Using Motor Babbling and Hebb Rules for Modeling the Development of Reaching with Obstacles and Grasping

Daniele Caligiore and Tomassino Ferrauto and Domenico Parisi and
Neri Accornero and Marco Capozza and Gianluca Baldassarre

***Abstract*— An influential hypothesis of developmental psychology states that, in the first months of their life, infants perform exploratory/random movements ("motor babbling") in order to create associations between such movements and the resulting perceived effects. These associations are later used as building blocks to tackle more complex sensorimotor behaviours. Due to its underlying simplicity, motor babbling might be a learning strategy widely used in the early phases of child development. Various models of this process have been proposed that focus on the acquisition of *reaching* skills based on the *synchronous* association between the positions of the seen hand (or grasped object) and the proprioception of the postures that cause them. This research tries to understand, on a computational basis, if the principles underlying motor babbling can be extended to the acquisition of behaviours more complex than reaching, such as the execution of non-linear movement trajectories for avoiding obstacles or the acquisition of movements directed to grasp objects. These behaviours are challenging for motor babbling as they involve the execution of movements, or sequences of movements, in time, and so they cannot be learned on the basis of simple synchronous associations between their neural representations and perceptive neural representations. The paper aims to show that infants might still use motor babbling for the development of these behaviours by overcoming its time-limits on the basis of complementary mechanisms such as Pattern Generators and innate reflexes. The computational viability of this hypothesis is demonstrated by testing the proposed models with a 3D simulated dynamic eye-arm-hand robot working on a plane.**

***Index Terms*— Circular reactions, enclosure reflex, neural networks, Hebb learning, Pattern Generators**

I. INTRODUCTION

One influential idea proposed by Piaget within the developmental psychology studies of sensorimotor development is the "primary circular-reaction hypothesis" [1]. According to this hypothesis, in the first few months of life infants repeatedly perform exploratory movements which are "centred on themselves". This means that behaviour is mainly directed to produce effects on own body rather than in the external environment. On the basis of this hypothesis, both the experimental literature [2] and the literature on computational biomimetic models (see below) have proposed that "motor babbling" inherent to primary circular-reactions (e.g., the performance of random hand movements in front of the eyes) has the function of enhancing the formation of associations between efferent motor patterns and re-afferent perceptive/proprioceptive patterns. These associations are important prerequisites for the later development of goal-directed behaviours (e.g., grasping objects in the world).

This work focusses on primary circular-reaction hypothesis and motor babbling because they might play a central role in the development of humans due to the effectiveness and simplicity of their underlying principles. With this respect, motor babbling requires: (a) to generate *unstructured experience* (e.g. random movements and behaviours); (b) to form *associations* (e.g., via Hebb learning rules) between performed actions and the perceived consequences of them; (c) to re-activate the internal representations of the "consequences" of actions so that they become "desired consequences", that is *goals*; (d) to "invert" the association "action→effect" with respect its acquisition time order in that the re-activation of the representation of the (desired) consequences has to trigger the execution of the related action. From a computational and neural point of view, the main "difficulties" that motor babbling poses are related to the implementation of mechanisms (c) and (d). As we shall see here, however, these difficulties are not too severe as, for example, the activation of *part* of the representation of the consequences by the external environment (e.g. the vision of an object previously grasped) can reactivate the consequences; moreover, the "action→effect" relation can be "inverted" during acquisition by directly learning "effect→action" neural associations (note that all these issues are tackled in detail within the psychological literature on the "Ideomotor Principle", a theory with interesting complementarities with respect to the circular-reaction hypothesis [3]).

In the last twenty years, several neural-network systems have been proposed with the aim of modeling motor babbling. For example, the seminal paper [4] showed how the execution of random movements can allow a neural model controlling an eye/arm system to form associations, via a supervised learning rule, between a perceived object-in-hand position and the corresponding arm postures; the model later uses this knowledge to perform reaching tasks. In [5] a neural model controlling a robotic arm moving on

This research was supported by the EU Projects *ICEA*, contract no. FP6-IST-027819-IP, and *MindRACES*, contract no. FP6-511931-STREP, and by Dipartimento di Psicologia, Universita' di Bologna, Progetto PRIN 2006 MODELLO B - prot. 2006119711_002

D. Caligiore, T. Ferrauto, D. Parisi and G. Baldassarre are with Laboratory of Autonomous Robotics and Artificial Life, Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche (LARAL-ISTC-CNR), Via San Martino della Battaglia 44, I-00185 Roma, Italy, {daniele.caligiore, tomassino.ferrauto, domenico.parisi, gianluca.baldassarre}@istc.cnr.it

N. Accornero and M. Capozza are with Dipartimento di Scienze Neurologiche, Universita' "La Sapienza", Viale dell'Universita' 30, I-00185 Roma, Italy, neri.accornero@uniroma1.it, marco_capozza@tin.it

a plane uses motor babbling to train, with an error-back propagation algorithm, a forward model later used to train an inverse model capable of performing reaching actions. Other neural-network models have been proposed in the last years to further specify the detail functioning of the circular-reaction hypothesis at a neural level (e.g., [6]), or to exploit associations formed with motor babbling and supervised algorithms to control complex robotic plants (e.g., [7]).

In [8] it is shown that if some biological constraints (equilibrium-point muscle mechanisms, leaky-neurons, population codes, and Hebb learning rules) are fulfilled when building models based on motor babbling, interesting phenomena emerge such as connection weights with a high-contrast Mexican hat distribution, stabilization effects on arm's movements, and movements with bell-shaped speed profiles similar to human movements.

In most of these works motor babbling was used to tackle relatively simple reaching tasks. The reason of this is likely that motor babbling is suitable for acquiring behaviours on the basis of the formation of associations between sensorial and behavioural events which are *synchronous* and *punctual*, for example between the current arm posture proprioception and the current hand/object sight. On the contrary, motor babbling is not suitable for acquiring behaviours that require associating events which take place asynchronously or last for long. The central hypothesis of this research is that organisms exploit the aforementioned advantages of motor babbling even for acquiring complex motor behaviours as they can overcome its time-limits by using suitable complementary mechanisms in synergy with it.

The paper shows the computational viability of this hypothesis by proposing two models that use motor babbling for acquiring two relatively-complex behaviours: *reaching with obstacles* and *object grasping*. Both behaviors pose the aforementioned time-problems. Reaching with obstacles involves the production of *non-linear trajectories*, that is the fine regulation of movement in time. Object grasping requires the production of a sequence of building-block movements in time (e.g., "bringing the hand in proximity of the object" and "closing the hand around the object"). The specific hypotheses investigated here are that: (a) in reaching with obstacles, the arm's non-linear trajectories can be generated on the basis of *Pattern Generators (PGs)*: motor babbling forms associations between the PGs' parameters and visual percepts; (b) in object grasping, motor babbling allows learning specific sensorimotor building-blocks later "chained" by an *innate reflex* that triggers the execution of a building block on the basis of the consequences produced by the execution of the previous building block.

In the following, Sect. II presents the experimental setup used to test the two models. Sect. III explains the model of infants' development of reaching with obstacles, and the related tests. Sect. IV explains the neural-network architectures used to model grasping in young infants and then illustrates the results of the tests of the model. Finally, Sect. V highlights the strengths and weaknesses of the models and draws conclusions.

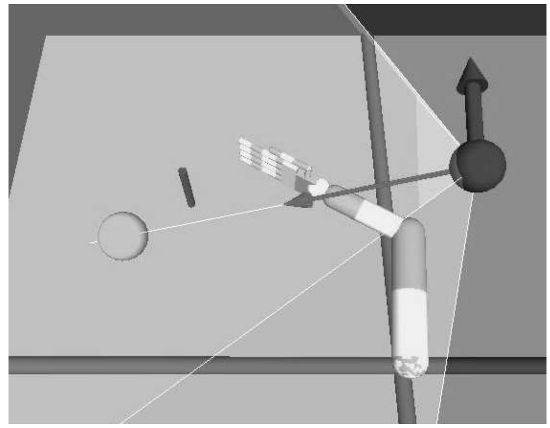

Fig. 1. The robotic setup used to test the models. The picture focuses on the north-west quadrant of the four working plane quadrants. The sphere with the arrows indicates the position of the eye (thin arrow: current gaze; thick arrow: camera up-vector). The sphere in front of the arm is the object that the robot tries to reach (first model) or grasp (second model: in this case it is the "big object"). The cylinder represents the obstacle to be avoided (first model).

II. THE SIMULATED ROBOTIC SETUP

We now analyse in detail the three components forming the robot used to test the two models: a visual system, a dynamic 3D arm-hand system, and a muscle system used to generate joints' torques.

A. Visual System

The visual system is composed of an "eye" (an RGB camera with a resolution of 630×630 pixels covering a 120° pan angle field and a 120° tilt angle field) mounted 25*cm* above the shoulder and "leaning forward" 10*cm*. The eye movement is controlled by a reflex that tends to foveate red objects (the objects to be reached or grasped). In particular, the pan and tilt angles $\mathbf{pt}$ are changed with a vector of values $\Delta\mathbf{pt}$:

$$\Delta\mathbf{pt} = 120^\circ \frac{(\mathbf{E}-\mathbf{O})}{630} \qquad (1)$$

where $\mathbf{E}$ is a vector with two elements equal to 315 (i.e. the center of the image in pixels), $\mathbf{O}$ is the vector whose two elements are equal to the weighted average of the x-y components of the positions of the red pixels in the retina, and 630 and 120° are respectively the image size (in pixels) and the camera aperture.

The assumption for which the eye always foveates the target is very important for the functioning of the models. For example it allows the two models to use the eye's proprioception (i.e. the current camera's pan and tilt angles normalised in $(0,1)$) to identify the position of the target object relative to the robot's body ("where"). Moreover, the assumption implies that the object is always centred in the image. This facilitates the processing of the grasping model that uses the information coming from the whole eye's image in order to identify the shape of the object ("what") so as to perform suitable movements with the hand. Note that from a computational point of view these advantages have been studied within the *active vision* literature [9]. The assumption is also in line with the current neuroscientific literature suggesting that primates tend to foveate the target objects with which they interact and that their brain exploits gaze centred reference frames as much as possible for sensorimotor coordination (see [10] for a review).

B. Muscle Model

The the robotic arm-hand is moved on the basis of joint torques generated by a simulated muscle system that receives as input the desired arm-hand postures (joints' angles) from the "inverse kinematic" module of the reaching model (see Sect. III-A) or directly from the neural output maps in the grasping model (see Sect. IV-A).

Muscles are often modeled as either a linear or a non linear spring element on the basis of the *equilibrium point hypothesis* for which they implement local-feedback mechanisms that lead the controlled limb to progressively approach a stable desired angle [11], [12]. The present work, similarly to what is done in [13], models single muscles as simple *Proportional Derivative controllers* (PD, [14]; see also [15]). As shown in [13] muscle models as simple as the one used here allow reproducing infants' movements quite accurately.

The equation of the PD muscle model is as follows:

$$\mathbf{T} = \mathbf{K}_P \tilde{\mathbf{q}} - \mathbf{K}_D \dot{\mathbf{q}} \tag{2}$$

where $\mathbf{T}$ is the vector of torques applied to the joints, $\mathbf{K}_P$ is diagonal matrix with elements equal to 50 (reaching model) or 200 (grasping model), $\tilde{\mathbf{q}}$ is the difference vector between the desired angular joint position and the current angular joint position, $\mathbf{K}_D$ is a diagonal matrix with elements equal to 40 (reaching model) or 200 (grasping model), and $\dot{\mathbf{q}}$ is the vector of current angular speed at joints.

A further assumption is that the PD's action is integrated by *gravity-compensation mechanisms* [14]. This is implemented in a simple fashion by ignoring the effects of gravity on the arm and the hand in the dynamic simulation of the arm-hand (see Sect. II-C).

C. Dynamic Arm and Hand

The robot's arm and hand were simulated with a physical engine software simulator developed at ISTC-CNR[1].

The arm and hand have the same parameters of the *iCub* robot[2] that is respectively 2 segments/7DOF and 21 segments/19DOF. In particular the arm is composed of an upper arm (15*cm*; 3DOF with respect to the shoulder), a forearm (13*cm*; 1DOF with respect to the upper arm) and a wrist (3DOF). The hand is composed of a thumb and four same size fingers. The thumb is formed by a first segment pivoting inside the hand's palm (2*cm*; 1DOF to allows the thumb's contraposition to the other fingers) and three segments (2*cm*, 1.7*cm* and 1.4*cm*; 1DOF each for flexion/extention). The same size fingers are formed by a first segment pivoting inside the hand's palm (1*cm*; 1DOF to allow fingers' expansion) and three segments (2*cm*, 1.7*cm* and 1.4*cm*; 1DOF each for flexion/extention).

In the simulations the arm moves on the plane and two of the 3DOF of the shoulder and all the 3DOF of the wrist are kept still (so only one DOF of the shoulder and one DOF of the elbow are controlled). As argued in [13], this assumption is in line with what is observed in experiments: infants tend to accomplish reaching movements by using mainly two degrees of freedom. The explanation is that this strategy reduces the complexity of movements and accelerates learning [16]. Note that here the assumption also simplifies the interpretation of the results and makes it possible to compute the inverse kinematics of the arm as needed by the reaching model (see Sect. III-A).

Another assumption regards the control of the hand. In the reaching model it is always kept fully stretched and reaching is performed with the tips of the four equal-size fingers. In the grasping model, the controller acts only on the thumb (its 4DOF angles are progressively "closed") and on a "virtual finger" which is supposed to correspond to the four same sizefingers (whose joint angles are again progressively "closed") acting as a whole "functional unit" [17]. In particular, a flexion of the virtual finger's DOFs causes a proportional flexion of the real fingers's DOFs so that they progressively "envelop" the target (spherical) objects. It is likely that primates, especially when young, use similar functional units to ease the control of the high number of DOFs of the hand [17].

III. Reaching with Obstacles

The robotics literature has proposed three methods to model reaching with obstacles [18]. The first method, called *global method*, performs *off-line motion planning* before triggering robots' motion. For example the algorithm proposed in [19] builds a *potential function* around the obstacle and uses it to plan the motion of a manipulator, in terms of positions with minimum energy, before performing the movement. The second method, called *local method*, checks *on-line* (i.e. during motion) potential collisions with obstacles and triggers an adjusting strategy when necessary. Several authors have proposed a number of solutions within this framework such as the *potential field approach* [20], [21], [22]. This assumes that each object in the environment exerts a repulsive force on the manipulator's end-point whereas the target exerts an attractive force on it. The resulting force is then computed at each step and used to determine the direction of movement. The third method is *imitation learning*. During a training phase several sample movements are generated by the experimenter by directly or indirectly acting on the robot's manipulator [23], [24], [25]. The robots store postures and forces of limbs during movements and associate them with obstacles' and targets' positions. In a later stage this information allows the robots to perform similar movements autonomously. The biological plausibility of these approaches is still debated. The main reason is probably that some fundamental aspects of motor control

[1]This software is based on the open-source software OPAL (Open Physics Abstraction Layer) updated to interface the NEWTON physical engines – originally OPAL could only be used to interface the open-source ODE physical engine.

[2]http://www.robotcub.org

are still unclear. For example, it is not clear if the central nervous system controls the kinematics of movements (e.g. on the basis of equilibrium points, [11]) or the dynamics of them [26], or both [6] (see [10] for a review).

This work proposes a new model on how young infants might use motor babbling to learn reaching with obstacles. The idea is that infants use motor babbling to create associations between the target and obstacle sight on one side and the *parameters* of *Pattern Generators (PGs)* on the other, and that the PGs are then capable of generating non-linear hand trajectories. The idea behind this is that, while reducing the space of dynamic movement control to the setting of few parameters, PGs allow implementing a repertoire of *motor primitives* that, when suitably combined, can still generate a wide range of different behaviours [27].

A. The Architecture and Functioning of the Model

Fig. 3 shows the architecture of the model. The model's core is formed by an input and an output 2D maps of 21×21 neurons each. As we shall see, these maps use *population codes* to encode input and output signals. This assumption is based on the *population code hypothesis* for which often organisms' brains use large populations of neurons with broad tuning curves to encode sensory and motor variables on the basis of their spatial location [8], [28], [29].

The input map encodes the extrinsic obstacle position in eye-centered coordinates with the eye assumed to always fixate the target position before and during the reaching movements [10], [30]. Note that, as the obstacle is always in the central position between the starting hand position and the target position, it is a perfect predictor of the seen hand position; as a consequence the latter information is redundant and not necessary to the system (as confirmed by pilot experiments: data not reported).

The input map encodes the input signal on the basis of a Gaussian function based on assumption of population code hypothesis for which stimuli with values close to the *preferred values* of neurons (assigned to them on the basis of their spatial locations, see [8]) cause a high activation of them, whereas farther stimuli cause lower activations:

$$g_i = f[\mathbf{x}_i, \mathbf{x}] = \exp\left[-\frac{|\mathbf{x}_i - \mathbf{x}|}{2\sigma^2}\right] \tag{3}$$

where g_i is the activity of neuron i, $\mathbf{x}_i$ is the two-element vector of the neuron's preferred values, $\mathbf{x}$ is the two-element vector of the stimuli, and σ is the standard deviation of the Gaussian set to 0.7.

The output map encodes two parameters, b_1 and b_2, used by a *Pattern Generator* to generate the hand trajectory. During learning (see below), random arm's movements (motor babbling) are generated by randomly drawing b_1 and b_2 values. These values are used to activate the output map on the basis of Eq. 3 analogously to what is done for the input map. During tests, the activation of the input map by the sight of the obstacle (with the eye centred on the target) activates the output map which on its turn produces the b_1

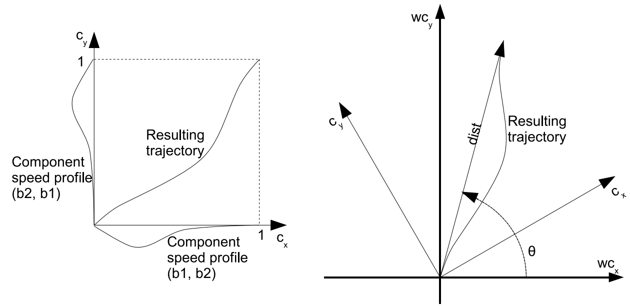

Fig. 2. Sketch of how the TBGs generate a hand non-linear trajectory. **Left**: example of trajectory generated by Eq. 5. **Right**: same trajectory after a rotation and scaling (see Eq. 6).

and b_2 values. These are "read out" from the output map as follows [28]:

$$\mathbf{b} = \frac{\sum_j \mathbf{y_j} \cdot a_j}{\sum_j a_j} \tag{4}$$

where $\mathbf{b} = (b_1, b_2)$, $\mathbf{y_j}$ is the two-element vector of the neuron's preferred values, and $a_j = max[tanh[\sum_i w_{ji} \cdot g_j], 0]$ where w_{ji} are the two maps' connection weights.

The vectors $\mathbf{b} = (b_1, b_2)$ and $\mathbf{b}' = (b_2, b_1)$ are used as parameters of two PGs used to control respectively the orthogonal c_x and c_y components of the trajectory which starts from $c_x = c_y = 0$ and terminates at $c_x = c_y = 1$ (see Fig. 2, left). During the execution of the movement, the PGs generate the coordinates c_x and c_y of this trajectory as follows:

$$\begin{aligned} \dot{c_x} &= \gamma {c_x}^{b_1} (1 - c_x)^{b_2} \\ \dot{c_y} &= \gamma {c_y}^{b_2} (1 - c_y)^{b_1} \end{aligned} \tag{5}$$

In [26], two-unit recurrent-neural-network PGs like those of Eq. 5 were used to generate "1D" time-series signals with controllable finite durations and bell-shaped velocity profiles as those observed in humans. Here the parameters of the two PGs of Eq. 5 are used to regulate the duration and velocity profile of the 2D hand trajectory. In Eq. 5 c_x and c_y are the current value of the trajectory coordinates, $\dot{c_x}$ and $\dot{c_y}$ are their variations, γ is a parameter regulating the overall duration of the movement and is set on the basis of $\mathbf{b}$ and $\mathbf{b}'$ as in [26]. The parameters $\mathbf{b}$ and $\mathbf{b}'$ allow regulating the symmetry/asymmetry of such shape: if $b_1 = b_2$ the shape is symmetric otherwise it is asymmetric. The fact that $\mathbf{b}$ and $\mathbf{b}'$ are specular implies that the neural network can produce movements with a left or right curvature (with a different curvature factor) by suitably setting b_1 and b_2 (Fig. 2, right).

Notice that, as indicated in Fig. 2, once the (c_x, c_y) points of the trajectory are obtained their actual positioning $\mathbf{wc} = (wc_x, wc_y)$ in the workspace is computed using the following (hardwired) rotation and scaling operations:

$$\mathbf{wc} = dist \mathbf{R} (c_x, c_y)^T \tag{6}$$

where *dist* is the distance from the hand position to the target position and $\mathbf{R}$ is the rotation matrix of the trajectory direction (see angle θ in Fig. 2, right). The trajectory orientation $\mathbf{R}$ and distance to cover *dist* are assumed to be directly obtained from the initial hand position (known through the

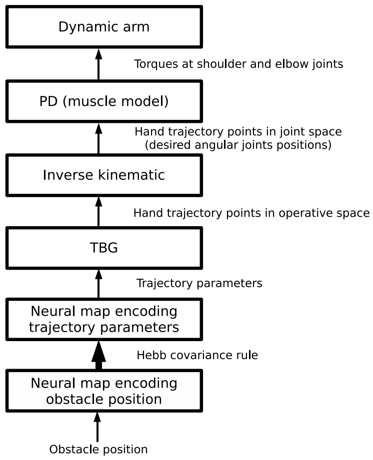

Fig. 3. The architecture of the system. Plain arrows refer to information flows whereas bold arrows represents all-to-all connection weights trained on the basis of a Hebb rule.

arm's proprioception) and from the target position (known through the eye proprioception) [31].

A hardwired inverse kinematic transformation is then used to convert the hand desired trajectory points **wc** (in Euclidean space) in posture points (in joint space; see Fig. 3). The posture points are then used as desired joint angles for the PD muscle models to obtain the joints' torques.

The model learns as follows. Random b_1 values (for simplicity $b_2 = 1 - b_1$) are generated to perform random arm trajectories (motor babbling). Importantly, learning uses only the random parameters that produce a "legal" trajectory, that is a trajectory that (a) does not lead the arm to collide the obstacle with any part of the arm and (b) does not lead the arm violate any angle range of the joints. Learning lasts 20.000 legal trajectories (on average, about 1 out of 5 trajectories is legal; futher 1.000 legal trajectories are generated to be used in a later generalisation test, see Sect. III-B). For each trajectory, the input map is activated on the basis of the eye-centred sight of the obstacle (Eq. 3), the random **b** values that generated the movement are used to activate the output map (Eq. 3), and the all-to-all weights between these two maps are updated according to a *covariance Hebb rule* [8] [32]:

$$\Delta w_{ji} = \eta (a_j - \overline{a}_j)(g_i - \overline{g}_i)(w_{max} - |w_{ji}|) \qquad (7)$$

where η is the learning rate set to 1, w_{max} is set to 0.2 and is a parameter that keeps the weights within the range of $[-0.2, 0.2]$, a_j is the activation of the post-synaptic neuron j, g_i is the activation of the pre-synaptic neuron i, $\overline{a}_j$ and $\overline{g}_i$ are moving decaying averages of the neurons' activations, calculated as $\overline{a} = \xi \overline{a} + (1 - \xi) a$ with ξ set to 0.2. This rule strengthens the connections between each couple of neurons that have both an activation above or below their own average activation, and weakens their connections in the other cases.

B. Results

After training, the model develops a good capacity to produced curved trajectories in order to reach the target while avoiding the obstacle. Fig. 4 shows some trajectories the system performs while reaching a target. As desired, the system not only produces trajectories with a curvature suitable for avoiding the obstacle with the tip of the hand, but it also learns to curve the trajectory so as to avoid that any other part of the arm collides it.

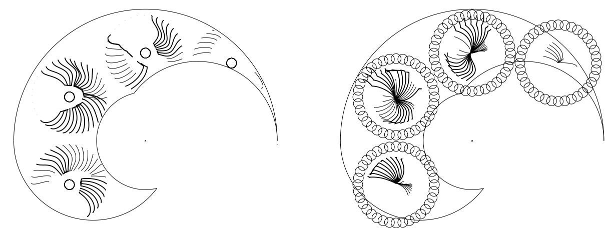

Fig. 4. **Left**: trajectories of the arm during reaching tests after learning, where an obstacle (not reported for clarity) is set between four different targets (circles) and 36 hand initial positions around each of them. Bold trajectories correspond to successful reaching movements, dotted trajectories to failures (i.e. obstacle collisions or joint-range violations). **Right**: the same test repeated with switched obstacle and initial hand positions.

In order to have a quantitative measure of the accuracy of skills, the model was tested with 1.000 couples of target and hand initial positions not used during training (generalisation test). The results are reported in Table I. This Table shows that the system has a high number of successes (67%) in comparison to the number of collisions (23%) or violations of the joints' angle range (10%).

TABLE I

RESULTS OF 1,000 REACHING TESTS WITH OBSTACLES.

Outcome	Frequency	Percent
Successes	670	67%
Collisions	228	23%
Violations	102	10%

The analysis of the distribution of weights from each neuron of the input map (encoding the obstacle position) to the output map (parameters' of trajectory curvature) shows that two distinct and symmetrical patterns emerge during learning for the various obstacle positions and hence for the various initial hand positions (see Fig. 5): one pattern causes the hand trajectory to curve left and the other to curve right. The neural network selects one of the two types of trajectories in correspondence to two compact subsets of possible initial hand positions, as indicated clearly in Fig. 4.

IV. OBJECT GRASPING

When performing actions directed to grasp objects, adults adjust the distance between the thumb and other fingers according to the perceived orientation and size of the target during the hand transport. With this respect, in [33] it was shown that primates' pre-motor cortex encodes a sophisticated repertoire of different types of grasping actions. However, before nine months infants lack the anticipatory movements seen in adults [34] and adjust their grip only *after* touching the target objects. Indeed, a fine fractionated control

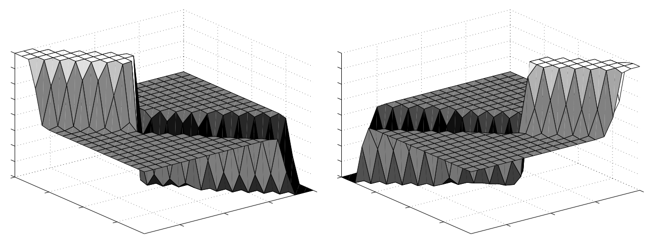

Fig. 5. The reaching model develops weight from each neuron of the input map to neurons of the output map belonging to two different patterns. **Left**: example of weights from one input neuron causing the trajectory to curve left. **Right**: example of weights causing the trajectory to curve right.

of the fingers is not possible at this age as it involves cortico-motoneuronal systems which are not yet fully developed by the age of acquisition of voluntary grasping [35]. At this stage of development, when infants contact objects they will occasionally try to grasp them. This behaviour is supported by the *enclosure reflex* for which grasping movements are triggered when the hand contacts objects. This reflex operates until infants are about six months old. This overall behavioural pattern likely "scaffolds" the formation of more stable grasping behaviours that will take a few more weeks to fully develop [36]. Infants younger than nine months are physically able to vary the grip size, as indicated by the fact that they can spread their fingers apart once they have touched a large object [37]. Likely, these types of adjustments allow the formation of associations between the perceived size of objects and the corresponding hand postures that later will support a full development of the anticipatory grasping patterns observed in adults.

This section proposes a model of the development of these processes based on motor babbling. The simulated development is composed of these phases: (a) when the hand touches the object with the palm, the enclosure reflex causes the closure of the hand with constant torques; (b) the systems moves the arm randomly (motor babbling) with the object in hand, and forms two types of associations: (b1) between the *locations of objects* in space (eye posture) and the corresponding *arm postures*; (b2) between the *foveal perception of objects* held in hand and the corresponding *hand postures*: this mimics the development of the different types of grasping (e.g., power grip, precision grip, etc., see [33]); (c) the sight of a target object re-activates the arm posture corresponding to it, and hence a reaching movement, while the hand's contact with the objects triggers, again via the enclosure reflex, the re-activation the hands's posture corresponding to the perceived object: this mimics the development of the different types of grips from the initial enclosure reflex; (d) suitable more sophisticated processes, such as learning by trial-and-error, support the further development of the different grips and form "chunked" reaching-opening-grasping action sequences *on the basis of their success* (this phase is not modeled here).

A. The Architecture and Functioning of the Model

Fig. 6 shows the components of the model. The neural components of the model are formed by four 2D maps of

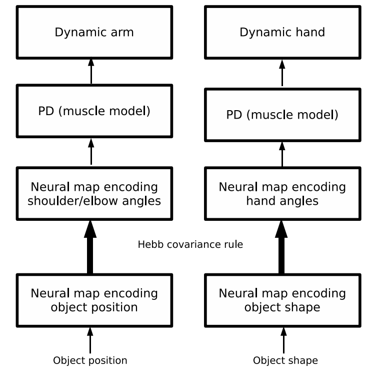

Fig. 6. The architecture of the grasping model, formed by four neural maps for reaching and grasping, PD muscle models, and a dynamic arm-hand model. Plain arrows refer to information flows whereas bold arrows represent connection weights trained on the basis of an Hebb rule.

21×21 neurons each. These maps use *population codes* to encode input and output signals. In particular, the two input maps respectively encode, on the basis of the Gaussian function of Eq. 3, the following information: (a) the object position signaled by the eye's pan and tilt angles; (b) the shape of the foveated object obtained through a *Sobel filter* [38] applied to the *fovea* of the visual image (the central 63×63 pixel central portion of the image). The Sobel filter is a very simple image filter that can be used to mimic edge detection performed by primary visual cortex simple cells [39].

While learning, the two output maps encode the following information, basically corresponding to the random posture angles generated by motor babbling: (a) the arm posture (two angles); (b) the fingers posture (one value corresponding to the thumb's angles, and one value corresponding to the "virtual finger's" angles, see Sect. II). Motor babbling implies: (a) setting either a big (diameter: $30cm$) or small object (diameter: $12cm$) close to the system's hand palm; (b) causing the closure of the hand around the object to mimic the enclosure reflex (this is done by issuing suitable desired angles to the PDs muscle models); (c) issuing desired random postures to the arm. While this is done, the Hebb covariance learning rule of Eq. 7 is used to update the connection weights of the model so as to form associations between: (a) the eye-posture (signaling the position of the target) and the corresponding arm posture (this mimics the acquisition of reaching skills); (b) the object perception and the corresponding hand posture (this mimics the acquisition of different visually-triggered grips).

During later test stages, the sight of the object and the corresponding eye posture activate the two input maps which, on their turn, activate the two output maps: this latter activation sets the desired arm and hand postures (angles) on the basis of Eq. 4 (here b_1 and b_2 are the arm and hand desidered angles). Importantly, however, the reaching and grasping movements are triggered in different times. In particular, it is assumed that the simple *sight* of the object *triggers only reaching* whereas *grasping* is triggered only by the sight of the object plus the later *hand contact* with the object caused by the execution of reaching. This implies that the *composition* of the reaching-grasping sequence is done on the basis of the enclosure reflex.

B. Results

With learning the model develops a relatively effective reaching ability with both small and big objects (see Fig. 7, thin lines). On the contrary, the system succeed to accomplish the whole sequence (reaching *and* grasping) in few trials (see Fig. 7, thick lines). This result was expected as in order to perform an effective grasping the hand needs to move to the object with the hand palm facing: with an unplanned reaching this happens only rarely. Table II show

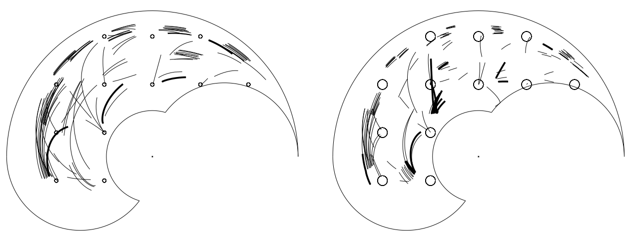

Fig. 7. **Left**: final portions of the hand trajectories performed by the grasping model in 132 tests that used 12 positions of the object and 11 hand initial positions distributed on the vertexes of a regular grid within the reachable working space. Thin lines indicate a successfull reaching toward a small object whereas bold lines indicates a successfull reaching *and* grasping. **Right**: the same test with big objects.

a quantitative measure of performance of the grasping and reaching. The data indicate that the system tends to close too much the thumb with big objects and not enough with small objects: this is due to the interference between the learning processes related to the fovea image of the two objects. Interestingly, the performance is lower with small objects than with big ones. Direct observation of behaviour indicated that this happens as reaching is based only on the eye posture (the object "where"). This leads the system to learn desired postures for reaching that average between the postures corresponding to grasped small objects and grasped big objects. As a consequence, the reaching behaviour tends to drive the hand "inside" the big objects and a bit far away from the small objects: the latter condition implies that the enclosure reflex does not succeed to trigger the grasping behaviour. Fig. 8 shows the weights emerged with learning. The values of the outward weights relative to reaching have assumed a typical Mexican hat distribution: this result confirm the results found in [8]. The same figure shows the configuration of the weights related to grasping: also these weights have developed a contrast-enhancing distribution due to the covariance Hebb learning rule.

TABLE II

RESULTS OF 144 REACHING AND GRASPING TESTS WITH SMALL AND BIG OBJECTS. $\langle EPs \rangle$: MEAN OF EQUILIBRIUM POINTS AFTER THE LEARNING PHASE; TV: IDEAL VALUES FOR EPS; STD: STANDARD DEVIATION; *Succ*%: PERCENTAGE OF SUCCESSES ON OVERALL TESTS.

		Grasp			Reach	Grasp/Reach
		$\langle EPs \rangle$	TV	STD	*Succ*%	*Succ*%
Small	thumb	36.99	20	0.027	71.5	2.8
	index	17.30	20	0.012		
Big	thumb	12.31	50	0.017	93.1	11.1
	index	46.34	50	0.050		

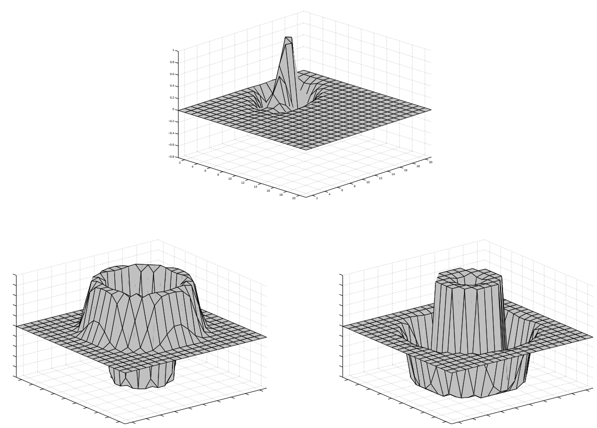

Fig. 8. **Top**: activation potential (z-axis) of the reaching output map (xy-plane) when the reaching input map is activated by an object. The outward weights of each input-map neuron have a similar configuration (data not shown). **Bottom**: inward weights of the two neurons of the grasping output map corresponding to the average grasping posture reported in Table II for the big (left graph) and small (right graph) objects.

V. CONCLUSIONS

Motor babbling might play an important role in motor development as it is based on simple and effective principles. So far, models on motor babbling have been mainly used to tackle the acquisition of simple reaching behaviours. This might be due to the fact that motor babbling alone might allow forming associations only related to events that are *instant* and *synchronous*. The main message of the paper is that infants might still use motor babbling for acquiring behaviours more complex than simple reaching by overcoming its time limits on the basis of other complementary mechanisms. To show the computational soundness of this hypothesis, the paper proposed two models that use motor babbling to acquire a behaviour of reaching with obstacles and a behaviour of grasping. The first model overcomes the time limits of motor babbling by using it to acquire the *parameters* used to control *Pattern Generators*, whereas the second model overcomes such limits by using *reflexes* to compose behavioral sequences in time.

The model of reaching with obstacles has another strength (a novelty with respect to previous models) in that it uses two TBGs to produce non-linear reaching trajectories while motor babbling allows learning the parameters regulating them on the basis of a simple Hebb rule. On the other side, it also has important limitations, in particular it has many hardwired parts which are not biologically plausible, such as the process that rotates and scales the trajectory produced by the TBGs and the inverse kinematics process that transforms the trajectory values produced by the TBGs from work space to joint space coordinates.

The interest of the grasping model is that it proposes a working hypothesis of the development of grasping behaviours based on motor babbling and the enclosure reflex. In particular, this reflex (a) supports the acquisition of basic reaching and grasping capabilities on the basis of motor babbling and a simple Hebb rule, and (b) supports a first rudimentary composition of the behaviours so acquired. A drawback of the model is its rather poor performance (this might be probably overcome with a more sophisticated representation of the foveated objects) and the absence of more sophisticated learning mechanisms for refining the behaviours' quality and their composition.

Notwithstanding their limits, the models show that motor babbling might indeed be used by organisms to develop complex behaviours, especially if they have not the function of achieving a high efficiency but rather to "scaffold" the later development of more accurate behaviours on the basis of other more sophisticated mechanisms.

REFERENCES

[1] J. Piaget, *The origins of intelligence in children*, I. U. Press, Ed., Newyork, 1952.

[2] C. von Hofsten, "Eye-hand coordination in newborns," *Developmental Psycology*, no. 18, pp. 450–461, 1982.

[3] G. Pezzulo, G. Baldassarre, M. V. Butz, C. Cristiano, and J. Hoffmann, "From actions to goals and vice-versa: theoretical analysis and models of the ideomotor principle and tote," in *Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior*, M. Butz, O. Sigaud, G. Pezzulo, and G. Baldassarre, Eds. Berlin: Springer-Verlag, 2007, pp. 73–93.

[4] M. Kuperstein, "Neural model of adaptive hand-eye coordination for single postures." *Science*, vol. 239, no. 4845, pp. 1308–1311, 1988.

[5] M. I. Jordan and D. E. Rumelhart, "Forward models: supervised learning with a distal teacher," *Cognitive science*, vol. 16, pp. 307–354, 1992.

[6] P. Morasso and V. Sanguineti, "Self organizing body schema fo motor planning," *Journal of Motor Behaviour*, vol. 27, no. 1, pp. 52–66, 1995.

[7] G. Asuni, F. Leoni, E. Guglielmelli, A. Starita, and P. Dario, "A neuro-controller for robotic manipulators based on biologically inspired," in *Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering*, Y. Demiris, B. Scassellati, and D. Mareschal, Eds., Capri Island-Italy, 2003, pp. 450–453.

[8] D. Caligiore, D. Parisi, and G. Baldassarre, "Toward an integrated biomimetic model of reaching," in *Proceedings of 6th IEEE International Conference on Development and Learning (ICDL 2007)*, Y. Demiris, B. Scassellati, and D. Mareschal, Eds. London: Imperial College, 2007, pp. E1–6.

[9] D. Ballard, "Animate vision," *Artificial Intelligence*, vol. 48, pp. 57–86, 1991.

[10] R. Shadmehr and S. P. Wise, Eds., *The Computational Neurobiology of Reaching and Pointing*. Cambridge, MA: The MIT Press, 2005.

[11] A. G. Feldman, "Once more on the equilibrium-point hypothesis (lambda model) for motor control." *J Mot Behav*, vol. 18, no. 1, pp. 17–54, 1986.

[12] J. Flanagan, D. Ostry, and A. Feldman, "Control of trajectory modifications in target-directed reaching." *J Mot Behav*, vol. 25, no. 3, pp. 140–152, 1993.

[13] N. E. Berthier, M. T. Rosenstein, and A. G. Barto, "Approximate optimal control as a model for motor learning," *Psychological Review*, vol. Vol. 112, no. No. 2, p. 329?346, 2005.

[14] L. Sciavicco and B. Siciliano, *Modeling and Control of Robot Manipulators*, McGraw-Hill, Ed., 1996.

[15] D. Bullock and S. Grossberg, "Vite and flete: neural modules for trajectory formation and postural control." in *Volitional Action*, W. Hershberger, Ed. Amsterdam: Elsevier., 1989, pp. pp. 253–298.

[16] N. E. Berthier, R. K. Clifton, D. D. McCall, and D. J. Robin, "Proximodistal structure of early reaching in human infants." *Exp Brain Res*, vol. 127, no. 3, pp. 259–269, Aug 1999.

[17] M. A. Arbib, "Grounding the mirror system hypothesis for the evolution of the language-ready brain," in *Simulating the Evolution of Language*, A. Cangelosi and D. Parisi, Eds. London: Springer Verlag, 2002, ch. 11, pp. 229–254.

[18] P. Bendahan and P. Gorce, "A neural network architecture to learn arm motion planning in grasping tasks with obstacles avoidance," *Robotica*, vol. 24, pp. 197–204, 2006.

[19] R. Volpe and P. Khosla, "Artificial potentials with elliptical isopotential contours for obstacle avoidance," in *26th IEEE Conference on Decision and Control*, 1987, pp. 180–185.

[20] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," *The International Journal of Robotics Research*, vol. 5, pp. 90–98, 1986.

[21] Y. Koren and J. Borentstein, "Potential field methods and their limitations for mobile robot navigation," in *Proceedings of the IEEE International Conference on Robot and Automation*, 1991, pp. 1398–1404.

[22] I. Iossifidis and G. Schöner, "Dynamical system approach for the autonomous avoidance of obstacle and joint-limits for an redundant robot arm," in *Proceedings fo the IEEE 2006 International Conference on Intelligent Robots and Systems (IROS)*, 2006, pp. 9–15.

[23] J. Demiris and G. Hayes, "Imitation as a dual-route process featuring predictive and learning components: A biologically-plausible computational model," *Imitation in Animals and Artifacts*, pp. 327–361, 2002.

[24] S. Schaal, A. Ijspeert, and A. Billard, "Computational approaches to motor learning by imitation," *Philosophical Transaction of the Royal Society of London*, vol. 358, pp. 537–547, 2003.

[25] J. Saunders, L. N. Chrystopher, K. Dautenhahn, and A. Alissandrakis, "Self-imitation and environment scaffolding for robot teaching," *International Journal of Advanced Robotic Systems*, vol. 4, no. 1, pp. 109–124, 2007.

[26] Y. Tanaka, T. Tsuji, V. Sanguineti, and P. G. Morasso, "Bio-mimetic trajectory generation using a neural time-base generator," *J. Robot. Syst.*, vol. 22, no. 11, pp. 625–637, 2005.

[27] S. Vahdat, A. Maghsoudi, M. Haji Hasani, F. Towhidkhah, S. Gharibzadeh, and M. Jahed, "Adjustable primitive pattern generator: a novel cerebellar model for reaching movements." *Neurosci Lett*, vol. 406, no. 3, pp. 232–234, Oct 2006.

[28] A. Pouget and P. E. Latham, "Population codes," in *The Handbook of Brain Theory and Neural Networks*, 2nd ed., M. A. Arbib, Ed. cambridge, MA, USA: The MIT Press, 2003.

[29] A. P. Georgopoulos, R. E. Kettner, and A. B. Schwartz, "Primate motor cortex and free arm movements to visual targets in three-dimensional space. ii. coding of the direction of movement by a neuronal population." *J Neurosci*, vol. 8, no. 8, pp. 2928–2937, 1988.

[30] C. Buneo, M. Jarvis, A. Batista, and R. Andersen, "Direct visuomotor transformations for reaching," *Nature*, no. 416, pp. 632–636, 2002.

[31] L. Natale, F. Nori, G. Sandini, and G. Metta, "Learning precise 3d reaching in a humanoid robot," in *Proceedings of 6th IEEE International Conference on Development and Learning (ICDL 2007)*, Y. Demiris, B. Scassellati, and D. Mareschal, Eds. London: Imperial College, 2007, pp. E1–6.

[32] T. J. Sejnowski, "Storing covariance with nonlinearly interacting neurons." *J Math Biol*, vol. 4, no. 4, pp. 303–321, 1977.

[33] G. Rizzolatti, R. Camarda, L. Fogassi, M. Gentilucci, G. Luppino, and M. Matelli, "Functional organization of inferior area 6 in the macaque monkey. ii. area f5 and the control of distal movements." *Exp Brain Res*, vol. 71, no. 3, pp. 491–507, 1988.

[34] D. A. Rosenbaum, "Human motor control," Academic Press. Amherst, Massachusetts, Tech. Rep., 1991.

[35] C. Lantz, K. Melen, and H. Forssberg, "Early infant grasping involves radial fingers," *Developmental Medicine and Child Neurology*, vol. 38, pp. 668–674, 1996.

[36] R. K. Clifton, W. M. Darwin, D. H. Ashmead, and M. G. Clarkson, "Is visually guided reaching in early infancy a myth?" *Child Development*, vol. 64, pp. 1099–1110, 1993.

[37] C. von Hofsten and L. Ronnqvist, "Preparation for grasping an object: a developmental study." *J Exp Psychol Hum Percept Perform*, vol. 14, no. 4, pp. 610–621, Nov 1988.

[38] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing," Presentation for Stanford Artificial Project, 1968.

[39] D. H. Hubel, *Eye, brain and vision*. New York: Scientific American Books, 1988.