The Organization of Cortex-Ganglia-Thalamus to Generate Movements From Motor Primitives: a Model for Developmental Robotics

Alessio Mauro Franchi[1], Danilo Attuario[2], and Giuseppina Gini[3]

Department of Elettronica, Informazione and Bioingegneria Politecnico Di Milano - Milan, Italy
[1]alessiomauro.franchi@polimi.it
[2]danilo.attuario@mail.polimi.it
[3]giuseppina.gini@polimi.it

Abstract. The advent of humanoid robots has posed new challenges and opportunities to control complex movements; their bodies have an high number of degrees of freedom, and methods used up to now to control them are no longer efficient. The purpose of this work is to create a system that could approach these challenges. We present a bioinspired model of the cortex-basal ganglia circuit for movement generation. Our model is able to learn and control movements starting from a set of motor primitives. Experiments on the NAO robot show that the system can be a good starting point for a more complex motor system to be integrated in a bioinspired cognitive architecture.

Keywords: developmental robotics, motor primitives, ganglia, cortex, thalamus, movement, learning

1 Introduction

Our task is to study and model the natural pathway to generate movements; in the brain it is based on the interaction between the cortex, the base ganglia and the thalamus [4]. Indeed in our model we concentrate on the interactions between cortex and ganglia, since the thalamus has only the role of information exchange.

Our solution is based on motor primitives, which represent simple movements used as building blocks to generate more complex motions [6]. Our model is divided into two areas; the first one derives from our reference bioinspired architecture, namely IDRA, able to model the sensorial cortex by a mechanism of dimensionality reduction and compact state representation [5]. Each internal state has an associated value of interest that can derive from the experience or be innate (a priori). The second area represents the motor cortex and the ganglia; it combines the motion primitives and a module devoted to learning such composition through reinforcement learning.

The novel result we want to obtain within this research is to use a reduced set of reflex movements, only those available in humans since birth. The model has been tested on the NAO robot to evaluate its value in reaching-like tasks.

2 Motor Primitives and the Definition Chosen

In the embodied view of robotics movement is a way to develop knowledge, since it enables active perception, but also the vice versa is true [11]. The human motor system is able to generate a large variety of motions starting from reflexes and voluntary movements [6].

There are two kinds of reflexes: simple “extensions and contractions”, present at the fetal state, and “primitive reflexes” of the newborns. Neuroscientific hypotheses state that simple primitives are innate and other are learned from experience inhibiting the innate reflexes [3]. There is a large literature about motor primitives, both in humans and robotics; these are considered as an important mechanism for motor learning and motor control [6] [9]. A common study method uses the EMG signals to derive those primitives [17]; other studies concentrate on how the nervous system is able to combine those simple primitives for coordinated intentional actions [7] [2]. It seems that each primitive has a specific aim, as for instance to control the distance from the hand to a target object during reaching. This last definition of motor primitive has inspired our model of the motor cortex, which is able to describe movements starting from a group of primitives all with the same aim.

Different mathematical formulations are available [14][2]. The model that in our opinion is the most bioinspired one is the Dynamic Movement Primitives (DMP) model that represents only kinematics [13]. DMPs are represented as differential equations. A movement can be represented as a mapping from a state vector to a command vector to the joints. This function depends on parameters that are specific for the activity to accomplish and that should be learned, typically using reinforcement learning [16]. Since learning for a large space of action-states is impractical, the combination of basic functions is a possible solution. In a bioinspired solution those basic functions are exactly the motor primitives.

Some applications of DMP can be seen in robotics. In [14] the human robot Sarcos has been trained to perform tennis forehand and backend; a similar experiment is proposed in [8] where a human teacher shows to a seven DOF robotics arm how to play ping-pong; [12] studied a quite different motor primitive representation which allowed the creation of a system able to learn different tasks and re-use shared knoledge. All of these experiments showed good results, but are different from our proposed study mainly in the type of primitives: they all use task-specific motor primitives, whereas we defined a few number of generic primitive reflexes, those presents in newborns. Our work aim at demonstrating the generic human-like primitive reflexes are sufficient for motor skill development.

According to the definitions given in [13], we use a purely kynematics notation, so the output is the velocity and acceleration of joints. This formulation makes it possible to ignore the non linearity due to external forces that are solved by the controller during the execution of the motion, a task that the cerebellum account for in humans [15].

3 The Implemented Method

Each DMP can be formalized by a second order dynamical system (for discrete movements) and a basic point attractive system (for rhytmic movements) [14], respectively as in 1 and 2:

$$\tau\dot{v} = \alpha_v(\beta_v(g-y)-v), \quad \tau\dot{x} = v \tag{1}$$

$$\tau\dot{z} = \alpha_z(\beta_z(g-y)-z), \quad \tau\dot{y} = z + f \tag{2}$$

where g is the known target position of the moviment, α_z (α_v) and β_z (β_v) are time constants, τ is a time-scaling factor, y and $\dot{y}$ are position and velocity generated by the equations and x is a phase variable.

The first equation is linear and monotonically convergence to the goal g, and is necessary for dampening the second equation, which by itself may result in a very complex equation due to the nonlinearity of f.

We use these DMPs as model for the generation of angular trajectories in terms of velocity; for each degree of freedom we have a single transformation system, while the canonical system is unique in order to synchronize each trajectory in time.

In order to generate a movement we need to determine the parameters of f; we used imitiation learning for this, formulated in the following simplified form (Eq. 3, 4):

$$f_{target} = \tau\dot{y}_{target} - z_{target} \tag{3}$$

$$\tau\dot{z}_{target} = \alpha_z(\beta_z(g - y_{target}) - z_{target}) \tag{4}$$

where y_{target} and $\dot{y}_{target}$ are given, Eq. 3 is the target trajectory for the right part of Eq. 2 and z_{target} is computed by integrating the left part of Eq. 4.

It can be demonstrated that Eq. 1, 2, 3 and 4 converge to g in time T.

For movement learning we search for a policy π binding a state vector x to a vector of command q in terms of position or speed or acceleration of joints. Learning this kind of policy is computationally intractable; for semplification we can write the problem as a composition of N simpler policies π_k, i.e. motor primitives, as in Eq. 5.

$$q = \pi(x,t) = \sum_{k=1}^{N} \pi_k(x,t) \tag{5}$$

Now, given a set of functional primitives $\pi_{k,i}(x)$, each with the same objective i, we define a new policy combining these primitives, as in Eq. 6.

$$\pi(x) = \sum_{k=1}^{N_i} \frac{\gamma_k \pi_{k,i}(x)}{\sum_{h=1}^{N_i} \gamma_h} \tag{6}$$

We can also create new more complex movements by combining primitives with different objectives $i \in [1 \dots j]$ (Eq. 7):

$$\pi(x) = \sum_{k=1}^{N_1} \frac{\gamma_k \pi_{k,1}(x)}{\sum_{h=1}^{N_1} \gamma_h} + \dots + \sum_{k=1}^{N_j} \frac{\gamma_k \pi_{k,j}(x)}{\sum_{h=1}^{N_j} \gamma_h} \tag{7}$$

The input of this system are the trajectories generated from DMPs, the output is a final novel trajectory. Learning this model means finding the optimal weights γ_i; we can model these weights with a linear equation (Eq. 8):

$$\gamma_i(x) = \Theta_i^T \phi_i(x) \tag{8}$$

where Θ_i is the vector of parameters and ϕ_i is the vector of the basis functions. As basis function we used simple gaussian function so that the policy is deterministic; as the policy is deterministic we have then added a $\epsilon \sim N(0, \Sigma)$ term for exploration. For weights optimization we use the reinforcement learning Natural-Actor-Critic algorithm, with SARSA(1) for the Critic [10]. In this algorithm it is important to have a learning rate of the Critic greater than that of the Actor. All the other parameters of this model have been esperimentally determined.

The rewards the robot receives after each movement are weighted by the amount of contribute of the i^{th} primitive in the final movement (Eq. 9).

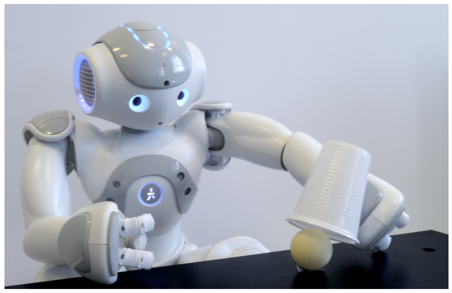
$$R_i = R \frac{\gamma_i}{\sum_{j=1}^{N} \gamma_j} \tag{9}$$

where R is the total reward after the execution of an action and R_i corresponds to the i_{th} primitive.

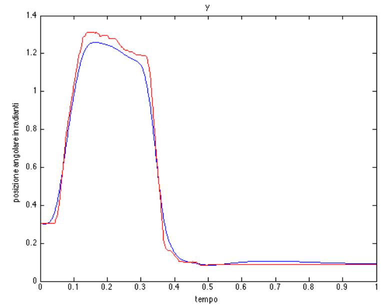
4 The Experiments

The task for our experimentation is to ask the NAO robot to cover a ball with a glass fixed in the hand starting with the left arm in a random position (Fig. 1(a)). The basic primitives are only seven neonatal reflexes and two acquired primitives for rotation [3], as listed in Table 1. Trajectories have been obtained in a kinesthetic fashion; joints values are recorded at time step of 0.4 seconds, for a total of 3 seconds (Fig. 1(b)).

The first experiment is to evaluate the capability of composing primitives to generate new movements. The modules used are only cortex and ganglia, the input is the position of the ball expressed in the joints space; the learning rate is set to 0.85 for the critic and 0.35 for the actor. To evaluate the learning we checked the two values of the reward, Cartesian and Angular. The first evaluates the distance to the ball, taking the $x^{\frac{1}{3}}$ of the distance (x); the second the hand orientation with respect to a target orientations. This splitting is due to the fact that the seven neonatal reflexes are used to make the reaching, the other two primitives for orientation.



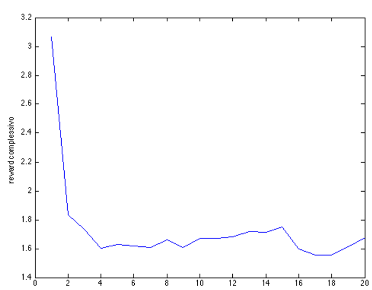
(a) The NAO robot trying to perform the reaching task



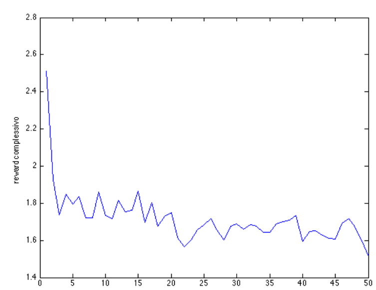
(b) Trajectory of a DMP approximating the shoulder rotation for the swimming reflex

Fig. 1. The NAO robot performing the task (1(a)) and the recorded shoulder rotation values (1(b))

The total reward for the first experiment is reported in Fig. 2(a); as the number of iterations of the experiment increases, the reward get lower, meaning that the robot is getting closer to its goal. It never reaches zero mainly due to noisy values of the joints; in human a visual feedback is used to correct trajectory at runtime [1]. We may also see that it is a little unstable, probably due to the choice of a Gaussian policy with a fixed variance.



(a) Total reward as computed from the first experiment



(b) Total reward for the robot from the experiment with the IDRA module

Fig. 2. The rewards the robot received after the first and second experiment

The experiment has been repeated with the integration in the IDRA architecture; the architecture is fully reported in [5]. In this case we wanted to use the Intentional Module to evaluate the interest of the robot for the state. We add also the visual input, with the application of a log-polar filter for simulating human vision. The ball position is obtained from the central pixel of the ball. A set of a-priori recorded images of the environment has been used for learning the IDRA modules.

Table 1. The learned composition of primitives to accomplish the reaching task

Weight	0.45	0.413	0.082	0.452	0.85	0.12	0.847	0.63	0.2
Primitive	Cervical	Galant	Moro1	Moro2	fall	swim1	swim2	Rot1	Rot2

In the new experiment no substantial differences in the rewards are found (see Fig. 2(b)). This means that the motion composition in-se may work without any visual feedback; moreover the integration into a more general cognitive architecture preserves its properties. We may expect advantages from vision and our cognitive architecture in more complex tasks that we will devise and experiment.

5 Conclusion

In this work we started from a known and quite accepted view about how to generate movements only from linear combination of motion primitives. Using this approach we have shown that the limited set of innate reflexes are able, almost alone, to generate a complex reaching trajectory with good results. Moreover, our technical implementation is quite new, using a mixture of experts and the Actor-Critic algorithm for learning. It is out of the scope of this paper to compare the performance of the obtained movement to other models in literature, mainly due to the different set of primitives and the difficulty in replicating the same experiments.

References

1. Ajemiana, R., DAusilio, A., Moormand, H., Bizzi, E.: A theory for how sensorimotor skills are learned and retained in noisy and nonstationary neural circuits. In: Proceedings of the National Academy of Sciences of the USA. vol. 110 (2013)
2. Alessandro, C., Delis, I., Nori, F., Panzeri, S., Berret, B.: Muscle synergies in neuroscience and robotics: from input-space to task-space perspectives. Frontiers in Computational Neuroscience 7(43) (April 2013)
3. Camaioni, L., Di Blasio, P.: Psicologia dello sviluppo. Il Mulino (2007)
4. Doya, K.: What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? Neural Networks 12(7-8), 961–974 (1999)
5. Franchi, A.M., Gini, G., Mutti, F.: A bioinspired intentional architecture for situated agents to explore human cognitive abilities. In: Proc. IARP Conference, Frascati. pp. 37–43 (May 2014)
6. Konczak, J.: On the notion of motor primitives in humans and robots. In: Proc. Modeling Cognitive Development in Robotic Systems. pp. 47–53 (2005)
7. Lee, W.A.: Neuromotor synergies as a basis for coordinated intentional action. Journal of Motor Behavior 16(2), 135–170 (1984)
8. Muelling, K., Kober, J., Kroemer, O., Peters, J.: Learning to select and generalize striking movements in robot table tennis. International Journal of Robotics Research 32(3), 263–279 (2013)
9. Paine, R.W., Tani, J.: Adaptive motor primitive and sequence formation in a hierarchical recurrent neural network. In: From Animals to Animats 8. pp. 274–283 (2004)

10. Peters, J., Schaal, S.: Natural actor-critic. Neurocomputing 71(7-9), 1180–1190 (2008)
11. Pfeifer, R., Lungarella, M., Iida, F.: Self-organization, embodiment, and biologically inspired robotics. Science 318(5853), 1088–1093 (2007)
12. Rückert, E., d'Avella Andrea: Learned parametrized dynamic movement primitives with shared synergies for controlling robotic and musculoskeletal systems. Frontiers in Computational Neuroscience 7(138) (2013)
13. Schaal, S.: Dynamic movement primitives - a framework for motor control in humans and humanoid robotics. International Symposium on Adaptive Motion of Animals and Machines pp. 261–280 (2006)
14. Schaal, S., Peter, J., Nakanishi, Jun ahd Ijspeert, A.: Control, planning, learning, and imitation with dynamic movement primitives. IROS pp. 1–21 (2003)
15. Schweighofer, N., Arbib, M.A., Kawato, M.: Role of the cerebellum in reaching movements in humans. i. distributed inverse dynamics control. European Journal of Neuroscience 10(1), 1088–1093 (2009)
16. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT Press (1998)
17. Tresch, M.C., Saltiel, P., dAvella, A., Bizzi, E.: Coordination and localization in spinal motor systems. Brain Research Reviews 40(1-3), 66–79 (2002)