the arm, and (b) of the hand. (c) The architecture of the arm neural controller. In (a) and (b), cylinders represent rotational DoFs; the axes of cylinders indicate the corresponding axis of rotation; the links among cylinders represents the rigid connections that make up the arm structure. In (c) the circles refer to the artificial neurons. Continuous line arrows indicate the efferent connections for the first neuron of each layer. Dashed line arrows indicate the correspondences between joints and tactile sensors and input neurons. The labels on the dashed line arrows refer to the notation used in equation 1a to indicate the readings of the corresponding sensors.

IV. The Robot's Sensors, Controller, and Actuators

The agent controller consists of a continuous time recurrent non-linear network (CTRNN) with 22 sensory neurons, 8 internal neurons, and 18 motor neurons [see Figure 1c and also 28]. At each time step, the activation values y_i of sensory neurons $i = 1, .., 7$ is updated on the basis of the state of the proprioceptive sensors of the arm and of the wrist which encode the current angles, linearly scaled in the range $[-1, 1]$, of the seven corresponding joints located on the arm and on the wrist (see joints J_1, J_2, J_3, J_4, J_5, J_6, and J_7 in Figure 1a). The activation values y_i of sensory neurons $i = 8, .., 17$ is updated on the basis of the state of tactile sensors distributed over the hand. These sensors are located on the palm (see label T_1 in Figure 1b), on the second phalange of the thumb (see label T_2 in Figure 1b), and on the first phalange (see labels T_4, T_6, T_8, T_{10} in Figure 1b) and the third phalange (see labels T_3, T_5, T_7, T_9 in Figure 1b) of each finger. These sensors return 1 if the corresponding part of the hand is in contact with any another body (e.g., the table, the sphere, the ellipsoid, or other parts of the arm), otherwise 0. The activation values y_i of sensory neurons $i = 18, .., 22$ is updated on the basis of the state of the hand proprioceptive sensors which encode the current extension/flexion of the five corresponding fingers (see joints J_8, J_9, J_{10}, J_{11}, and J_{12} in Figure 1b). The readings of the hand proprioceptive sensors are linearly scaled in the range $[0, 1]$ (with 0 for fully extended and 1 for fully flexed finger). To take into account the fact that sensors are noisy, tactile sensors return, with 5% probability, a value different from the computed one, and 5% uniform noise is added to proprioceptive sensors.

Internal neurons are fully connected. Additionally, each internal neuron receives one incoming synapse from each sensory neuron. Each motor neuron receives one incoming synapse from each internal neuron. There are no direct connections between sensory and motor neurons. The values of sensory neurons are updated using equation 1a, the values of internal neurons with equation 1b, and the values of motor neurons with equation 1c.

$$\tau_i \dot{y}_i = \begin{cases} -y_i + gI_i; & \text{for i=1,..,22} \quad (1a) \\ -y_i + \sum_{j=1}^{30} \omega_{ji}\sigma(y_j + \beta_j); & \text{for i=23,..,30;} \quad (1b) \\ -y_i + \sum_{j=23}^{30} \omega_{ji}\sigma(y_j + \beta_j); & \text{for i=31,..,48;} \quad (1c) \end{cases}$$

with $\sigma(x) = (1 + e^{-x})^{-1}$. In these equations, using terms derived from an analogy with real neurons, y_i represents the cell potential, τ_i the decay constant, g is a gain factor, I_i the intensity of the perturbation on sensory neuron i, ω_{ji} the strength of the synaptic connection from neuron j to neuron i, β_j the bias term, $\sigma(y_j + \beta_j)$ the firing rate. τ_i with $i = 23, .., 30$, β_i with $i = 1, .., 48$, all the network connection weights ω_{ij}, and g are genetically specified networks' parameters. τ_i with $i = 1, .., 22$ and $i = 31, .., 48$ is equal to the integration time step $\Delta T = 0.01$. There is one single bias for all the sensory neurons.

The activation values y_i of motor neurons determine the state of the simulated muscles of the arm. In particular, the total force exerted by a muscle is the sum of three forces $T_A(\sigma(y_i + \beta_i), x) + T_P(x) + T_V(\dot{x})$, which are calculated on the basis of the following equations:

$$T_A = \sigma(y_i + \beta_i)\left(-\frac{A_{sh}T_{max}(x - R_L)^2}{R_L^2} + T_{max}\right) \quad (2)$$

$$A_{sh} = \frac{R_L^2}{(L_{max} - R_L)^2}$$

$$T_P = T_{max}\frac{exp\left\{K_{sh}\frac{x-R_L}{L_{max}-R_L}\right\} - 1}{exp\{K_{sh}\} - 1}$$

$$T_V = b \cdot \dot{x}$$

where $\sigma(y_i + \beta_i)$ is the firing rate of output neurons $i = 31, .., 46$. x is the current elongation of the muscle; L_{max} and R_L are the maximum and the resting length of the muscle; T_{max} is the maximum force that can be generated; K_{sh} is the passive shape factor and b is the viscosity coefficient. The parameters of the equation are identical for all fourteen muscles controlling the seven DoFs of the arm and have been set to the following values: $K_{sh} = 3.0$, $R_L = 2.5$, $L_{max} = 3.7$, $b = 0.9$, $A_{sh} = 4.34$ with the exception of parameter T_{max} which is set to $3000N$ for joint J_2, to $300N$ for joints J_1, J_3, J_4, and J_5, and to $200N$ for joints J_6 and J_7. Muscle elongation is simulated by linearly mapping within specific angular ranges the current angular position of each DoF [see 27, for details].

The joints of the hand are actuated by a limited number of independent variables through a velocity-proportional controller. That is, for the extension/flexion, the force exerted by the MP, PIP, and DIP joints (MP-A, MP-B, and PIP in the case of the thumb) are controlled by a two step process: first, θ is set equal to the firing rate $\sigma(y_i + \beta_i)$ (with $i = 45$ for the thumb movement, and $i = 46$ for the other finger movement), linearly mapped into the range $[-90°, 0°]$; second, the desired angular positions of the finger joints MP, PIP, DIP are set to θ, θ, and $(2.0/3.0) \cdot \theta$ respectively. For the thumb, its movement towards the other fingers (i.e., the extra DoF in MP joints) corresponds to the desired angle of $-(2.0/3.0) \cdot \theta$. The DoFs that regulate the abduction/adduction movements of the fingers are not actuated.

The activation values y_i of output neurons $i = 47, 48$ are used to categorise the shape of the object (i.e., to produce different output patterns for different object types, see also Section VI).

V. The Evolutionary Algorithm

A simple generational genetic algorithm is employed to set the parameters of the networks [see 29]. The initial population contains 100 genotypes. Generations following the first one are produced by a combination of selection with elitism, and mutation. For each new generation, the 20 highest scoring individuals ("the elite") from the previous generation are retained unchanged. The remainder of the new population is generated by making 4 mutated copies of each of the 20

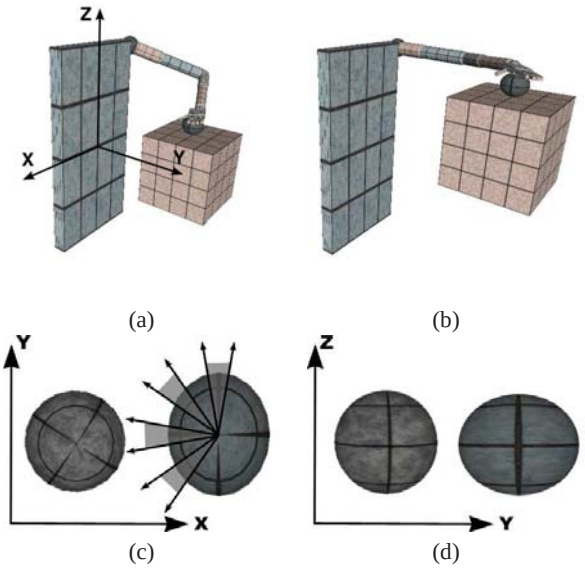


Fig. 2: (a) Position A; angle of joints $J_1, .., J_7$ are $\{-50°, -20°, -20°, -100°, -30°, 0°, -10°\}$ (b) Position B; angle of joints $J_1, .., J_7$ are $\{-100°, 0°, 10°, -30°, 0°, 0°, -10°\}$; (c) the sphere and the ellipsoid viewed from above; (d) the sphere and the ellipsoid viewed from the side. The radius of the sphere is 2.5 cm. The radii of the ellipsoid are 2.5, 3.0 and 2.5 cm. In (c) the arrows indicate the intervals within which the initial rotation of the ellipsoid is set.

highest scoring individuals. Each genotype is a vector comprising 420 parameters. Each parameter is encoded with 16 bits. Initially, a random population of vectors is generated. Mutation entails that each bit of the genotype can be flipped with a 1.5% probability. Genotype parameters are linearly mapped to produce network parameters with the following ranges: biases $\beta_i \in [-4, -2]$, weights $\omega_{ij} \in [-6, 6]$, gain factor $g \in [1, 10]$ for all the sensory neurons; decay constants τ_i with $i = 23, .., 30$ are exponentially mapped into $[10^{-2}, 10^{0.3}]$ with the lower bound corresponding to the integration step-size used to update the controller and the upper bound, arbitrarily chosen, corresponds to about half of a trial length (i.e., 2 s). Cell potentials are set to 0 when the network is initialised or reset, and circuits are integrated using the forward Euler method [see 30].

VI. The Fitness Function

During evolution, each genotype is translated into an arm controller and evaluated 8 times in position A and 8 times in position B, for a total of $K = 16$ trials (see Figure 2a and 2b). For each position, the arm experiences 4 times the ellipsoid and 4 times the sphere. Moreover, the rotation of the ellipsoid with respect to the z-axis is randomly set in the range $[350°, 10°]$ in the first presentation, $[35°, 55°]$ in the second presentation, $[80°, 100°]$ in the third presentation, and $[125°, 145°]$ in the fourth presentation (see also Figure 2c). At the beginning of each trial, the arm is located in the corresponding initial position (i.e., A or B), and the state of the neural controller is reset. A trial lasts 4 simulated seconds ($T = 400$ time step). A trial is terminated earlier in case the object falls off the table.

In each trial k, an agent is rewarded by an evaluation function which seeks to assess its ability to recognise and distinguish the ellipsoid from the sphere. Note that, rather than imposing a representation scheme in which different categories are associated with *a priori* determined state/s of the categorisation neurons (i.e., neurons 47 and 48), we leave the robot free to determine how to communicate the result of its decision. That is, the agents can develop whatever representation scheme as long as each object category is clearly identified by a unique state/s of the categorisation neurons. This system has also the advantage that it scales up to categorisation tasks with objects of more than two categories, without having to introduce structural modifications to the agent's controller. More precisely, we score agents on the basis of the extent to which the categorisation outputs produced for objects of different categories are located in non-overlapping regions of a two dimensional categorisation space $C \in [0, 1] \times [0, 1]$. The categorisation and the evaluation of the agent's discrimination capabilities is done in the following way:

- in each trial k, the agent represents the experienced object (i.e., the sphere S or the ellipsoid E) by associating to it a rectangle R_k^S or R_k^E whose vertices are:

  the bottom left vertex:

  $$(\min_{0.95T<t<T} \sigma(y_{47}(t) + \beta_{47}), \min_{0.95T<t<T} \sigma(y_{48}(t) + \beta_{48}))$$

  the top right vertex:

  $$(\max_{0.95T<t<T} \sigma(y_{47}(t) + \beta_{47}), \max_{0.95T<t<T} \sigma(y_{48}(t) + \beta_{48}))$$

- the sphere category, referred to as C^S, corresponds to the minimum bounding box of all R_k^S; the ellipsoid category, referred to as C^E, corresponds to the minimum bounding box of all R_k^E.

The final fitness FF attributed to an agent is the sum of two fitness components F_1 and F_2. F_1 rewards the robots for touching the objects, and corresponds to the average distance over a set of 16 trials between the centre of the palm and the experienced objects. F_2 rewards the robots for developing an unambiguous category representation scheme on the basis of the position in a two-dimensional space of C^S and C^E. F_1 and F_2 are computed as follows:

$$F_1 = \frac{1}{K} \sum_{k=1}^{K} \left(1 - \frac{d_k}{d_{max}}\right), \text{ with } K = 16; \quad (3)$$

$$F_2 = \begin{cases} 0 & \text{if } F_1 \neq 1; \\ 1 - \frac{area(C^S \cap C^E)}{\min\{area(C^S), area(C^E)\}} & \text{otherwise;} \end{cases} \quad (4)$$

with d_k the euclidean distance between the object and the centre of the palm at the end of the trial k; d_{max} the maximum distance the centre of the palm can reach from the object when located on the table. $F_2 = 1$ if C^S and C^E do not overlap

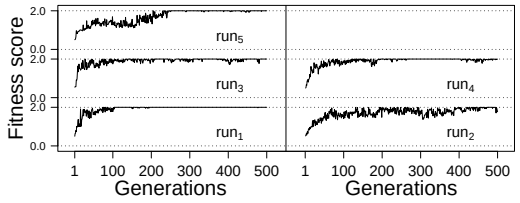


Fig. 3: Graph showing the fitness of the best agents at each generation of the five evolutionary runs that managed to generate highest score individuals for at least 10 consecutive generations: run_1, run_2, run_3, run_4, run_5.

(i.e., if $C^S \cap C^E = \emptyset$). The fact that, for each individual, F_1 must be 1 to be rewarded with F_2, constrains evolution to work on strategies in which the palm is constantly touching the object. This condition has been introduced because we thought it represents a pre-requisite for the ability to perceptually discriminate the shape of the objects. However, alternative formalisms which encode different evolutionary selective pressures may work as well.

VII. Results

Ten evolutionary simulations, each using a different random initialisation, were run for 500 generations. Figure 3 shows the fitness of the best agent at each generation for the five evolutionary runs that managed to generate highest score individuals for at least 10 consecutive generations. The other five runs failed to achieve this first objective. A quick glace at these curves indicates that run_1 reaches very quickly (in about 100 generations) a plateau on the highest fitness score and keeps on generating highest score agents until the end of evolution. run_2 run_3, run_4, run_5 also generate highest score agents but they need more generations and the solutions seem to be more sensitive to the effect produced by those parameters of the task randomly initialised and/or by noise. Although all the agents with the highest fitness are potentially capable of accomplishing the task, the effectiveness and the robustness of their collective strategies have to be further estimated with more severe post-evaluation tests. In the next Section, we show the results of a series of post-evaluation tests aimed at estimating the robustness of the best evolved discrimination strategies chosen from run_1, run_2, run_3, run_4, and run_5. In Section VII-B, we show the results of post-evaluation tests aimed at estimating the role of different sensory channels for categorisation. Finally, in Section VII-C, we analyse the dynamics of the best evolved agents categorisation strategy. It is important to note that, although all the post-evaluation analyses have been carried out on all the best evolved agents, for the sake of space, for several tests we include only the results concerning the performances of one of these agents[1]

[1]An exhaustive description of the analyses carried out on all the best evolved agents, results of tests not shown in the paper, further simulations as well as movies of the bests evolved strategies can be found at http://laral.istc.cnr.it/esm/active_perception.

A. Robustness

To verify to what extent the robots are able to discriminate between the two types of object regardless the initial orientation of the ellipsoid object, we run post-evaluation tests (referred to as test P) in which we systematically vary the ellipsoid initial orientation. More precisely in test P, an agent is demanded to distinguish for 360 times the two objects placed in position A, and for 360 times placed in position B. In each position, the agent experiences half of the times the sphere (i.e., for 180 trials) and half of the times the ellipsoid (i.e., for 180 trials). Moreover, trial after trial, the initial orientation of the ellipsoid around the z-axis changes of $1°$, from $0°$ in the first trial to $179°$ in the last trial. For each run, we selected and post-evaluated 10 agents chosen among those with the highest fitness. It is important to note that these agents are selected from evolutionary phases in which the run managed to generate highest score individuals for at least 10 consecutive generations. Table I shows the results of the best agent A_j chosen from run_i, with $j, i = 1, ..., 5$.

Note that, compared to the evolutionary conditions, in which the agents are allowed to perceive the ellipsoid only 4 times with 4 different initial orientations, P is a severe test. The results unambiguously tell us whether or not the five selected highest fitness agents are capable of distinguishing and categorising the ellipsoid from the sphere in a much wider range of initial orientations of the former object. For each selected agent, test P is repeated 5 times (i.e., P_i with $i = 1, .., 5$), with each repetition differently seeded to guaranteed random variations in the noise added to sensors readings.

The performance of the agent A_j at test P_i is quantitatively established by considering all the responses given by A_j over 3600 trials (i.e., 720 trials per test P_i, repeated 5 times, with $i, j = 1, ..., 5$). In each post-evaluation trial, the response of the agent is based on the firing rates of neurons 47 and 48 during the last 20 time steps (i.e., $0.95T < t < T$) of each trail k. In particular, the smallest and the highest firing rates recorded by both neurons are used to define the bottom left and the top right vertices of a rectangle, as illustrated in Section VI. At the end of each test P_i, we have 360 rectangles associated to trials in which the agent experienced the sphere (i.e., rectangles R^S_k with $k = 1, .., 360$), and 360 rectangles associated to trials in which the agent experienced the ellipsoid (i.e., rectangles R^E_k with $k = 1, .., 360$). At the end of the five post-evaluation tests P_i, we build five pairs of non-overlapping minimal bounding boxes (i.e., C^S_i and C^E_i), a pair for each test i, as explained in Section VI. At this point, we take as a quantitative estimate of the robustness of an agent categorisation strategy, the highest number of R^S_k and R^E_k rectangles that can be included in C^S_i and C^E_i respectively, by fulfilling the condition that none of the C^S_i overlaps with any of the C^E_i. Table I shows, for each selected agent and for each test P_i, the number of rectangles (R^S_k and R^E_k) for post-evaluated agent, and for post-evaluation test, that can be included in C^S_i and C^E_i by fulfilling the condition that none of the C^S_i overlaps with any of the C^E_i. The last row of this Table tells us that, for agent A_1, A_3, A_4, and A_5, the total number of rectangles that can be included by the minimal

bounding boxes without breaking the non-overlapping rule is extremely high, with a percentage of success over 97%. These four agents are quite good in discriminating and categorising the sphere and the ellipsoid in a much wider range of initial orientations of the ellipsoid. Agent A_2, whose performance is slightly worst, is excluded from all further post-evaluation tests.

The agents with a performance at the first test P above 95% (i.e., A_1, A_3, A_4, and A_5) undergo a further series of tests P in circumstances in which i) the length of the longest radius of the ellipsoid progressively increases/decreases (see Figure 4a); ii) the length of the radius of the sphere progressively increases/decreases (see Figure 4b); iii) the initial position of the object and of the hand varies (see Figure 4c). In these as well as in all the other post-evaluation tests we describe from now on concerning A_1, A_3, A_4, and A_5, a trial k can: (i) successfully terminate if the R_k^E, built as illustrated above, completely falls within the agent's two-dimensional space delimited by the five bounding boxes C_i^E built during the first test P; (ii) unsuccessfully terminate with a sphere response if the R_k^E completely falls within the agent's two-dimensional space delimited by the five bounding boxes C_i^S built during the first test P; (iii) unsuccessfully terminate with a none response, if the R_k^E, completely falls outside the agent's two-dimensional space delimited by the ten bounding boxes $C_i^S \cap C_i^E$ built during the first test P.

As far as it concerns tests in which the length of the longest radius of the ellipsoid progressively increases/decreases, we notice that distortions that further increase the longest ellipsoid radius up to 1 cm, are rather well tolerated by the agents, with A_1 and A_5 that manage to reliably differentiate the two objects with a success rate higher than 90%. Distortions that tend to reduce the longest radius of the ellipsoid are clearly disruptive for all the agents, with an expected 50% success rate when the ellipsoid is reduced to a sphere. In tests in which the ellipsoids have a radius progressively shorter that the radius of the sphere, the performance of all the agents are quite disrupted (see Figure 4a).

As far as it concerns tests in which the length of the radius of the sphere progressively increases/decreases, we notice that these distortions are particularly disruptive for all the agents except for A_5. This agent is not as disrupted as the other agents in those tests in which the sphere becomes progressively smaller, and it is very successful in tests in which the radius of the sphere is at least 7 millimetres longer than the longest radius of the ellipsoid (see Figure 4b).

Finally, in a further series of post-evaluation tests we estimated the robustness of the best evolved strategies in tests in which the initial positions of object and of the arm change. To simplify our analysis, we focused only on those circumstances in which the movement of the arm respect to the initial positions experienced during evolution are determined by displacements of only one joint at time (see Figure 4c). Although the results are quite heterogeneous, there are some features which are shared by all the agents. First, displacements of joint J_1 for position A are tolerated quite well. Second, wider the displacement, bigger the performance drop, with the exception of J_4 for agents A_1 A_3 A_4, in which displacements that tend to progressively bring the hand/object closer to the body result in a better performance for both positions. It is important to note that, A_4 is particularly sensitive to disruptions to joint J_1 and J_2 for position B, and joint J_6 for position A.

TABLE I: The table shows, for post-evaluated agent (A_j with $j = 1, ..., 5$), and for post-evaluation test (P_i with $i = 1, .., 5$), the number of rectangles R_k^E and R_k^S that can be included in bounding boxes C_i^E and C_i^S, respectively, by fulfilling the condition that none of the C_i^E overlaps with any of the C_i^S. The last row indicates the total number of correct categorisation choices and percentage of success over 3600 evaluation trials. See the text for further details.

	A_1		A_2		A_3	
	R_k^E	R_k^S	R_k^E	R_k^S	R_k^E	R_k^S
P_1	357	360	310	351	340	358
P_2	359	360	311	347	342	358
P_3	356	360	312	349	343	356
P_4	357	360	304	353	341	355
P_5	358	360	303	348	349	356
Tot./(%)	3587 / 0.99%		3288 / 0.91%		3498 / 0.97%	

	A_4		A_5	
	R_k^E	R_k^S	R_k^E	R_k^S
P_1	347	356	355	354
P_2	356	358	356	355
P_3	348	355	356	354
P_4	342	354	354	355
P_5	349	355	353	353
Tot./(%)	3520 / 0.98%		3545 / 0.98%	

B. The role of different sensory channels for categorisation

To understand the mechanisms which allow agents A_1, A_3, A_4, and A_5, to solve their task, we first established the relative importance of the different types of sensory information available through arm proprioceptive sensors (i.e., I_i with $i = 1, ..., 7$, see also Figure 1c), tactile sensors (i.e., I_i with $i = 8, ..., 17$, see also Figure 1c), and hand proprioceptive sensors (i.e., I_i with $i = 18, ..., 22$, see also Figure 1c). This has been accomplished by measuring the performance displayed by the agents in a series of *substitution tests* in which one type of sensory information experienced by each agent during the interaction with an ellipsoid has been replaced with the corresponding type of sensory information previously recorded in trials in which the agent was interacting with a sphere. In these tests, each agent experiences the ellipsoid in all the initial rotations (i.e., from $0°$ to $179°$) excluding those for which, given the randomly chosen seed for the tests, its responses turned out to be wrong in the absence of any type of substitution (i.e., the rectangle R_k^E did not fall within any of the five bounding boxes C_i^E resulted from the test P described in Section VII-A). For each ellipsoid initial orientation, each *substitution tests* is repeated 180 times. The rational behind these tests is that any performance drop caused

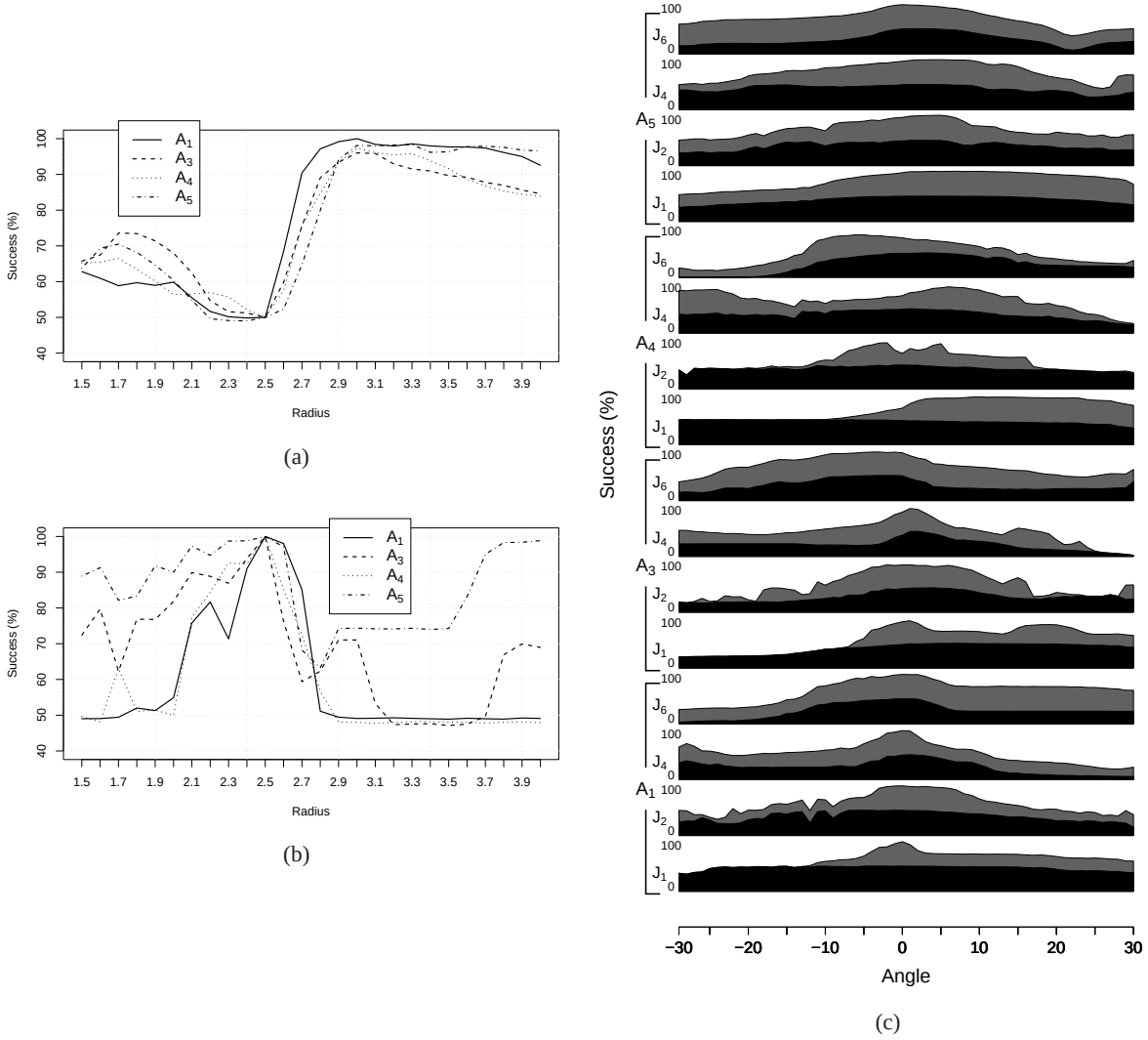


Fig. 4: Graphs showing the percentage of success in post-evaluation tests in which (a) the length of the longest radius of the ellipsoid progressively increases/decreases; (b) the length of the radius of the sphere progressively increases/decreases; (c) the initial position of the object and of the hand varies. Black is for position A, and grey for position B. See also the text for further details.

by the replacement of different type of sensory information provides an indication of the relative importance of that sensory channel on the categorisation process.

The results of this first series of *substitution tests* tell us that, for all the agents, the replacement of the sensory information originated by the arm proprioceptive sensors and by the hand proprioceptive sensors in position A, only marginally interfere with their performance. That is, for position A, the agents undergo a substantial performance drop only due to replacement of tactile sensation (see Figure 5 black columns in correspondence of tactile sensors). The clear performance drop in these *substitution tests* concerning tactile sensation clearly indicates that, for position A, the agents heavily rely on tactile sensation to distinguish the ellipsoid from the sphere and to correctly perform the categorisation task.

For position B, the results are slightly more heterogeneous. For agent A_1, the results of *substitution tests* indicate that both the replacement of tactile sensations and of the hand proprioceptive sensor produce about 20% performance drop (see Figure 5 white columns in correspondence of tactile and

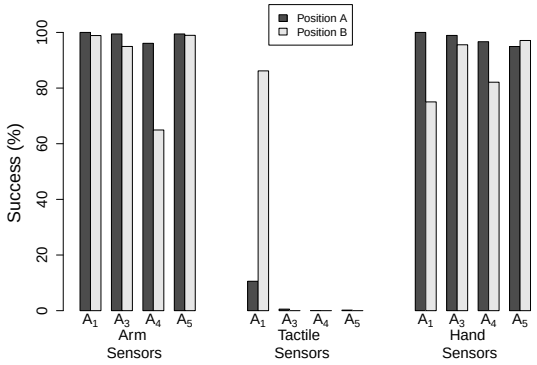


Fig. 5: Graphs showing, for agents A_1, A_3, A_4, and A_5, the results of *substitution tests* concerning the readings of arm proprioceptive sensors, tactile sensors, and hand proprioceptive sensors for position A (see black columns) and for position B (see grey columns).

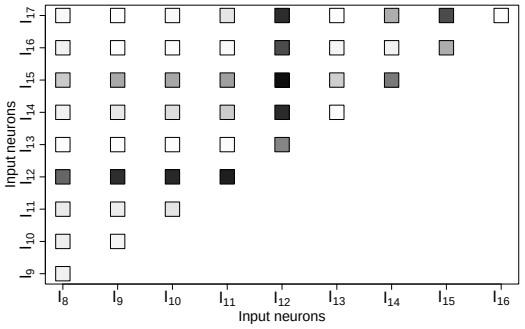


Fig. 6: Graphs showing the results of substitution tests concerning the readings I_i with $i = 8, .., 17$ of all the possible combinations of two elements of the tactile sensors for position A. Each square is coloured in shades of grey. The grey scale is proportional to the percentage of success, with white indicating combinations in which the agent is 100% successful, and black combinations in which the agent is 100% unsuccessful.

hand sensors). For the other agents, tactile sensation keeps on being extremely important for the correct categorisation of the objects (see Figure 5 white columns in correspondence of tactile sensors). However, for agent A_4, the replacement of the arm and of the hand proprioceptive sensor produces a performance drop of about 40% in the case of the arm and 20% in the case of the hand sensors (see Figure 5 white columns in correspondence of arm and hand sensors). Thus, we conclude that, for agent A_1 the categorisation of the ellipsoid in position B is performed by exploiting information distributed over two sensory channels, that is tactile and hand sensors. The information provided by the two sensory channels seems to be fused together in a way that, for several orientations, the lack or the unreliability of information from one channel can be compensated by the availability of reliable information from the other channel (data not shown). The other agents seem to strongly rely on tactile sensation, with agent A_4 that makes also use of arm and hand sensation to discriminate the objects.

Given that, tactile sensation is the major source of discriminating cues in order to distinguish spheres from ellipsoids in position A, for all the selected agents, and in position B for A_3, and A_5, we pursue further investigations, to see whether among the tactile sensors, there are any whose activations play a predominant role in the categorisation task. We begin by running *substitution tests* in which we applied the kind of replacements described above only to single tactile sensors. It turned out that the categorisation abilities of the agents are not hindered by replacements which selectively hit the functioning of single tactile sensors. The performance of all the agents remain largely above 90% success rate (data not shown).

Thus, we proceeded by running *substitution tests* in which we applied replacements to all the possible combinations of two elements of the tactile sensors. Although this analysis have been carried out on all the agents for position A, and on agents A_3, and A_5, for position B, in the following we illustrate in details only the results of A_1 (i.e., the best performing agent, see Table I) for position A[1]. The results are shown in Figure 6, in which, the grey scale of the small squares is proportional to the percentage of success, with white indicating combinations in which the agent is 100% successful, and black combinations in which the agent is 100% unsuccessful. This *substitution tests* did not produce clear cut results. However, by looking at Figure 6 we can see that there are specific sensors which, when disrupted in combination with any other sensor, produce a clear performance drop. In particular, disruptions applied to the reading of the tactile sensors placed on the third phalange of the middles finger (i.e., I_{12}), and in minor terms, disruption applied to the reading of the tactile sensors placed on the first phalange of the ring finger (i.e., I_{15}) induce the agent to mistake the ellipsoid for the sphere. We conclude that, agent A_1 heavily relies on the patterns of activation of tactile sensors in which the reading of I_{12} and I_{15} are particularly important to distinguish the ellipsoid from the sphere. For what concerns the other agents, the performance of agent A_3 drops in position A when substitutions concern the reading of I_{10} in combination with any other tactile sensor. In position B, a performance drop is recorded when substitutions concern the reading of I_8 or I_{12} in combination with any other sensor. Agent A_4 in position A is particularly disrupted by substitutions concerning the reading of I_{11} or I_{12} in combination with any other sensor. Agent A_5 in position A is disrupted by substitution concerning the reading of I_{12} with any other sensor, and of I_{12} or I_{17} with any other sensor in position B. In conclusion, in those circumstances in which we observed a predominance of tactile sensation to carry out the categorisation task, the agents tend to rely on combinations of tactile sensors, with the tactile sensor placed on the third phalange of the middles finger basically more relevant than the other sensors for all the agents (data not shown).

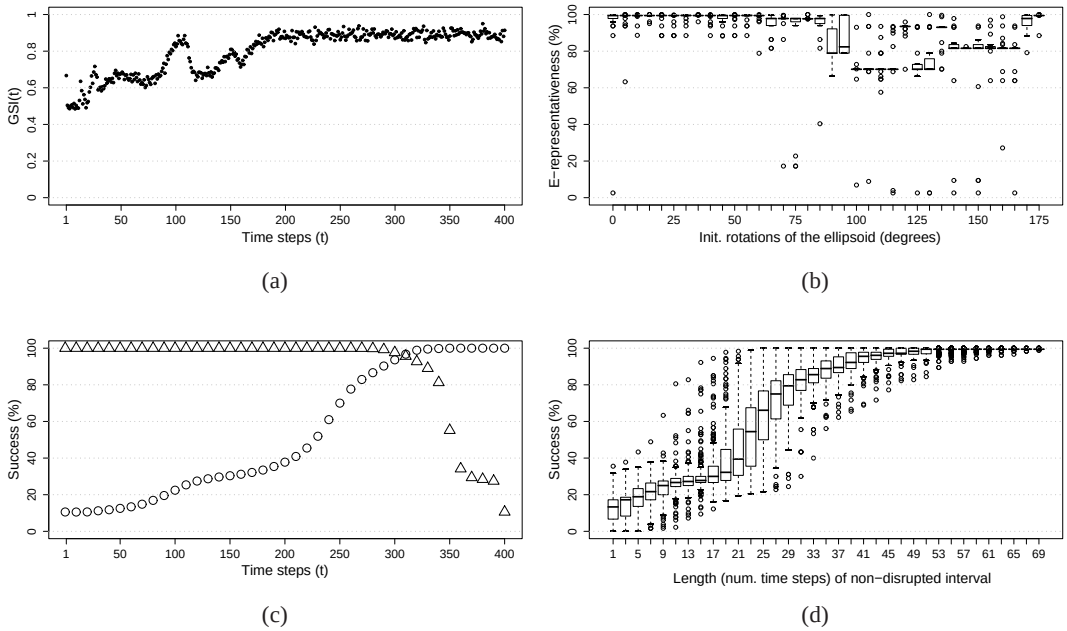


Fig. 7: Graphs showing: (a) the Geometric Separability Index (GSI); (b) the *E-representativness* of the tactile sensors patterns recorded in the last 20 time steps of 180 different trials with the ellipsoid; (c) the percentage of success in *pre-substitution tests* (see triangles) and *post-substitution tests* (see empty circles); (d) the percentage of success at the *window-substitution tests*.

C. On the dynamics of the categorisation process

In this section, we focus our attention on the dynamics of the categorisation process. More specifically, we analyse: (i) to what extent the sensory stimuli experienced while the agents interact with the objects provide the regularities required to categorise the objects; (ii) to what extent the agents succeed in self-selecting discriminative stimuli (i.e., stimuli that can be unambiguously associated with either category); (iii) how long the agents need to interact with the object before being able to tell whether they are touching a sphere or an ellipsoid; (iv) whether the categorisation process occurs instantaneously by exploiting the regularities provided by single unambiguous sensory patterns or whether it occurs over time by integrating the regularities provided by several stimuli.

To answer these questions we run qualitative and quantitative tests. The former are just observations of the trajectories of the categorisation outputs in the two-dimensional categorisation space $\{\sigma(y(t)_{47}+\beta_{47}), \sigma(y(t)_{48}+\beta_{48})\}$, in single trials. The latter are tests that further explore the dynamics of the categorisation processes by taking advantage of the fact that in both positions almost all the best evolved agents exploit tactile sensation to carry out the task. The quantitative tests have been carried out on all the agents for position A, and on agents A_3, and A_5, for position B. In the following, we illustrate in details only the analysis concerning A_1 (i.e., the best performing agent, see Table I) for position A. However, it turned out that, successful categorisation strategies are very similar from a behavioural point of view, and in terms of the mechanisms exploited to perform the task. Therefore, the reader should consider the operational description of A_1 representative of the categorisation strategies of A_3, A_4, and A_5 in position A, and of A_3 and A_5 in position B[1].

The first two tests aim at establishing to what extent the stimuli experienced by A_1 during its interactions with the objects provide the regularities required to categorise the objects. We begin our analysis by computing a slightly modified version of the Geometric Separability Index (hereafter, referred to as *GSI*). The *GSI*, originally proposed by Thornton [31], is an estimate of the degree to which tactile sensors readings associated with the sphere or with the ellipsoid are separated in sensory space. We built four hundred data sets, one for each time step with the ellipsoid (i.e., $\{\tilde{I}_k^E\}_{k=1}^{180}$), and four hundred data sets, one for each time step with the sphere (i.e., $\{\tilde{I}_k^S\}_{k=1}^{180}$). Where, $\tilde{I}_k^E$ is the tactile sensors reading experienced by the agent while interacting with the ellipsoid at time step t of trial k; and $\tilde{I}_i^E$ is the tactile sensors reading experienced by the agent while interacting with the sphere at time step t of trial k. Recall that, trial after trial, the initial rotation of the ellipsoid around the z-axis changes of $1°$, from $0°$ in the first trial to $179°$ in the last trial. Each trial is differently seeded to guaranteed random variations in the noise added to sensors readings. At each time step t, the *GSI*

is computed in the following:

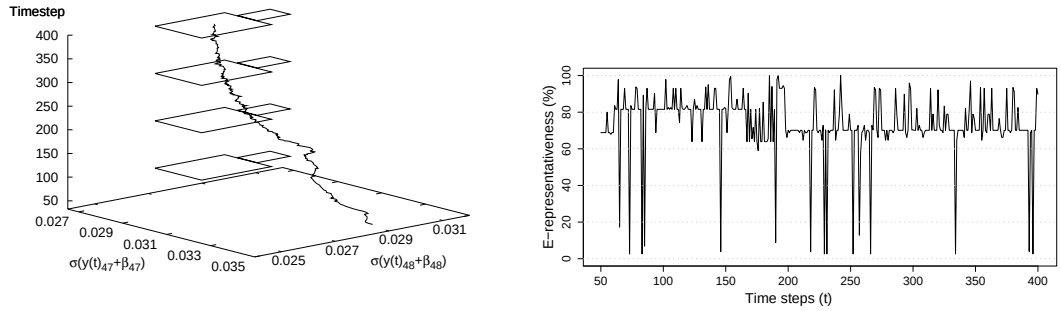
$$
\begin{aligned}
GSI(t) &= \frac{1}{K}\sum_{k=1}^{K} z_k, \text{ with } K = 180; \\
z_k &= \begin{cases} 1 & if m^{EE} < m^{ES}; \\ 0 & if m^{EE} > m^{ES}; \\ \frac{u}{u+v} & otherwise; \end{cases} \\
m^{EE} &= \min_{\forall j \neq k}(H(\tilde{I}_k^E, \tilde{I}_j^E)) \\
m^{ES} &= \min_{\forall j}(H(\tilde{I}_k^E, \tilde{I}_j^S)) \\
u &= |\{\tilde{I}_j^E : H(\tilde{I}_k^E, \tilde{I}_j^E) = m^{EE}\}_{\forall j \neq k}| \\
v &= |\{\tilde{I}_j^S : H(\tilde{I}_k^E, \tilde{I}_j^S) = m^{ES}\}_{\forall j}|
\end{aligned}
\tag{5}
$$

where H(x, y) is the Hamming distance between tactile sensors readings. $|x|$ means the cardinality of the set x. *GSI* equal to 1 means that at time step t the closest neighbourhood of each $\tilde{I}_k^E$ is one or more elements of the set $\tilde{I}_k^E$. *GSI* equal to 0 means that at time step t the closest neighbourhood of each $\tilde{I}_k^E$ is one or more elements from the set $\tilde{I}_k^S$. As shown in Figure 7a, for agent A_1 position A, the *GSI(t)* tends to increase from about 0.5 at time step 1 to about 0.9 at time step 200, and remains around 0.9 until time step 400. This trend suggests that during the first 200 time steps, the agent acts in a way to bring forth those tactile sensors readings which facilitate the object identification and classification task. In other words, the behaviour exhibited by the agent allows it to experience two classes of sensory states which tend to become progressively more separated in the sensory space. However, the fact that the *GSI* does not reach the value of 1.0 indicates that the two groups of sensory patterns belonging to the two objects are not fully separated in the sensory space. In other words, some of the sensory patterns experienced during the interactions with an ellipsoid are very similar or identical to sensory patterns experienced during interactions with the sphere and vice versa.

To analyse in more details to what extent the stimuli experienced by the agent could be associated to the correct or the wrong category we calculated the *E-representativness*. The latter refers to the probability with which a single tactile sensors pattern is associated to the category ellipsoid. The *E-representativness* is computed on a set of 32.400 trials, given by repeating 180 times each the 180 trials corresponding to 180 different ellipsoid initial orientations, from 0° to 179°. During these trials, for each single tactile sensors pattern, we recorded the number of times each pattern appears during interaction with the ellipsoid (N) and during interactions with the sphere (M). The *E-representativness* of a single pattern is given by ($\frac{N}{N+M}$). It is important to notice that an *E-representativness* of 1.0 or 0.0 corresponds to fully discriminative stimuli that can be unambiguously associated with the ellipsoid or the sphere category, respectively, while 0.5 *E-representativness* corresponds to fully ambiguous stimuli. The graph in Figure 7b refers to the *E-representativness* of the last 20 patterns (i.e, patterns recorded from time step 380 to time step 400) of single successful trials of test P described in Section VII-A. Each trial refers to a different initial orientation of the ellipsoid. A quick glance at Figure 7b indicates that there are trials in which the agent has to deal with tactile sensors patterns that have very low *E-representativness*. That is, they are very weakly associated with the ellipsoid. Patterns with very low *E-representativness* tend to appear in trials in which the initial orientation of the ellipsoid is chosen in the interval 75°, ..., 175°. These patterns may have at least two not mutually excluding origins: i) they may come from the fact that the agent is not able to effectively position the object in a way to unequivocally say whether is a sphere or an ellipsoid; ii) they may be determined by the noise injected into the system. The fact that agent A_1 succeeds in correctly discriminating the category of the objects also during trials in which it does not experience fully discriminating stimuli indicates that the problem is solved by integrating over time the partially conflicting evidences provided by sequences of stimuli. In fact, if the agent employs a reactive strategy (i.e., no need of memory structure), it would be deceived by those sensor patterns, very strongly associated with the sphere, that appear in interaction with the ellipsoid. Under this circumstance an agent that employs a reactive strategy would mistake the ellipsoid for a sphere. Since, in spite of the deceiving patterns, the agent is 100% successful, it looks like the agent is employing a discrimination strategy which uses the dynamic properties of its controller.

Other evidence that supports the integration over time hypothesis come from additional analyses conducted employing further types of *substitution tests*. In particular, we substitute, for a certain time interval, tactile sensors patterns experienced by A_1 in interaction with the ellipsoid with those experienced in interaction with the sphere. In a first series of tests, referred to as *pre-substitution tests*, substitutions have been applied from the beginning of each trial up to time step t where t = 1,...,400. In a second series of tests, referred to as *post-substitution tests*, substitutions have been applied from time step t, where t = 1, ..., 400, to the end of a trial t=400. Each test has been repeated at intervals of 10 time steps. For agent A_1 position A, the results of *pre-substitution tests* and *post-substitution tests* are illustrated in Figure 7c. This graph shows that, regardless of the rotation of the ellipsoid, pre-substitutions which do not affect the last 100 time steps do not cause any performance drop. For *pre-substitution tests* that involve more than 300 time steps the amount of performance drop is higher for longer substitution periods (see triangles in Figure 7c). Similarly, the agent does not incur in any performance drop if post-substitutions affect less than 100 time steps. For *post-substitution tests* that affect more than the last 100 time steps the amount of performance drop is higher for longer substitution periods (see empty circles in Figure 7c).

By looking at the results of *pre-substitution tests* and *post-substitution tests*, we suppose that the agent is integrating sensory states over time for a certain amount of time around time step 310. In particular, the results shown in Figure 7c seem to indicate that, for what concerns agent A_1 position A, the interactions between the agent and the objects can be divided into three temporal phases that are qualitatively different from the point of view of the categorisation process: (i) an initial phase whose upper bound can be approximately fixed at time step 250, in which the categorisation process



(a) (b)

Fig. 8: Graphs showing: (a) trajectories of the decision outputs in the two-dimensional categorisation space ($\sigma(y(t)_{47} + \beta_{47})$, $\sigma(y(t)_{48} + \beta_{48})$), with (a) t = 50, ..., 400, recorded in a successful trial with the ellipsoid initially orientated at $115°$. Big and small rectangles at 100, 200, 300, and 400 time steps indicate the bounding box of the ellipsoid and sphere category, respectively; (b) the *E-representativness*of the tactile sensory patterns recorded in a successful trial with the ellipsoid initially orientated at $115°$.

begins but in which the categorisation answer produced by the agent is still reversible; (ii) an intermediate phase whose upper bound can be approximately fixed at time step 350, in which very often a categorisation decision is taken on the basis of all previously experienced evidences; and (iii) a final phase in which the previous decision (which is now irreversible) is maintained. The fact that the categorisation decision formed by A_1 during the initial phase is not definitive yet is demonstrated by the fact that substitutions of the critical sensory stimuli performed during this phase do not cause any performance drop (see Figure 7c, triangles). The fact that the intermediate phase corresponds to a critical period is demonstrated by the fact that *pre-substitution tests* and *post-substitution tests* affecting this phase produce a significant performance drop (see Figure 7c). The fact that A_1 takes an ultimate decision during the intermediate phase is demonstrated by the fact that *post-substitution tests* affecting the last 80 time steps, approximately, do not produce any drop in performance (see Figure 7c, empty circles).

In a further series of tests, we looked at whether there is and eventually how big it is the hypothesised temporal phase in which the agent is supposed to integrate tactile sensors states. To look at this issue, we employ the *window-substitution tests*. In these tests, substitutions are applied before and after a temporal window centred around time step 310. The length of the temporal window with no substitutions can varies from 1 time step (i.e., no substitution at time step 310) to 69 time steps (i.e., no substitution from time step 276 to 344). As shown in Figure 7d, wider the window with no substitutions higher the performance of the agent, with 100% success rate when no substitutions are applied to a temporal phase of about 50 time steps or longer. Although the graph in Figure 7d does not exclude the possibility that the agent employs an instantaneous categorisation process, the graph seems to suggest that the performance of the agent is in a way correlated to the amount of empirical evidences it manages to gather over time starting from about time step 270 until time step 340.

Finally, additional evidence in support of a dynamic categorisation process based on the integration of tactile sensation over time come from a qualitative analysis of the trajectories of the categorisation outputs in the two-dimensional categorisation space $\{\sigma(y(t)_{47} + \beta_{47}), \sigma(y(t)_{48} + \beta_{48})\}$, in single trials. Figure 8a shows the trajectory recorded by A_1 in a trial in which the initial orientation of the ellipsoid was $115°$. As we can see, A_1 moves rather smoothly in the categorisation space by reaching in slightly less than 2 s (200 time steps) the corresponding bounding box. If we now look at Figure 8b, we see that during the interaction with the ellipsoid A_1 experiences: (i) few stimuli with an high percentage of *E-representativness* (i.e., stimuli that are experienced in interaction with an ellipsoid object most of the times); (ii) several stimuli with an intermediate level of *E-representativness* (i.e., stimuli that are experienced in interaction with the ellipsoid and the sphere in about the 3/4 and 1/4 of the cases, respectively); and (iii) few stimuli with a low percentage of *E-representativness* (i.e., stimuli that are experienced in interaction with a sphere object most of the times). If we visually compare Figure 8a with Figure 8b, we notice that the experienced sensory patterns with different percentage of *E-representativness* appear to drive the categorisation output in different regions of the the categorisation space, corresponding to the ellipsoid and the sphere bounding box, respectively. The final position of the categorisation output (i.e., the categorisation decision) therefore is not determined by a single or few selected patterns but is rather the result of a process extended over time in which partially conflicting evidence provided by the experienced tactile sensation is integrated over time. Similar dynamics have been observed by inspecting all other trials. Given this evidence, we conclude that the performance of all best evolved agents in position A, and of agent A_3 and A_5 in position B, is the result of a dynamic categorisation process based on the integration of tactile sensation over time.

VIII. Discussion and Conclusions

In this paper, we described an experiment in which a simulated anthropomorphic robotic arm acquires an ability to categorise un-anchored spherical and ellipsoid objects placed in different positions and orientations over a planar surface. The agents neural controller has been trained through an evolutionary process in which the free parameters of the neural networks are varied randomly and in which variations are retained or discarded on the basis on their effects on the overall ability of the robots to carry out their task. This implies that the robots are left free to determine (i) how to interact with the external environment (by eventually modifying the environment itself); (ii) how the experienced sensory stimuli are used to discriminate the two categories; and (iii) how to represent in the categorisation space each object category.

The analysis of the obtained results indicates that the agents are indeed capable of developing an ability to effectively categorise the shape of the objects despite the high similarities between the two types of objects, the difficulty of effectively controlling a body with many DoFs, and the need to master the effects produced by gravity, inertia, collisions etc. More specifically, the best individuals display an ability to correctly categorise the objects located in different positions and orientations already experienced during evolution, as well as an ability to generalise their skill to objects positions and orientations never experienced during evolution. Moreover, the agents are robust enough to deal with categorisation tasks in which the longest radius of the ellipsoid is progressively increased. Other distortions on the original objects dimensions result more disruptive. These results prove that the method proposed can be successfully applied to scenarios which appear to be more complex than those investigated in previous works based on similar methodologies.

The analysis of the best evolved agents indicates that one fundamental skill that allows them to solve the categorisation problem consists in the ability to interact with the external environment and to modify the environment itself so to experience sensory states which are progressively more different for different categorical contexts. This result represents a confirmation of the importance of sensory-motor coordination, and more specifically of the active nature of situated categorisation, already highlighted in previous studies [e.g., 20, 23]. On the other hand, the fact that sensory-motor coordination does not allow the agents to experience fully discriminative stimuli demonstrates how in some cases sensory-motor coordination should be complemented by additional mechanisms. Such mechanism, in the case of the best evolved individuals, consists in an ability to integrate the information provided by sequences of sensory stimuli over time. More specifically, we brought evidence showing that agent A_1 categorise the current object as soon as it experiences useful regularities and that the categorisation process is realised during a significant period of time (i.e., about 50 time steps) in which the agent keeps using the experienced evidence to confirm and reinforce the current tentative decision or to change it. Similar strategies have been observed in the other three best evolved agents (data not shown[1]). On this aspect see also [22, 33, 34].

The importance of the ability to integrate the regularities provided by sequences of stimuli is also confirmed by the results obtained in a control experiment, replicated 10 times, in which the agents were provided with reactive neural controllers (i.e., neural networks without recurrent connections, with simple logistic internal neurons, and in which all other parameters were kept equal to those described in Section IV). Indeed the performance displayed by the best evolved individuals in this control experiment were significantly worse than those observed in the basic experiment in which the agents were allowed to keep information about previously experienced sensory states (data not shown[1]). Although we cannot exclude that different experimental scenarios (e.g., scenarios involving agents provided with different neural architecture and/or different physical characteristics of the agents) could lead to qualitatively different results, the analysis of the results obtained in this specific scenario overall indicates that the task does admit pure reactive solutions or alternatively that such solutions are hard to synthesise through an evolutionary process. This may also be due to functional constraints which limit the movements of the robotic arm (e.g., the fact that the fingers can not be extended/flexed separately, or that there was no adduction/abduction movement of the fingers), as well as other implementation details (e.g., the dimensions of the objects with respect to the hand). This issue will be definitely investigated in future works.

The analysis of the role played by different sensory channels indicates that the categorisation process in the best evolved individuals is primarily based on tactile sensors and secondarily on hand and arm proprioceptive sensors (with arm proprioceptive sensors playing a role only for agent A_4 position B, see Figure 5). It is interesting to note that at least one of the best evolved agents (i.e., A_1) does not only display an ability to exploit all relevant information but also an ability to fuse information coming from different sensory modalities in order to maximise the chance to take the appropriate categorisation decision [see also 32]. More specifically, the ability to fuse the information provided by the tactile and hand proprioceptive sensors, for objects located in position B, allows the robot to correctly categorise the shape of the object in the majority of the cases even when one of the two sources of information is corrupted (see Figure 5).

For the future, we plan to validate the obtained results by porting the best evolved controller on the I-CUB humanoid robotic platform [see 35]. Note that, the porting may require only few changes. In particular, while structurally the simulated arm described in Section III is identical to the real I-CUB, from the functional point of view, it may not match the dynamics of the tendon actuators moving the arm of the real I-CUB. The simulation-reality gap can be closed by firstly quantitatively estimating the mismatch between simulation model and real robot and by appropriately adjusting the system to undo this mismatch. Moreover, we plan to scale up the experiment to a larger number of object categories, and to study experimental scenarios in which the robots are rewarded for the ability to perform a manipulation task (e.g., grasping different type of objects) that presumably requires categorisation rather than directly for the ability to

perceptually categorise the shape of the objects.

ACKNOWLEDGEMENT

This research work was supported by the *ITALK* project (EU, ICT, Cognitive Systems and Robotics Integrating Project, grant n° 214668). The authors thank Massimiliano Schembri, Tomassino Ferrauto and their colleagues at LARAL for stimulating discussions and feedback during the preparation of this paper.

REFERENCES

[1] S. Harnad, Ed., *Categorical Perception: The Groundwork of Cognition*. Cambridge University Press, 1987.

[2] H. Cohen and C. Lefebvre, Eds., *Handbook of Categorisation in Cognitive Science*. Elsevier, 2005.

[3] R. Beer, "Dynamical approaches to cognitive science," *Trends in Cognitive Sciences*, vol. 4, pp. 91–99, 2000.

[4] S. Nolfi, "Behavior and cognition as a complex adaptive system: Insights from robotic experiments," in *Philosophy of Complex Systems, Handbook on Foundational/Philosophical Issues for Complex Systems in Science*, C. Hooker, Ed. Elsevier, In Press.

[5] J. J. Gibson, "The theory of affordances," in *Perceiving, Acting and Knowing. Toward an Ecological Psychology*, R. Shaw and J. Bransford, Eds. Hilldale, NJ: Lawrence Erlbaum Associates, 1977, ch. 3, pp. 67–82.

[6] A. Noë, *Action in Perception*. MIT Press, Cambridge, MA, 2004.

[7] R. Pfeifer and C. Scheier, *Understanding Intelligence*. MIT Press, Cambridge, MA, 1999.

[8] S. Nolfi and D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines*. MIT Press, Cambridge, MA, 2000.

[9] I. Harvey, E. Di Paolo, R. Wood, M. Quinn, and E. Tuci, "Evolutionary robotics: A new scientific tool for studying cognition," *Artificial Life*, vol. 11, no. 1-2, pp. 79 – 98, 2005.

[10] D. Floreano, P. Husband, and S. Nolfi, "Evolutionary robotics," in *Springer Handboook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer Verlag, Berlin, Germany, 2008, pp. 1423–1451.

[11] S. Takamuku, G. Gomez, K. Hosoda, and R. Pfeifer, "Haptic discrimination of material properties by a robotic hand," in *Proceedings of the IEEE 6th International Conference on Development and Learning (ICDL 2007)*, 2007, paper nr 76.

[12] M. Johnsson, R. Pallbo, and C. Balkenius, "Experiments with haptic perception in a robotic hand," in *Advances in Artificial Intelligence in Sweden*, P. Funk, T. Rognvaldsson, and N. Xiong, Eds. Västerás, Sweden: Mälardalen University, 2005, pp. 81–86.

[13] M. Johnsson and C. Balkenius, "A robot hand with t-mpsom neural networks in a model of the human haptic system," in *Proceedings of the International Conference Towards Autonomous Robotic Systems*, M. Witkowski, U. Nehmzow, C. Melhuish, E. Moxey, and A. Ellery, Eds. Springer Verlag, Berlin, Germany, 2006, pp. 80–87.

[14] ——, "Experiments with proprioception in a self-organizing system for haptic perception," in *Proceedings of the International Conference Towards Autonomous Robotic Systems*, M. Wilson, F. Labrosse, U. Nehmzow, C. Melhuish, and M. Witkowski, Eds. Springer Verlag, Berlin, Germany, 2007, pp. 239–245.

[15] ——, "Neural network models of haptic shape perception," *Robotics and Autonomous Systems*, vol. 55, pp. 720–727, 2007.

[16] P. Dario, C. Laschi, C. Carrozza, E. Guglielmelli, G. Teti, B. Massa, M. Zecca, D. Taddeucci, and F. Leoni, "An integrated approach for the design of a grasping and manipulation system in humanoid robotics," in *Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 1, 2000, pp. 1–7.

[17] L. Natale and E. Torres-Jara, "A sensive approach to grasping," in *Proceedings of the 6th International Workshop on Epigenetic Robotics*, F. Kaplan, P. Oudeyer, A. Revel, P. Gaussier, J. N. L. Berthouze, H. Kozima, C. Prince, and C. C. Balkenius, Eds., vol. 128. Lund University Cognitive Studies, Lund, Danemark, 2006, pp. 87–94.

[18] S. Stansfield, "A haptic system for a multifimgered harnd," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 1991, pp. 658–664.

[19] C. Scheier and D. Lambrinos, "Categorization in a real-world agent using haptic exploration and active perception," in *Proceedings of the 4th International Conference on Simulation of Adaptive Behavior (SAB96)*, P. Maes, M. Mataric, J. Meyer, J. Pollack, and S. Wilson, Eds. MIT Press, Cambridge, MA, 1996, pp. 65–74.

[20] C. Scheier, R. Pfeifer, and Y. Kunyioshi, "Embedded neural networks: exploiting constraints," *Neural Networks*, vol. 11, no. 7-8, pp. 1551–1596, 1998.

[21] S. Nolfi, "Power and limits of reactive agents," *Neurocomputing*, vol. 42, pp. 119–145, 2002.

[22] R. Beer, "The dynamics of active categorical perception in an evolved model agent," *Adaptive Behavior*, vol. 11, pp. 209–243, 2003.

[23] S. Nolfi and D. Marocco, "Active perception: A sensorimotor account of object categorisation," in *Proc. of the 7^{th} Inernational Conference on Simulation of Adaptive Behavior (SAB '02)*, B. Hallam, D. Floreano, J. Hallam, G. Hayes, and J.-A. Meyer, Eds. MIT Press, Cambridge, MA, 2002, pp. 266–271.

[24] T. Buehrmann and E. D. Paolo, "Closing the loop: Evolving a model-free visually-guided robot arm," in *Proceedings of the 9th International Conference on the Simulation and Synthesis of Living Systems*, J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. Watson, Eds. MIT Press, Cambridge, MA, 2004, pp. 63–68.

[25] E. Tuci, V. Trianni, and M. Dorigo, "Feeling the flow of time through sensory-motor coordination," *Connection Science*, vol. 16, no. 4, pp. 301–324, 2004.

[26] O. Gigliotta and S. Nolfi, "On the coupling between agent internal and agent/ environmental dynamics: Development of spatial representations in evolving autonomous

robots," *Adaptive Behavior*, vol. 16, pp. 148–165, 2008.
[27] G. Massera, A. Cangelosi, and S. Nolfi, "Evolution of prehension ability in an anthropomorphic neurorobotic arm," *Front. Neurorobot.*, vol. 1, pp. 1–9, 2007.
[28] R. Beer and J. Gallagher, "Evolving dynamical neural networks for adaptive behavior," *Adaptive Behavior*, vol. 1, no. 1, pp. 91–122, 1992.
[29] D. Goldberg, *Genetic algorithms in search, optimization and machine learning*. Reading, MA: Addison-Wesley, 1989.
[30] S. Strogatz, *Nonlinear Dynamics and Chaos*. Perseus Books Publishing, 2000.
[31] C. Thornton, "Separability is a learner's best friend," in *Proc. of the* 4^{th} *Neural Computation and Psychology Workshop: Connectionist Representations*, J. Bullinaria, D. Glasspool, and G. Houghton, Eds. Springer Verlag, London, UK, 1997, pp. 40–47.
[32] A. Waxman, "Sensor fusion," in *Handbook of brain theory and neural networks*, 2nd ed., M. Arbib, Ed. MIT Press, Cambridge, MA, 2002, pp. 1014–1016.
[33] J. Townsend and J. Busemeyer, "Dynamic representation of decision-making," in *Mind as motion: Explorations in the dynamics of cognition*, R. Port and T. van Gelder, Eds. MIT Press, Cambridge, MA, 1995, pp. 101–120.
[34] M. Platt, "Neural correlates of decisions," *Current Opinion in Neurobiology*, vol. 12, pp. 141–148, 2002.
[35] G. Sandini, G. Metta, and D. Vernon, "The icub cognitive humanoid robot: An open-system research platform for enactive cognition," in *50 Years of Artificial Intelligence*, M. Lungarella, F. Iida, J. Bongard, and R. Pfeifer, Eds. Springer Verlag, Berlin, GE, 2007, pp. 358–369.

Elio Tuci received a Laurea (Master) in Experimental Psychology from "La Sapienza" University, Rome (IT), in 1996, and a PhD in Computer Scinece and Artificial Intelligence from University of Sussex (UK), in 2004. His research interests concern the development of real and simulated embodied agents to look at scientific questions related to the mechanisms and/or the evolutionary origins of individual and social behaviour.

Gianluca Massera is a PhD student at the Plymouth University working under the supervision of Prof. A. Cangelosi and Dr S. Nolfi. His research interests are within the domain of evolutionary robotics, active perception, and sensory-motor coordination in artificial arms.

Stefano Nolfi is research director at the Institute of Cognitive Sciences and Tecnologies of the Italian National Research Council (ISTC-CNR) and head of the Laboratory of Autonomous Robots and Artificial Life. His research activities focus on Embodied Cognition, Adaptive Behaviour, Autonomous Robotics, and Complex Systems. He authored or co-authored more than 130 scientific publications and a book on Evolutionary Robotics published by MIT Press.

The Facilitatory Role of Linguistic Instructions on Developing Manipulation Skills

Gianluca Massera, Elio Tuci, Tomassino Ferrauto, and Stefano Nolfi
Institute of Cognitive Sciences and Technologies (ISTC), ITALY

Abstract–In this paper, we show how a simulated humanoid robot controlled by an artificial neural network can acquire the ability to manipulate spherical objects located over a table by reaching, grasping, and lifting them. The robot controller is developed through an adaptive process in which the free parameters encode the control rules that regulate the fine-grained interaction between the agent and the environment, and the variations of these free parameters are retained or discarded on the basis of their effects at the level of the behavior exhibited by the agent. The robot develops the sensory-motor coordination required to carry out the task in two different conditions; that is, with or without receiving as input a linguistic instruction that specifies the type of behavior to be exhibited during the current phase. The obtained results shown that the linguistic instructions facilitate the development of the required behavioral skills.

I. Introduction

In this paper, we describe a series of experiments in which a simulated iCub robot acquires through an adaptive process the ability to reach, grasp, and lift a spherical object. The robot develops the sensory-motor coordination required to carry out the whole task in two different conditions; that is, with or without receiving as input linguistic instructions that specify the type of behavior that should be exhibited during the current phase. These are binary input vectors associated with elementary behaviors that should be displayed by the robot during the task. The main objective of this study is to investigate whether the use of linguistic instructions facilitates the acquisition of a sequence of complex behaviors. The long term goal of this research is to verify whether the acquisition of elementary skills guided by linguistic instructions provides a scaffolding for more complex behaviors.

Digital Object Identifier 10.1109/MCI.2010.937321

The first theoretical assumption behind this work is that the activity of developing robots displaying complex cognitive and behavioral skills should be carried out by taking into account the empirical findings in psychology and neuroscience which show that there are close links between the mechanisms of action and those of language. As shown in [1], [2], [3], [4], [5] action and language develop in parallel, influence each other, and base themselves on each other. If brought into the world of robotics, the co-development of action and language skills might enable the transfer of properties of action knowledge to linguistic representations, and vice versa, thus enabling the synthesis of robots with complex behavioral and cognitive skills [6], [7].

The second theoretical assumption behind this work is that behavioral and cognitive skills in embodied agents are emergent dynamical properties which have a multi-level and multi-scale organization. Behavioral and cognitive skills arise from a large number of fine-grained[1] interactions occurring among and within the robot body, its control system, and the environment [8]. Handcrafting the mechanisms underpinning these skills may be a hard task. This is due to the inherent difficulty in figuring out from the point of view of an external observer, the detailed characteristics of the agent that, as a result of the interactions between the elementary parts of the agent and of the environment, lead to the exibition of the desired behavior. The synthesis of robots displaying complex behavioral and cognitive skills should instead be obtained through an adaptive process in which the detailed characteristics of the agent are subjected to variation and in which variations are retained or discarded on the basis of their effects at the level of the overall behavior exhibited by the robot situated in the environment [8]. Therefore, the role of the designer should be limited to the specification of the utility function, that determines whether variations should be preserved or discarded, and eventually to the design of the ecological conditions in which the adaptive process takes place [9], [10], [8].

II. Background and Literature Review

The control of arm and hand movements in human and non-human primates and in robots is a fascinating research topic actively investigated within several disciplines including psychology, neuroscience, and robotics. However, the task to model in detail the mechanisms underlying arm and hand movement control in humans and primates and the task of building robots able to display human-like arm/hand movements still represents an extremely challenging goal [11]. Moreover, despite the progress achieved in robotics through the use of traditional control methods [12], the attempt to develop robots with the dexterity and robustness of humans is still a long term goal. These difficulties can be explained by considering the need to take into account the role of several aspects including the morphological characteristics of the arm and of the hand, the bio-mechanics of the musculoskeletal system, the presence of redundant degrees of freedom and limits on the joints, non-linearity (e.g., the fact that small variations in some of the joints might have a strong impact on the hand position), gravity, inertia, collisions, noise, the need to rely on different sensory modalities, visual occlusion, the effects of movements on the next experienced sensory states, the need to coordinate arm and hand movements, the need to adjust actions on the basis of sensory feedback, and the need to handle the effects of the physical interactions between the robot and the environment. The attempt to design robots that develop their skills autonomously through an adaptive process permits, at least in principle, to delegate the solutions to some of these aspects to the adaptive process itself.

The research work described in this paper proposes an approach that takes into account most of the aspects discussed above, although often by introducing severe simplifications. More specifically, the morphological characteristics of the human arm and of the hand are taken into account by using a robot that reproduces approximately the morphological characteristics of a 3.5 year-old in term of size, shape, articulations, degrees of freedom and relative limits [13]. Some of the properties of the musculo-skeletal system have been incorporated into the model by using muscle-like actuators controlled by antagonistic motor neurons. For the sake of simplicity, the segments forming the arm, the palm, and the fingers are simulated as fully rigid bodies. However, the way in which the fingers are controlled, enable a certain level of compliance in the hand. The role of gravity, inertia, collision, and noise are taken into account by accurately simulating the physic laws of motion and the effect of collisions (see Section IV for details of the model).

One of the main characteristics of the model presented in this paper is that the robot controller adjusts its output on the basis of the available sensory feedback directly updating the forces exerted on the joints (see [14] for related approaches). The importance of the sensory feedback loop has been emphasized by other works in the literature. For example in [15] the authors describe an experiment in which a three-fingered robotic arm displays a reliable grasping behavior through a series of routines that keep modifying the relative position of the hand and of the fingers on the basis of the current sensory feedback. The movements tend to optimize a series of properties such as hand-object alignment, contact surface, finger position symmetry, etc.

In this work, the characteristics of the human brain that processes sensory and proprio-sensory information and control the state of the arm/hand actuators are modeled very loosely through the use of dynamical recurrent neural networks. The architecture of the artificial neural network employed is not inspired by the characteristics of the neuroanatomical pathways of the human brain. Also, many of the features of neurons and synapses are not taken into account (see [16], for an example of works that emulate some of the anatomical characteristics of the human brain). The use of artificial neural networks as robot

[1]The granularity refers to the extent to which the robot-environmental system is broken into small parts and to the extent to which the dynamics of the system is divided into short time periods. The term fine-grained interactions thus refer to interactions occurring at a high frequency between small parts.

controller provides several advantages with respect to alternative formalisms, such as robustness, graceful degradation, generalization and the possibility to process sensory-motor information in a way that is quantitative both in state and time. These characteristics also make neural networks particularly suitable to be used with a learning/adaptive process in which a suitable configuration of the free parameters is obtained through a process that operates by accumulating small variations.

Newborn babies display a rough ability to perform reaching, which evolves into effective reaching and grasping skills by 4/5 months, into adult-like reaching and grasping strategies by 9 months, up to precision grasping by 12/18 months [17], [18], [19]. Concerning the role of sensory modalities, the experimental evidence collected on humans indicates that young infants rely heavily on somatosensory and tactile information to carry out reaching and grasping action and they use vision to elicit these behaviors [20]. However, the use of visual information (employed to prepare the grasping behavior or to adjust the position of the hand by taking into account the shape and the orientation of the object) starts to play a role only after 9 months from birth [21]. On the basis of this, we provide our robot with proprioceptive and tactile sensors and with a vision system that only provides information concerning the position of the object but not about its shape and its orientation. Moreover, we do not simulate visual occlusions on the basis of the assumption that the information concerning the position of the object can be inferred in relatively reliable way even when the object is partially or totally occluded by the robot's arm and hand.

In accordance with the empirical evidence indicating that early manipulation skills in infants are acquired through self-learning mechanisms rather than by imitation learning [16], the robot acquires its skills through a trial and error process during which random variations of the free parameters of the robots' neural controller (which are initially assigned randomly) are retained or discarded on the basis of their effect at the level of the overall behavior exhibited by the robot in interaction with the environment. More precisely, the effect of variations is evaluated using a set of utility functions that determine the extent to which the robot manages to reach and grasp a target object with its hand, and the extent to which the robot succeeds in lifting the object over the table. The use of this adaptive algorithm and utility functions leaves the robot free to discover during the adaptive process its own strategy to reach the goals set by the experimenter. This in turn allows the robot to exploit sensory-motor coordination (i.e., the possibility to act in order to later experience useful sensory states) as well as the properties arising from the physical interactions between the robot and the environment. In [22] it is shown how this approach allows the robot to distinguish objects of different shapes by self-selecting useful stimuli through action, and in [23] it is shown how this approach allows for the exploiting of properties arising from the physical interaction between the robot body and the environment for the purpose of manipulating the object.

Finally, in this work we shape the ecological conditions in which the robot has to develop its skills by allowing the robot to access linguistic instructions that indicate the type of behavior that should be currently exhibited by the robot. We do not consider any other form of shaping, such as, for example, the possibility to expose the robot to simplified conditions in some of the trials (in which, for example, the object to be grasped is initially placed within the robot's hand) although we assume that other forms of shaping might favour the developmental process as well.

III. Experimental Setup

Our experiments involve a simulated humanoid robot that is trained to manipulate a spherical object located in different positions over a table in front of the robot by reaching, grasping, and lifting it. More specifically the robot is made up of an anthropomorphic robotic arm with 27 actuated degrees of freedom (DOF) on the arm and hand, 6 tactile sensors distributed over the inner part of the fingers and palm, 17 propriosensors encoding the current angular position of the joints of the arm and of the hand, a simplified vision system that detects the relative position of the object (but not the shape of the object) with respect to the hand and 3 sensory neurons that encode the category of the elementary behaviors that the robot is required to exhibit (i.e., reaching, grasping, or lifting the sphere). The neural controller of the robot is a recurrent neural network trained through an evolutionary algorithm for the ability: (i) to reach an area located above the object, (ii) to wrap the fingers around the object, and (iii) to lift the object over the table. The condition in which the linguistic instructions are provided has been compared with the condition in which the linguistic instructions are not provided. For each condition, the evolutionary process has been repeated 10 times with different random initializations. The robot and the robot/environmental interactions have been simulated by using Newton Game Dynamics (NGD, see: www.newtondynamics.com), a library for accurately simulating rigid body dynamics and collisions. For related approaches, see [23], [22], [24].

In section IV, we describe the structure and the actuators of the arm and hand. In section V, we describe the architecture of the robot controller and the characteristics of the sensors. In section VI, we describe the adaptive process that has been used to train the robot. In section VII, we describe the results obtained, and, finally, in section VIII, we discuss the significance of these results and our plans for the future.

IV. Robot Structure

A. Arm Structure

The arm consists mainly of three elements (the arm, the forearm, and the wrist) connected through articulations placed into the shoulder, the arm, the elbow, the forearm and wrist (see Figure 1).[2]

[2]Details about arm and hand dimensions are available at the supplementary web page http://laral.istc.cnr.it/esm/linguisticExps.

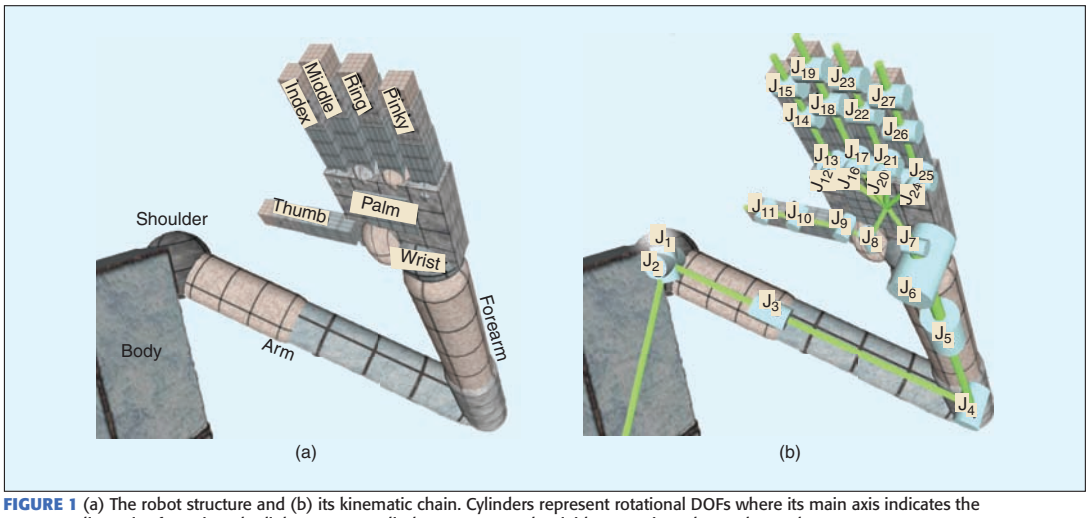


FIGURE 1 (a) The robot structure and (b) its kinematic chain. Cylinders represent rotational DOFs where its main axis indicates the corresponding axis of rotation; the links amongst cylinders represents the rigid connections that make up the arm structure.

The joints J_1, J_2 and J_3 provide *abduction/adduction, extension/flexion* and *supination/pronation* of the arm in the range $[-140^\circ, +100^\circ]$, $[-110^\circ, +90^\circ]$ and $[-110^\circ, +90^\circ]$, respectively. These three degrees of freedom (DOFs) acts like a ball-and-socket joint moving the arm in a way analogous to the human shoulder joint. J_4, located in the elbow, is a hinge joint which provides *extension/flexion* within the $[-170^\circ, +0^\circ]$ range. J_5 twists forearm providing *pronation/supination* of the wrist (and the palm) within $[-100^\circ, +100^\circ]$. J_6 and J_7 provide *flexion/extension* and *abduction/adduction* of the hand within $[-40^\circ, +40^\circ]$ and $[-100^\circ, +100^\circ]$ respectively (see Figure 1).

B. Arm Actuators

The arm joints $(J_1, \ldots, J_7)$ are actuated by two simulated antagonist muscles implemented accordingly to Hill's muscle model [25], [26]. More precisely, the total force exerted by a muscle is the sum of three forces $T_A(\alpha, x) + T_P(x) + T_V(\dot{x})$ which depend on the activity of the corresponding motor neuron (α) on the current elongation of the muscle (x) and on the muscle contraction/elongation speed $(\dot{x})$ which are calculated on the basis of the following equations:

$$
\begin{aligned}
T_A &= \alpha\left(-\frac{A_{sh} T_{\max}(x - R_L)^2}{R_L^2} + T_{\max} \right) \\
A_{sh} &= \frac{R_L^2}{(L_{\max} - R_L)^2} \\
T_P &= T_{\max} \frac{\exp\left\{ K_{sh}\left(\frac{x - R_L}{L_{\max} - R_L} \right) \right\} - 1}{\exp\{K_{sh}\} - 1} \\
T_V &= b \cdot \dot{x},
\end{aligned}
\tag{1}
$$

where $L_{\max}$ and R_L are the maximum and resting lengths of the muscle, $T_{\max}$ is the maximum force that can be generated, K_{sh} is the passive shape factor, and b is the viscosity coefficient.

The active force T_A depends on the activation of muscle α and on the current elongation/compression of the muscle. When the muscle is completely elongated/compressed, the active force is zero regardless of the activation α. At the resting length R_L, the active force reaches its maximum that depends on the activation α. The red curves in Figure 2 show how the active force T_A changes with respect to the elongation of the muscle for some possible values of α. The passive force T_P depends only on the current elongation/compression of the muscle (see the blue curve in Figure 2). T_P tends to elongate the muscle when it is compressed less than R_L and tends to compress the muscle when it is elongated above R_L. T_P differs from a linear spring for its exponential trend that produces a large opposition to muscle elongation and

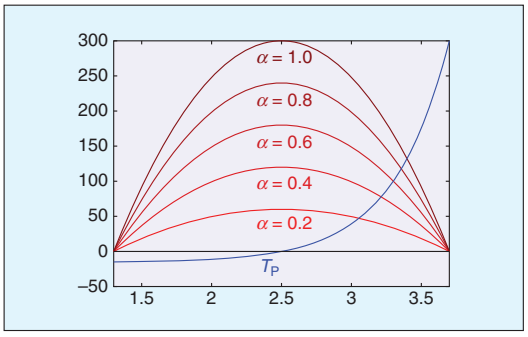


FIGURE 2 An example of the force exerted by a muscle; the graph shows how the force exerted by a muscle varies as a function of the activity of the corresponding motor neuron and of the elongation of the muscle for a joint in which $T_{\max}$ is set to 300 *N*.

little to muscle compression. T_V is the viscosity force. It produces a force proportional to the velocity of the elongation/compression of the muscle.

The parameters of the equation are identical for all 14 muscles controlling the seven DOFs of the arm and have been set to the following values: $K_{sh} = 3.0$, $R_L = 2.5$, $L_{\max} = 3.7$, $b = 0.9$, $A_{sh} = 4.34$ with the exception of parameter $T_{\max}$ which is set to $3000\,N$ for joint J_2, $300\,N$ for joints J_1, J_3, J_4, and J_5, and $200\,N$ for J_6 and J_7.

Muscle elongation is computed by linearly mapping the angular position of the DOF, on which the muscle acts, into the muscle length range. For instance, in the case of the elbow where the limits are $[-170^o, +0^o]$, this range is mapped onto $[+1.3, +3.7]$ for the agonist muscle and $[+3.7, +1.3]$ for the antagonist muscle. Hence, when the elbow is completely extended (angle 0), the agonist muscle is completely elongated (3.7) and the antagonist muscle is completely compressed (1.3), and vice versa when the elbow is flexed.

C. Hand Structure

The hand is attached to the robotic arm just after the wrist (at joint J_7 as shown in Figure 1). One of the most important features of the hand is its compliance. In details, the compliance has been obtained setting a maximum threshold of $300\ N$ to the force exerted by each joint. When an external force acting on a joint exceeds this threshold, either the joint cannot move further, or the joint moves backward due to the external force.

The robotic hand is composed of a palm and 15 phalanges that make up the digits (three for each finger) connected through 20 DOFs, $J_8, \ldots, J_{27}$ (see Figure 1).

Joint J_8 allows the opposition of the thumb with the other fingers and it varies within the range $[-120^\circ, +0^\circ]$, where the lower limit corresponds to thumb-pinky opposition. The knuckle joints J_{12}, J_{16}, J_{20} and J_{24} allow the *abduction/adduction* of the corresponding finger and their ranges are $[0^\circ, +15^\circ]$ for the index, $[-2^\circ, +2^\circ]$ for the middle, $[-10^\circ, +0^\circ]$ for the ring, and $[-15^\circ, +0^\circ]$ for the pinky. All others joints are for the *extension/flexion* of phalanges and vary within $[-90^\circ, +0^\circ]$ where the lower limit corresponds to complete flexion of the phalanx (i.e., the finger closed).

D. Hand Actuators

The joints are not controllable independently of each other, but they are grouped. The same grouping principle used for developing the iCub hand [13] has been used. More precisely, the two distal phalanges of the thumb move together as do the two distal phalanges of the index and the middle fingers. Also, all *extension/flexion* joints of the ring and pinky fingers are linked as are all the joints of *abduction/adduction* of the fingers. Hence, only 9 actuators move all the joints of the hand, one actuator for each of the following group of joints: $\langle J_8 \rangle$, $\langle J_9 \rangle$, $\langle J_{10}, J_{11} \rangle$, $\langle J_{13} \rangle$, $\langle J_{14}, J_{15} \rangle$, $\langle J_{17} \rangle$, $\langle J_{18}, J_{19} \rangle$, $\langle J_{12}, J_{16}, J_{20}, J_{24} \rangle$ and $\langle J_{21}, J_{22}, J_{23}, J_{25}, J_{26}, J_{27} \rangle$. These actuators are simple motors controlled by position.

V. Neural Controller

The architecture of the neural controllers varies slightly depending on the ecological conditions in which the robot develops its skills. In the case of the development supported by linguistic instructions, the robot is controlled by a neural network which includes 29 sensory neurons, 12 internal neurons with recurrent connections and 23 motor neurons. In the case without the support of linguistic instructions, the neural network lacks the sensory neurons dedicated to the linguistic instructions. Thus, it is composed of 26 sensory neurons instead of 29. The sensory neurons are divided into four blocks.

The Arm Sensors encode the current angles of the 7 DOFs located on the arm and on the wrist normalized in the range $[0, 1]$.

The Hand Sensors encode the current angles of hand's joints. However, instead of feeding the network with all joint angles of the hand, the following values are used:

$$\left\langle a(J_8), a(J_9), \frac{a(J_{10}) + a(J_{11})}{2}, a(J_{13}), \frac{a(J_{14}) + a(J_{15})}{2}, a(J_{17}), \frac{a(J_{18}) + a(J_{19})}{2}, a(J_{21}), \frac{a(J_{22}) + a(J_{23})}{2}, a(J_{12}) \right\rangle,$$

where $a(J_i)$ is the angle of the joint J_i normalized in the range $[0, 1]$ with 0 meaning fully extended and 1 fully flexed. This way of representing the hand posture mirrors the way in which the hand joints are actuated (see section IV-D).

The Tactile Sensors encode how many contacts occur on the hand components. The first tactile neuron corresponds to the palm and its activation is set to the number of contacts normalized in the range $[0, 1]$ between the palm and another body (i.e., an object or other parts of the hand). Normalization is performed using a ramp function that saturates to 1 when there are more than 20 contacts. The other five tactile neurons correspond to the fingers and are activated in the same way.

The Target Position Sensors can be seen as the output of a vision system (which has not been simulated) that computes the relative distance in cm of the object with respect to the hand over three orthogonal axes. These values are fed into the networks as they are without any normalization.

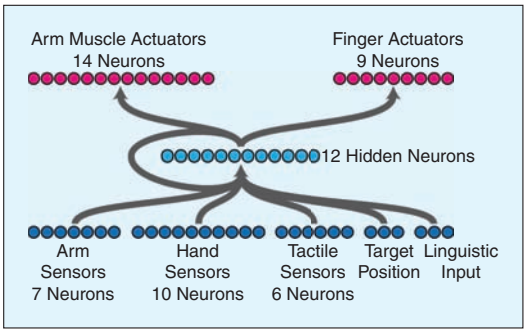


FIGURE 3 The architecture of the neural controllers. The arrows indicated blocks of fully connected neurons

The Linguistic Instruction Sensors is a block of three neurons each of which represents one of the commands *reach, grasp* and *lift*. Specifically, the vector $\langle 50, 0, 0\rangle$ corresponds to the linguistic instruction *"reach the object"*, $\langle 0, 50, 0\rangle$ corresponds to the linguistic instruction *"grasp the object"* and $\langle 0, 0, 50\rangle$ corresponds to the linguistic instruction *"lift the object"*. The way in which the state of these sensors is set is determined by equation 4 explained below.

Note that the state of the Linguistic Instruction and Target Position Sensors varies on a larger interval than the other sensors in order to increase the relative impact of these neurons. Indeed, control experiments in which all sensory neurons were normalized within the $[0, 1]$ interval led to significantly lower performance (result not shown).

The outputs $H_i(t)$ of the Hidden Neurons are calculated on the basis of following equation:

$$\gamma_i(t) = \sigma\left(\sum_{j=1}^{29} w_{ji} I_j(t) + \beta_i\right)$$
$$H_i(t) = \delta_i \cdot \gamma_i(t) + (1 - \delta_i) \cdot \gamma_i(t-1), \qquad (2)$$

where $I_j(t)$ is the output of the *jth* sensory neuron, w_{ji} is the synaptic weight from the *jth* sensory neuron to the *ith* hidden neuron, β_i is the bias of the *ith* hidden neuron, δ_i is the decay-factor of the *ith* hidden neuron, and $\sigma(x)$ is the logistic function with a slope of 0.2.

The output neurons are divided into two blocks, the Arm Muscle Actuators and the Finger Actuators. All outputs of these neurons are calculated in the same way using the following equation:

$$O_i(t) = \sigma\left(\sum_{j=1}^{12} w_{ji} H_j(t)\right), \qquad (3)$$

where $H_j(t)$ is the output of hidden neuron j as described in 2, w_{ji} is the synaptic weight from the *jth* hidden neuron to the *ith* output neuron and $\sigma(x)$ is the logistic function with slope 0.2. With respect to the hidden neurons, the output neurons do not have any bias or decay-factor.

The Arm Muscle Actuators output sets the parameter α used in equation 1 to update the position of the arm as described in section IV-B while the Finger Actuators output sets the desired *extension/flexion* position of the nine hand actuators as described in IV-D. The state of the sensors, the desired state of the actuators, and the internal neurons are updated every 10 ms.

This particular type of neural network architecture has been chosen to minimize the number of assumptions and to reduce, as much as possible, the number of free parameters. Also, this particular sensory system has been chosen in order to study situations in which the visual and tactile sensory channels need to be integrated.

VI. The Adaptive Process

The free parameters of the neural controller (i.e., the connection weights, the biases of internal neurons and the time constant of leaky-integrator neurons) are set using an evolutionary algorithm [27], [28].

The initial population consists of 100 randomly generated genotypes, which encode the free parameters of 100 corresponding neural controllers. In the conditions in which Linguistic Instruction Sensors are employed (hereafter, referred to as Exp. A), the neural controller has 792 free parameters. In the other condition without the Linguistic Instruction Sensors (hereafter, referred to as Exp. B) there are 756 free parameters. Each parameter is encoded into a binary string (i.e., a gene) of 16 bits. In total, a genotype is composed of $792 \cdot 16 = 12672$ bits in Exp. A and $756 \cdot 16 = 12096$ bits in Exp. B. In both experiments, each gene encodes a real value in the range $[-6, +6]$, but for genes encoding the decay-factors δ_i the encoded value is mapped in the range $[0, 1]$.

The 20 best genotypes of each generation are allowed to reproduce by generating five copies each. Four out of five copies are subject to mutations and one copy is not mutated. During mutation, each bit of the genotype has a 1.5% probability to be replaced with a new randomly selected value. The evolutionary process is repeated for 1000 generations.

A. Fitness Function

The agents are rewarded for reaching, grasping and lifting a spherical object of radius 2.5 cm placed on the table in exactly the same way in both Exp. A and Exp. B. Each agent of the population is tested 4 times. Each time the initial position of the arm and the sphere change. Figure 4 shows the four initial positions of the arm and of the sphere superimposed on one another. For each initial arm/object configuration, a random displacement of $\pm 1^o$ is added to each joint of the arm and a random displacement of ± 1.5 cm is added on the x and the y coordinates of the sphere position. Each trial lasts 6 sec corresponding to 600 simulation steps. The sphere can move freely and it can eventually fall off the table. In this case, the trial is stopped prematurely.

The fitness function is made up of three components: *FR* for reaching, *FG* for grasping and *FL* for lifting the object. Each trial is divided in 3 phases in which only a single fitness component is updated. The conditions that define the current phase at each timestep and consequently which component has to be updated are the following:

$$\begin{aligned} r(t) &= 1 - e^{(-0.1 \cdot ds(t))} \\ g(t) &= e^{(-0.2 \cdot \text{grasp}Q(t))} \\ l(t) &= 1 - e^{(-0.3 \cdot \text{contact}(t))} \\ \text{Phase}(t) &= \begin{cases} \text{reach} & r(t) > g(t) < 0.5 \\ \text{grasp} & \text{otherwise} \\ \text{lift} & g(t) > 0.7 \wedge l(t) > 0.6, \end{cases} \end{aligned}$$

where $ds(t)$ is the distance from the center of the palm to a point located 5 cm above the center of the sphere. $\text{grasp}Q(t)$ is the distance between the centroid of the fingertips-palm polygon and the center of the sphere. $\text{contacts}(t)$ is the number of contacts between the fingers and the sphere. The shift between

the three phases is irreversible (i.e. the reach phase is always followed by the reach or grasp phases and the grasp phase is always followed by the grasp or lift phases).

Essentially, the current phase is determined by the values $r(t)$, $g(t)$ and $l(t)$. When $r(t)$ is high (i.e., when the hand is far from the object) the robot should reach the object. When $r(t)$ decreases and $g(t)$ increases (i.e., when the hand approaches the object from above) the robot should grasp the object. Finally, when $l(t)$ increases (i.e., when the number of activated contact sensors are large enough) the robot should lift the object. The rules and the thresholds included in equation 4 have been set manually on the basis of our intuition and have not been adjusted through a trial and error process. In Exp. A, the phases are used to define which linguistic instruction the robot perceives.

The three fitness components are calculated in the following way:

$$FR = \sum_{t \in T_{\text{Reach}}} \left(\frac{0.5}{1 + ds(t)/4} + \frac{0.25}{1 + ds(t)} (\text{fingersOpen}(t) + \text{palmRot}(t)) \right)$$

$$FG = \sum_{t \in T_{\text{Wrap}}} \left(\frac{0.4}{1 + \text{graspQ}(t)} + \frac{0.2}{1 + \text{contacts}(t)/4} \right)$$

$$FL = \sum_{t \in T_{\text{Lift}}} \text{objLifted}(t),$$

where T_{Reach}, T_{Wrap} and T_{Lift} are the time ranges determined by equation 4. $\text{fingersOpen}(t)$ correspond to the average degree of extension of the fingers, where 1 occurs when all fingers are extended and 0 when all fingers are closed. $\text{palmRot}(t)$ is the dot product between the normals of the palm and the table, with 1 referring to the condition in which the palm is parallel to the table and 0 to the condition in which the palm is orthogonal to the table). $\text{objLifted}(t)$ is 1 only if the sphere is not touching the table and it is in contact with the fingers, otherwise it is 0.

The total fitness is calculated at the end of four trials as: $F = \min(500, FR) + \min(720, FW) + \min(1600, FL) + bonus$, where *bonus* adds 300 for each trial where the agent switches from reach phase to grasp phase only, and 600 for each trial where the agent switches from reach to grasp phase and from grasp to lift phase.

During the reach phase the agent is rewarded for approaching a point located 5 cm above the center of the object with the palm parallel to the table and the hand open. Note that the rewards for the hand opening and the rotation of the palm are relevant only when the hand is near the object (due to $0.25/(1 + ds(t))$ factor); in this way the agent is free to rotate the palm when the hand is away from the sphere allowing any reaching trajectory.

During the grasp phase, the centroid of the fingertips-palm polygon can reach the center of the sphere only when the hand wraps the sphere with the fingers, producing a potential power grasp. During the lift phase, the reward is given when the agent effectively moves the sphere upward of the table.

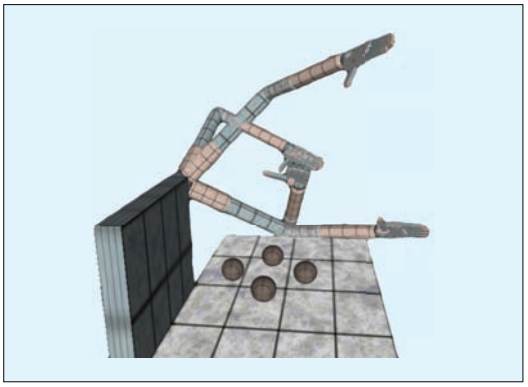


FIGURE 4 Initial positions of the arm and the sphere over imposed; the joints $J_1, \ldots J_4$ are initialized to $\langle -73, -30, -40, -56 \rangle$, $\langle -73, -30, -40, -113 \rangle$, $\langle -6, +30, -10, -56 \rangle$ and $\langle -73, -30, +45, -113 \rangle$; the initial sphere positions are $\langle -18, +10 \rangle$, $\langle -26, +18 \rangle$, $\langle -18, +26 \rangle$ and $\langle -10, +18 \rangle$.

VII. Results

For both Exp. A (with linguistic instructions) and Exp. B (without linguistic instructions), we run 10 evolutionary simulations for 1,000 generations, each using a different random initialization. Looking at the fitness curves of the best agents at each generation of each evolutionary run, we noticed that, for Exp. A, there are three distinctive evolutionary paths (see Figure 5a). The most promising is run 7, in which the last generation's agents have the highest fitness. The curve corresponding to run 2 is representative of a group of seven evolutionary paths which, after a short phase of fitness growth, reach a plateau at $F = 2{,}000$. The curve corresponding to run 9 is representative of a group of two evolutionary paths which are characterized by a long plateau slightly above $F = 1{,}000$. Generally speaking, these curves progressively increase by going through short evolutionary intervals in which the fitness grows quite rapidly followed by a long plateau[3]. For Exp. B, all the runs show a very similar trend, reaching and constantly remaining on a plateau at about $F = 3{,}000$ (see Figure 5b).

Due to the nature of the task and of the fitness function, it is quite hard to infer from these fitness curves what could be the behavior of the agents during each evolutionary phase. However, based on what we know about the task, and by visual inspection of the behavior exhibited by the agents, we found out how the agents behave at different generations of each evolutionary run. In Exp. A, the phases of rapid fitness growth are determined by the *bonus* factor, which substantially rewards those agents that successfully

[3]The fitness curves of the runs not shown are available at the supplementary web page http://laral.istc.cnr.it/esm/linguisticExps.

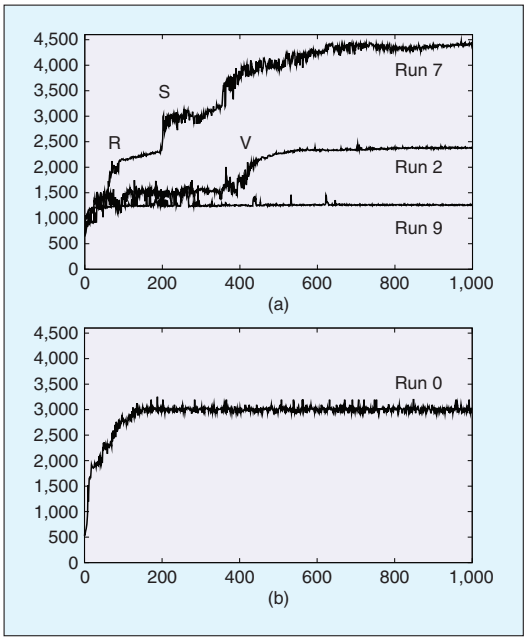


FIGURE 5 Fitness of the best agents at each generation of (a) run 2, run 7, and run 9 of Exp. A, and (b) run 0 of Exp. B.

accomplish single parts of the task. The first fitness jump is due to the *bonus* factor associated to the execution of a successful reaching behavior. This jump corresponds to the phase of fitness growth observed in run 7 in correspondence of label R Figure 5a, and in run 2 in correspondence of label V Figure 5a. The agents generated after these fitness jumps are able to systematically reach the object. Run 9 does not go through the first fitness jump, and the agents of this run lack the ability to systematically carry out a successfull reaching behavior.

The second fitness jump is due to the *bonus* factor associated with the execution of a successful grasping behavior. Only in run 7 is it possible to observe a phase of rapid fitness growth corresponding to a second fitness jump (see label S Figure 5a). The agents generated after this jump are able to successfully carry out reaching and grasping. Note also that, in run 7, the fitness curve keeps on growing until the end of the evolution. This growth is determined by the evolution of the capability to lift the object. Thus, in run 7, the best agents following generation 400 are capable of reaching, grasping, and lifting the object. The constant increment of the fitness is determined by the fact that the agents become progressively more effective in lifting the object. Run 2 does not go through a second fitness jump. The agents of this run lack the ability to systematically carry out a successfully grasping behavior.

In summary, only run 7 has generated agents (i.e., those best agents generated after generation 400) capable of successfully accomplishing reaching, grasping, and lifting.[4] The best agents of run 2, and of the other six runs that show a similar evolutionary trend, are able to systematically reach but not grasp the object and completely lack the ability to lift the object. The best agents of run 9, and of the other run that show a similar evolutionary trend, are not even able to systematically reach the object. In Exp. B, they are able to successfully reach and grasp the object, but not lift it.

A. Robustness and Generalization

In this section, we show the result of a series of post-evaluation tests aimed at establishing the effectiveness and robustness of best agents' behavioral strategies of the four runs show in Figure 5. In these tests, the agents, from generation 900 to generation 1000 of each run, are subjected to a series of trials in which the position of the object as well as the initial position of the arm are systematically varied. For the position of the object, we define a rectangular area (28 cm $\times$ 21 cm) divided in 11 $\times$ 11 cells. The agents are evaluated for reaching, grasping and lifting the object positioned in the center of each cell of the rectangular area. For the initial position of the arm, we use the four initial positions employed during evolution as prototypical cases (see Figure 4). For each prototypical case, we generate 100 slightly different initial positions with the addition of a $\pm 10^{\circ}$ random displacement on joints J_1, J_2, J_3, and J_4. Thus, this test is comprised of 48400 trials, given by 400 initial positions ($4 \cdot 100$) for each cell, repeated for 121 cells corresponding to the different initial positions of the object during the test. In each trial, reaching is considered successful if an agent meets the conditions to switch from the *reach* phase to the *grasp* phase (see equation 4). Grasping is considered successful if an agent meets the conditions to switch from the *grasp* phase to the *lift* phase (see equation 4). Lifting is considered successful if an agent manages to keep the object at more than 1 cm from the table until the end of the trial. In this section, we show the results of a single agent for each run. However, agents belonging to the same run obtained very similar performances. Thus, the reader should consider the results of each agent as representative of all the other agents of the same evolutionary run.

All the graphs in Figure 6 show the relative position of the rectangular area and the cells with respect to the agent/table system. Moreover, each cell of this area is colored in shades of grey, with black indicating 0% success rate, and white indicating 100% success rate. As expected from the previous section, the agent chosen from run 7 Exp. A proved to be the only one capable of successfully accomplishing all the three phases of the task. This agent proved capable of successfully reaching the object placed almost anywhere within the rectangular area. Its grasping and lifting behavior are less robust than the reaching behavior. Indeed, the grasping and lifting performances are quite good everywhere except in

[4]Movies of the behavior and corresponding trajectories are available at the supplementary web page http://laral.istc.cnr.it/esm/linguisticExps.

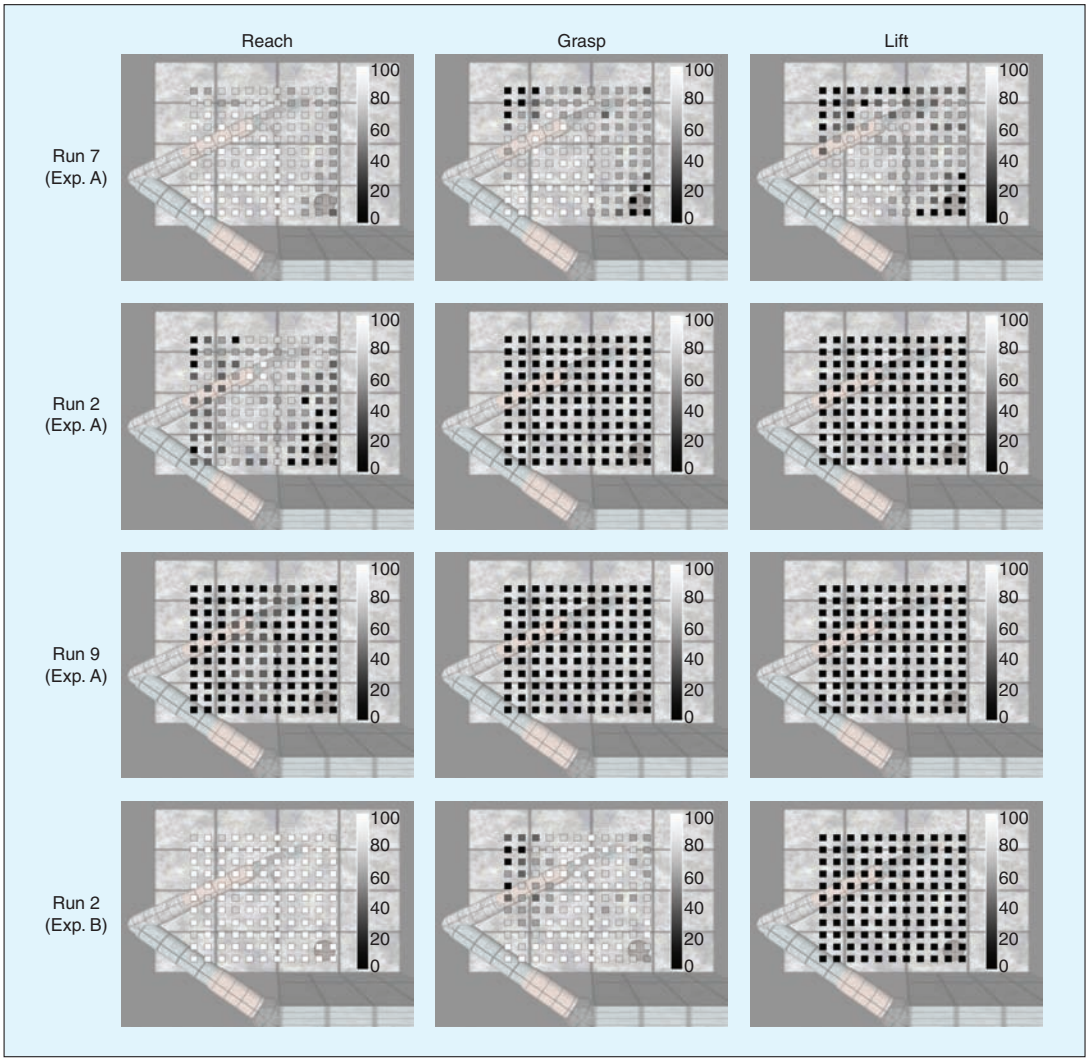


FIGURE 6 Results of post-evaluation tests on the robustness of reaching, grasping and lifting behavior of the best agent at generation 1,000 of run 7, run 2, and run 9 in Exp. A and run 0 in Exp. B. The cells in shades of grey indicate the percentage of successful trials (from 0% success rate in black, to 100% success rate in white), with the object located in the center of each cell.

two small zones located in the top left and bottom right of the rectangular area in which cells are colored black. The agent chosen from run 2 Exp. A proved to be capable of successfully performing reaching behavior for a broad range of object initial positions, and completely unable to perform grasping and lifting behavior. The agent chosen from run 9 Exp. A does not even manage to systematically bring the hand close to the object regardless of the object's initial position. The agent chosen from run 0 Exp. B, proved capable of successfully performing reaching and grasping behavior but not lifting behavior.

VIII. Conclusion

In this paper, we showed how a simulated humanoid robot controlled by an artificial neural network can acquire the ability to manipulate spherical objects located over a table by reaching, grasping and lifting them. The agent is trained through an adaptive process in which the free parameters encode the control rules that regulate the fine-grained interaction between the agent and the environment, and the variations of these free parameters are retained or discarded on the basis of their effects at the level of the behavior exhibited by the agent. This means that the agents develop their skills autonomously in interaction

with the environment. Moreover, this means that the agents are left free to determine the way in which they solve the task within the limits imposed by i) their body/control architecture, ii) the characteristics of the environment, and iii) the constraints imposed by the utility function that rewards the agents for their ability to reach an area located above the object, wrap the fingers around the object, and lift the object. The analysis of the best individuals generated by the adaptive process shows that the agents of a single evolutionary run manage to reach, grasp, and lift the object in an reliable and effective way. Moreover, when tested in new conditions with respect to those experienced during the adaptive process, these agents proved to be capable of generalising their skills with respect to new object positions never experienced before. The comparison of two experimental conditions (i.e., with or without the use of linguistic instructions that specify the behaviors that the agents are required to exhibit during the task) indicates that the agents succeed in solving the entire problem only with the support of linguistic instructions (i.e., in Exp. A). This result confirms the hypothesis that the possibility to access linguistic instructions, representing the category of the behavior that has to be exhibited in the current phase of the task, might be a crucial prerequisite for the development of the corresponding behavioral skills and for the ability to trigger the right behavior at the right time. More specifically, the fact that the best agents of Exp. B succeed in exhibiting the reaching and then the grasping behavior but not the lifting behavior suggests that the linguistic instructions represent a crucial pre-requisite in situations in which the agent has to develop an ability to produce different behaviors in similar sensory-motor circumstances. The reaching to grasping transitions are marked by well differentiated sensory-motor states, which are probably sufficient to induce the agents to stop the reaching phase and to start the grasping phase, even without the support of a linguistic instruction. The grasping to lifting transition is not characterized by well differentiated sensory-motor states. Thus, in Exp. A, it seems to be that the valuable support of the linguistic instruction induces successful agents to move on to the lifting phase.

In future work, we plan to verify whether these agents can be trained to self-generate linguistic instructions and use them to trigger the corresponding behaviors autonomously (i.e., without the need to rely on external instructions). In other words, we would like to verify whether the role played by linguistic instructions can be later internalized in agents' cognitive abilities [29], [30], [31]. Moreover, we plan to port the experiments performed in simulation in hardware by using the iCub robot and the compliant system recently developed [32]. Even though the iCub joints are stiff, the implementation of the muscle model used in this article is still possible. Two 6 axis force sensors placed on the arms and a module developed by the robotcub consortium allow the joints to react as if they were compliant. In this way, it is possible to move the joint applying a torque on its axis and thanks to the opensource aspect of the project, it would be possible to implement muscle actuation directly on the motor control boards.

IX. Acknowledgment

This research work was supported by the ITALK project (EU, ICT, Cognitive Systems and Robotics Integrating Project, grant no 214668). The authors thank their colleagues at LARAL for stimulating discussions and feedback during the preparation of this paper.

References

[1] S. F. Cappa and D. Perani, "The neural correlates of noun and verb processing," *J. Neurolinguistics*, vol. 16, no. 2–3, pp. 183–189, 2003.
[2] A. Glenberg and M. Kaschak, "Grounding language in action," *Psychon. Bull. Rev.*, vol. 9, pp. 558–565, 2002.
[3] O. Hauk, I. Johnsrude, and F. Pulvermuller, "Somatotopic representation of action words in human motor and premotor cortex," *Neuron*, vol. 41, no. 2, pp. 301–307, 2004.
[4] F. Pulvermuller, *The Neuroscience of Language. On Brain Circuits of Words and Serial Order.* Cambridge, U.K.: Cambridge Univ. Press, 2003.
[5] G. Rizzolatti and M. A. Arbib, "Language within our grasp," *Trends Neurosci.*, 1998.
[6] A. Cangelosi, V. Tikhanoff, J. F. Fontanari, and E. Hourdakis, "Integrating language and cognition: A cognitive robotics approach," *IEEE Comput. Intell. Mag.*, vol. 2, no. 3, pp. 65–70, 2007.
[7] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. L. Nehaniv, K. Fischer, J. Tani, G. Sandini, L. Fadiga, B. Wrede, K. Rohlfing, E. Tuci, K. Dautenhahn, J. Saunders, and A. Zeschel, "Integration of action and language knowledge: A roadmap for developmental robotics," *Tech. Rep.*, 2010.
[8] S. Nolfi, "Behaviour as a complex adaptive system: On the role of self-organization in the development of individual and collective behaviour," *Complexus*, vol. 2, no. 3–4, pp. 195–203, 2005.
[9] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, no. 5504, pp. 599–600, 2001.
[10] J. Weng, "Developmental robotics: Theory and experiments," *Int. J. Humanoid Robot.*, vol. 1, no. 2, pp. 199–236, 2004.
[11] S. Schaal, "Arm and hand movement control," in *Handbook of Brain Theory and Neural Networks*, 2nd ed., M. Arbib, Ed. Cambridge, MA: MIT Press, 2002, pp. 110–113.
[12] M. Gienger, M. Toussaint, N. Jetchev, A. Bendig, and C. Goerick, "Optimization of fluent approach and grasp motions," in *Proc. 8th IEEE-RAS Int. Conf. Humanoid Robots*. IEEE Press, 2008, pp. 111–117.
[13] G. Sandini, G. Metta, and D. Vernon, "Robotcub: An open framework for research in embodied cognition," *Int. J. Humanoid Robot.*, 2004.
[14] S. Schaal, J. Peters, J. Nakanishi, and A. Ijspeert, "Learning movement primitives," in *Proc. Int. Symp. Robotics Research (ISRR2003)*, S. verlag, Ed. 2004, pp. 1–10.
[15] J. Felip and A. Morales, "Robust sensor-based grasp primitive for a three-finger robot hand," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2009.
[16] E. Oztop, N. S. Bradley, and M. A. Arbib, "Infant grasp learning: A computational model," *Exp. Brain Res.*, vol. 158, no. 4, pp. 480–503, 2004.
[17] C. von Hofsten, "Eye-hand coordination in the newborn," *Dev. Psychol.*, vol. 18, pp. 450–461, 1982.
[18] C. von Hofsten, "Developmental changes in the organization of prereaching movements," *Dev. Psychol.*, vol. 20, pp. 378–388, 1984.
[19] C. von Hofsten, "Structuring of early reaching movements: a longitudinal study," *J. Mot. Behav.*, vol. 23, pp. 280–292, 1991.
[20] P. Rochat, "Self-perception and action in infancy," *Exp. Brain Res.*, vol. 123, pp. 102–109, 1998.
[21] M. K. McCarty, R. K. Clifton, D. H. Ashmead, P. Lee, and N. Goulet, "How infants use vision for grasping objects," *Child Dev.*, vol. 72, pp. 973–987, 2001.
[22] E. Tuci, G. Massera, and S. Nolfi, "Active categorical perception of object shapes in a simulated anthropomorphic robotic arm," *IEEE Trans. Evol. Comput.*, to be published.
[23] G. Massera, A. Cangelosi, and S. Nolfi, "Evolution of prehension ability in an anthropomorphic neurorobotic arm," *Front. Neurorobot.*, vol. 1, pp. 1–9, 2007.
[24] T. Buehrmann and E. A. Di Paolo, "Closing the loop: Evolving a model-free visually-guided robot arm," in *Proc. 9th Int. Conf. Simulation and Synthesis of Living Systems*, J. Pollack, M. Bedau, P. Husbands, T. Ikegami, and R. Watson, Eds. Cambridge, MA: MIT Press, 2004, pp. 63–68.
[25] T. G. Sandercock, D. C. Lin, and W. Z. Rymer, "Muscle models," in *Handbook of Brain Theory and Neural Networks*, 2nd ed., M. Arbib, Ed. Cambridge, MA: MIT Press, 2002, pp. 711–715.
[26] R. Shadmehr and S. P. Wise, *The Computational Neurobiology of Reaching and Pointing: A Foundation for Motor Learning.* Cambridge, MA: MIT Press, 2005.
[27] S. Nolfi and D. Floreano, *Evolutionary Robotics: The Biology, Intelligence, and Technology of Self-Organizing Machines.* Cambridge, MA: MIT Press, 2000.
[28] X. Yao and M. M. Islam, "Evolving artificial neural network ensembles," *IEEE Comput. Intell. Mag.*, vol. 3, no. 1, pp. 31–42, 2008.
[29] L. S. Vygotsky, *Thought and Language.* Cambridge, MA: MIT Press, 1962.
[30] L. S. Vygotsky, *Mind in Society.* Cambridge, MA: Harvard Univ. Press, 1978.
[31] M. Mirolli and D. Parisi. (2009). *Towards a vygotskyan cognitive robotics: The role of language as a cognitive tool. New Ideas Psychol.* [Online]. Available: http://www.sciencedirect.com/science/article/B6VD4-4X00P73-1/2/5eb2e93d' 3fc615eea3ec0f637af6fc89
[32] V. Mohan, J. Zenzeri, P. Morasso, and G. Metta, "Equilibrium point hypothesis revisited: Advances in the computational framework of passive motion paradigm," pp. 1–3.