

Learning and communication via imitation: an autonomous robot perspective

P. Andry, P. Gaussier, S. Moga, J. P. Banquet¹, J. Nadel²
Neurocybernetic team, ETIS Lab, UPRES A 8051 UCP-ENSEA
ENSEA, 6 avenue du Ponceau, 95014 Cergy, France
andry@ensea.fr, gaussier@ensea.fr

¹Neurosciences et modélisation INSERM U483 Paris.

²Equipe Développement et Psychopathologie UMR CNRS 7593
Hopital de la Salpêtrière, Pavillon Clérambault
54 bd de l’Hopital, 75005 Paris, France

Abstract—This paper proposes a neural network architecture designed to exhibit learning and communication capabilities via imitation. Our architecture allows a “proto imitation” behavior using the “perception ambiguity” inherent to real environments. In the perspective of turn-taking and gestural communication between two agents, new experiments on movement synchronization in an interaction game are presented. Synchronization is obtained as a global attractor depending on the coupling between agents’ dynamic. We also discuss the non-supervised context of the imitation process and we present new experiments in which the same architecture is able to learn perception-action associations without any explicit reinforcement. The learning is based on the ability to detect novelty or irregularities in the communication rhythm.

Keywords—Imitation, Neural Network, Implicit Reinforcement, Autonomous Learning, Interaction, Synchronization.

I. INTRODUCTION

IMITATION is a powerful learning paradigm for autonomous systems. As a capability of learning by observation, imitation can improve and speed up the learning of sensory-motor associations. In order to design a neural architecture able to learn via imitation, we are mainly interested by the first levels of imitation [30]. These levels concern low-level imitations, *i.e.* reproduction of meaningless and simple movements. This “proto imitation” level plays a key role in understanding the principles of the perception-action mechanism necessary to perform higher order behaviors. But imitation also has a communicative function [1]. As observed among nonverbal children [20], imitation is a powerful tool for gestural interactions consisting of initiations of motor sequences by one child and synchronous re-enactments by the other child (Fig. 1). The imitative capabilities involved in earlier stages of learning and communication lead us to find out how our previous proto imitation model can be used to understand interactions in a human-machine or a machine-machine loop. In section II, we present a brief account of recent knowledge of imitation in the field of developmental psychology. We focus on what is necessary for an architecture to learn and communicate via imitation. Therefore, we review only the early stages of imitation as observed among neonates



Fig. 1. Young children interacting and turn-taking. Imitation deserves a communication purpose, where synchronization and rhythm matter. Reproduction from [20].

and infants, rather than the more sophisticated capabilities involved in planning imitation at a program level. Section III summarizes the key features of the neural network architecture developed in previous works [12], [4], where imitation was used as a tool to simplify the learning of motor sequences. Based on this architecture, we present in section IV, a new architecture allowing synchronization capabilities in the perspective of gestural communication and turn-taking between two agents. We show that synchronization can be obtained from a simple modulation of the perception-action loop: perception can be seen as an energy exchange between the interacting systems. Finally we demonstrate in section V that the rhythm of the interaction carries out implicit information that can then be used to build a powerful reward signal. This signal is then used to learn an arbitrary set of sensory-motor associations in reinforcement learning [28].

II. DEVELOPMENTAL PSYCHOLOGY BACKGROUND

A. *The traditional view: deferred imitation as a building block for representational capabilities*

Up to the seventies, two claims have framed developmental studies on imitation. First, generations of developmen-

tal psychologists have followed Guillaume’s definition of imitation [14], which implies that imitation should require at least an elementary level of representation. As a consequence, most developmental psychologists in the fifties have denied any ontogenetic value to immediate imitation, and have emphasized instead the developmental role of deferred imitation. Defined as the delayed re-enactment of an absent motor model, deferred imitation was considered by Piaget [24] and most of contemporary child psychologists, as a determinant building block for emerging representational capabilities, thus serving mentalization. As a second claim, true imitation had to be novel and not already in the repertoire. This led to see imitation as a special case of observational learning occurring without incentives, without trial and error, and requiring no reinforcement [3]. Within this framework, the developmental benefits of imitation are to provide an alternative to expensive trial-and-error learning; to facilitate the rapid acquisition of adaptive behavior in the young; to allow a direct incorporation of the learned repertoire of the society [9].

B. Current view: immediate imitation as serving two functions: learning and communication

Developmental studies carried out since the 1970s have completely altered the theoretical basis for understanding the contribution of imitation to development. The emphasis has been on the innate origin of imitation in humans [16], [17], and on the contribution of perceptual systems as channels of information, which enable the infant to imitate through perception-action linkage [18]. Explaining neonatal imitation requires a theory of perception-action coupling accounting for how coordinated behaviors emerge between individuals and how more advanced forms of imitation develop. Immediate imitation has thus been rehabilitated. It is clear now that it is not the unintelligent process Piaget described, but rather a way to learn new “know-how” as well as to communicate. Through perceptual information exchange, immediate imitation does not only serve the social transmission of culturally accumulated characteristics but also individual identification [19], communication and role-taking [15], [21]. Indeed, as shown by Nadel and colleagues [20], preverbal children use imitation in their social exchanges as a way to take turns, switch roles (imitator versus imitated) and to share topics.

III. THE STARTING POINT FOR AN “IMITATING BEHAVIOR” IMPLEMENTATION

Inspired by developmental data, our model of an imitative autonomous system is designed to learn perception-action coupling, and to take part in a communication process via imitation. In a long-term perspective, the system should respond to the teacher’s expectancies like in the “do as I do” test proposed by Hayes and Hayes [30], communicating according to a “start and end” signal given by the teacher. Moreover, the system should also be capable of spontaneous imitation generated by an internal motivation or simply triggered by novelty detection. Spontaneous imitation may be an essential part of the social interactions

among a population of robots [7], [11].

A. Imitation triggered by a perception ambiguity

Intentions, emotions, identity, consciousness of self or supra modal levels are often mentioned as notions involved in imitation. But these words are not fully understood, and very hard to consider for designing autonomous systems. So the first question that arises is: is it possible to design a simple architecture (related to the complexity of the notions involved), which could exhibit basic imitation capabilities, e.g., a type of proto imitation? This architecture could be a starting point to explore more complex behaviors. We assume that imitation can be triggered by a perception ambiguity. “Perception ambiguity” must be understood here as a difficulty to discriminate objects (is this my arm or another’s one?), or to decide between different interpretations (is this a useful object, or an obstacle?). Perception ambiguity was first introduced by Gestaltists, assuming that local features in a perceived scene were always ambiguous (only the global contextual information and the dynamic of the perception-action loop allow to suppress ambiguity). In the present issue, we consider that the available information is egocentric (relative to the system’s position). For instance, an imitation behavior of an artificial arm controlled by vision can be obtained when the robot learns the visuo-motor coordination between its camera and its mechanical hand. During this learning phase, the robot creates a correspondence between a given position of its hand in the visual field and the angular position of the different joints. Assuming that the perception module of this system is quite simple (using for example movement detection), let us now suppose that this robot looks somewhere else and perceives (Fig. 2) another artificial arm, or simply a hand moving in its visual field (narrow field of view). Because our robot will try to reduce the differences between the representations it supposes to have of its arm, it will perform a similar movement as the hand. Moreover, if the sequence of movements is stored and associated with the satisfaction of a particular motivation, it can be triggered later. An external observer will consider that the robot has learned via imitation the behavior of the hand. To sum up, it is proposed that an imitation behavior can be induced by:

1. the ambiguity on the identification of the perceived extremity,
2. the minimization of the error between the visual and the motor positions (homeostasis principle [2]).

This proposed control principle is close to the “low-level resonance” mechanism, mentioned by Rizzolatti [26]. Experiments show that the same motor neurons (situated in the “rostral part of the inferior parietal lobule” of monkeys [26]) are activated when a monkey observes or produces meaningless arm movements, regardless of the execution context. According to Rizzolatti, such a “low-level resonance” account for low-level imitating faculties. This direct link between perception and action could allow imitation capabilities without high-level notions of “self” and “others”. Nevertheless, our approach is not aimed to per-

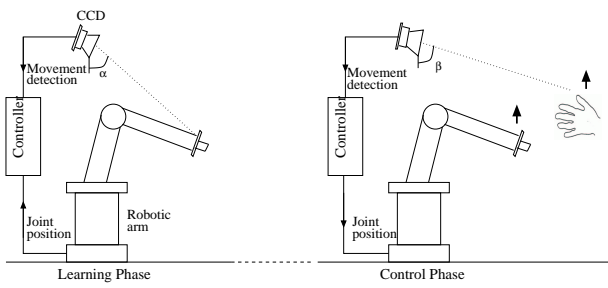


Fig. 2. The proto imitation principle applied to a robotic arm. In a learning phase, a controller robot learn the correspondence between its arm proprioception (the joint position) and its position in its visual field. To do this, the controller detects movement. Once the associations are learned, if the robot focuses its attention in a human teacher’s moving hand, it will reproduce the teacher’s simple movement just because it will perceive a difference between its proprioceptive and visual information. It will try to reduce the proprioceptive error of its arm position according to what it believes to be the visual information linked to its arm (the detection of movement in the visual field)! An external observer will then deduce the learner robot is imitating the teacher.

form exact physical motion reproduction between identical arms. We assume that some important information (such as the motion and the dynamic of the model) are sufficient to induce the premises of imitation. The perception ambiguity simply suppresses the need of recognizing others to initiate some simple group dynamics. It allows gestural imitation (or meaningless imitation) between systems having very different morphologies (robots, humans, animals...). The fidelity of the reproduction is of course dependent on the experimental setup, where distances, perspectives and positions can influence the perception and therefore the quality of the reproduction (amplitude of the movement’s trajectories). Short distances, same body orientations of the imitator and the imitated, will favor the quality of imitation, such as in sports and dance tutorials. Being able to imitate a movement, whatever the posture is, would imply for instance using a mapping mechanism allowing to store an invariant representation of the action to be performed (the trajectory should be stored independently of the effector and posture). The solution to a real and more complex imitation must take into account the “effect level” of imitation (where actions of the imitator achieved the same effect as the imitated [22]). These imitations require the understanding of the goal of the perceived actions. We assume that our proto imitation mechanism associated with an invariant representation of actions and the capability of learning objects affordances could be a way, in future works, to address correctly complex imitation problems (see for instance [10]). In the present works we only investigates the basic features necessary for simple unsupervised interactions between two autonomous agents.

B. Learning and reproducing temporal motor sequences

We summarize here previous work [12] using perception ambiguity to teach a robot different “dances”. This work is very important for the present paper since it is a prerequisite for the new architectures developed in sections IV

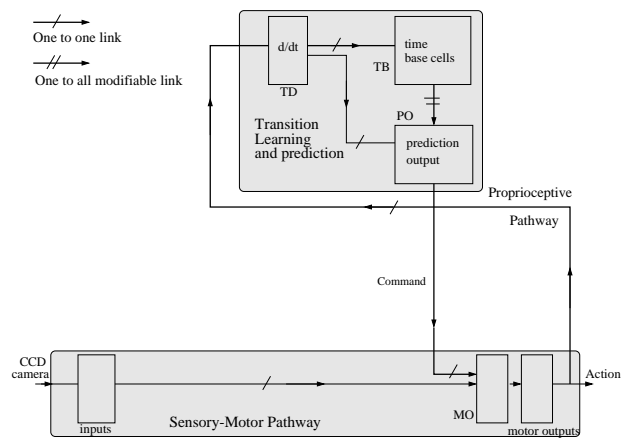


Fig. 3. The model, allowing proto imitation. The system is designed as a PerAc block. The “transition learning and prediction” mechanism is a perception level modulating the “sensory-motor pathway”. The PO group of neurons learns the transitions between the past event (activity of the TB neurons) and the present one (activity of the TD neurons). The link between PO and the MO motor group ensures the reproduction of a learned sequence.

and V for communication capabilities (synchronization of agents gestures in an imitation game) and rhythm prediction. We used a mobile (six-wheeled) robot, whose control architecture (Fig. 3) tries to reduce the difference of speed between the information of the visual flow and the information of the wheel speed (homeostasis principle [2]). If the robot is still, then any movement in the visual field will make the control system consider that this still position is an error. The control system will try to cancel these changes by modifying its wheel speed; a tracking behavior emerges. Our mobile robot is following the detected movements in its optical field, *i.e.* it “imitates” the moves of the human teacher. But, more than an simple reproduction, the system also learns the whole trajectory made by the teacher, in order to reproduce it later. By “whole trajectory”, we refer to the entire temporal sequence of motor actions. The robot learns to reproduce its own sequence of actions primarily induced by the tracking behavior. To do this, the control architecture of the robot is based on a neural network which learns and predicts the temporal succession of events. This is inspired by the functions of two brain structures involved in memory and time learning: the cerebellum and the hippocampus (see [4] for more neurobiological references). The network is made of three groups of neurons (Fig. 4): Time Derivation (TD) group, Time Base (TB) group and Prediction output (PO) group. The Time Base (TB) keeps a temporal trace of a past event. It is a set of neurons whose activity time course looks like a wavelet basis. The TB neurons are triggered in battery, each battery corresponding to a line of neurons. In the experiment, fifteen cells constitute a battery. Temporal activity of six cells of one of the batteries is presented in Fig. 5. The activation of a TB cell is computed as follows:

$$Act_{i,j}^{TB}(t) = \frac{m_0}{m_j} \cdot \exp - \frac{((t - \tau_i) - m_j)^2}{2 \cdot \sigma_j} \quad (1)$$

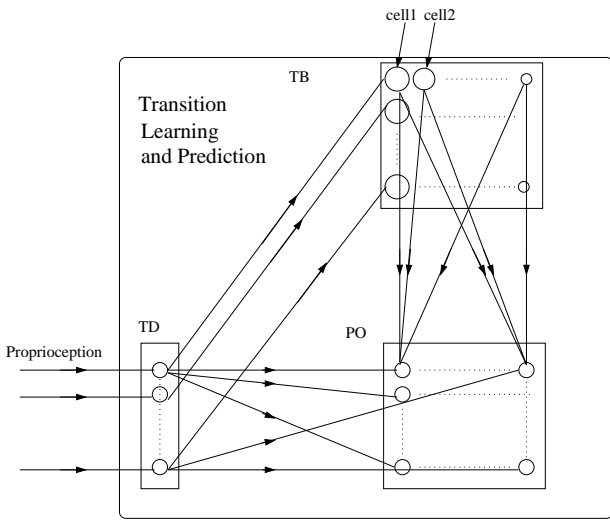


Fig. 4. Detailed connectivity of the event prediction network. The circle size in TB is associated to the time constants (m_j) of the neurons.

where i, j are the position of the cell in the group (the j th cell of the i th battery), m_j and σ_j are the time constant and the standard deviation associated to the j th cell. τ_i is the instant of the activation of the i th line of neurons. The Time Derivation (TD) group performs the derivation

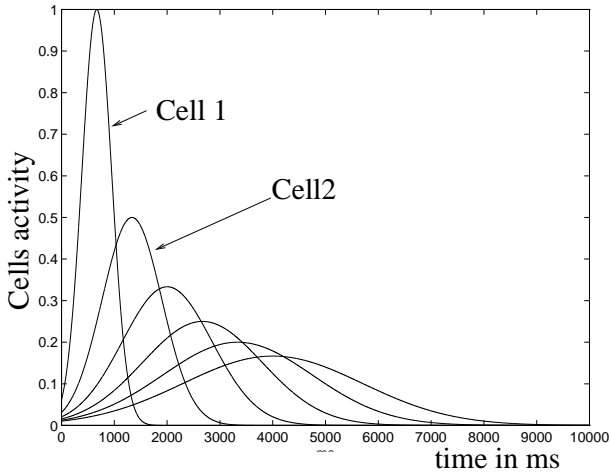


Fig. 5. Time activity of a group of cells that allows to measure time in the Time Base TB. Here TB is stimulated at $t=0$.

of the input signal. TD activates TB at the beginning of an input event. The Prediction Output (PO) group receives information from TD (the present event) and TB (a temporal trace of the past event). The timing between the past and present events is learned in the TB-PO connection weights. The strength between a PO neuron and a TB battery is modified according:

$$W_{PO(i,j)}^{TB(j,l)} = \begin{cases} \frac{Act_{i,l}^{TB}}{\sum_{j,l} (Act_{j,l}^{TB})^2} & \text{if } Act_j^{TD} \neq 0 \\ \text{unmodified} & \text{otherwise} \end{cases} \quad (2)$$

Because of the connectivity between TD, TB and PO, each PO neuron can learn a given transition between two events.

To n possible events corresponds $n \times n$ neurons in PO. The potential of a PO neuron is the sum of the information coming from TD and the delayed activity in TB. The potential of a PO neuron is computed as:

$$Pot_{i,j}^{PO} = \sum_l W_{PO(i,j)}^{TB(j,l)} \cdot Act_{j,l}^{TB} + W_{PO(i,j)}^{TD(i)} \cdot Act_j^{TD} \quad (3)$$

$Act_{j,l}^{TB}$ is the activity of the l cell of the j battery of TB. $W_{PO(i,j)}^{TB(j,l)}$ is the connection strength between the TB(j, l) neuron and the PO(i, j) neuron. One neuron of PO is linked with all of the neurons of a battery of TB. Act_j^{TD} is the corresponding TD neuron of a PO(i, j) neuron and $W_{PO(i,j)}^{TD(i)}$ the strength of link between them. A PO neuron only fires when its potential reaches its maximum value, which corresponds to the prediction of a new pattern:

$$Act_{i,j}^{PO} = f_{PO}(Pot_{i,j}^{PO}) \quad (4)$$

$$f_{PO}(x(t)) = \begin{cases} 1 & \text{if } \frac{dx(t)}{dt} < 0 \text{ and } \frac{dx(t-\tau)}{dt} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

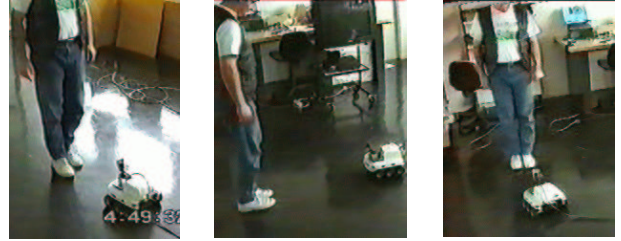


Fig. 6. Learning sequences of movements via imitation. The robot follows its human teacher, and learns the timing of its own movements. The sequence is then associated to a given motivation. Later, the robot will be able to reproduce the sequence according to the satisfaction of the motivation.

Hence, after a first sequence of actions (Fig. 6), a motor neuron will be activated by the input and will also receive information from the transition prediction group (PO). Then a simple conditioning rule allows the activated neuron to react the next time the action is predicted even if the input does not provide any information. Moreover, if the sequence of movements induces a positive reward, the past predicted transitions and their associated movements are reinforced. Thus, the robot learns to imitate the behavior (here, the sequence of movements) of the teacher according to an activated motivation. More details about the robotic issues such as the tracking system, the vision filtering, or how the system copes with multiple trajectories can be found in [12]. Such a robotic experiment is of course dependent of the experimental setup, the system being sensible to the perception ambiguity. Therefore a lot of parts of the experimental setup still have to be controlled by the human teacher: the speed of the demonstration, the distances, etc... Practically, the teacher also controls the succession of the perceptions of the robot. At learning phase, step by step perceptions induce the actions to be

learned. At reproduction phase, an initial perception triggers the demonstration. The start of an experiment is also explicitly signaled by the teacher (he touches a specific sensor of the robot), which induces the following consequences:

1. the teacher decides when the learning experiment begins and stops.
2. he explicitly gives the robot the information (as a reward) about the end of the learning.
3. he also gives a signal to induce the reproduction of the sequence.

These steps assume that the experimenter is a human agent. Adding the robot the capability to perceive or predict the end of the experiment would improve its autonomy, allowing it to decide spontaneous learning via imitation of others. Such an ability could be a great improvement among a group of robots. They could learn from each other new skills, as a kind of social transmission [30]. This statement lead us to turn to the communicative function of imitation in young children, who initiate actions so as to be imitated, observe and reproduce the other’s action, without any explicit “end of pattern” signal, and with internal capabilities to modulate self actions according to the other participant of the interaction.

IV. INTERACTION BETWEEN TWO SYSTEMS: SYNCHRONIZATION EFFECTS

In this section, we investigate interactions between two simulated robots that have prior knowledge of the same sequence of actions. This situation is inspired by games among young children where immediate gestural imitations are performed quasi-simultaneously (Fig. 1). Such interactions involve the capability of motor synchronization, where the same sequences of actions are executed at the same time. The synchrony can be seen as an attractor of a cyclic interaction game between two or more agents. We study here the overall dynamic of a loop made of two identical systems, with perception and action groups interconnected (Fig. 7), in order to understand which minimal features need to be added to our architecture. Sections IV-A and IV-B explain how minor improvements of the architecture can reach a trade-off between independent production and adaptive production according to the other, *i.e.* synchronization. The problem is how to use the sensory-motor pathway for synchronization, since it must be inhibited to avoid cyclic perturbation due to the interconnection. The developed solution is inspired by the “entrainment” phenomenon observed by C. Huyggens in 1665, in which two pendulum clocks placed on the same support synchronize themselves (“clock synchronization”). In our simulations, perception is similar to the physical wave transmitted by the support, it energizes the system’s actions.

A. Avoiding interferences in the reflex pathway

Let us suppose that both systems use the “transition learning and prediction” mechanism detailed in section III. To simplify the analysis and the simulations, perception and action signals are supposed to have binary values. The output of the first system is connected to the input of the

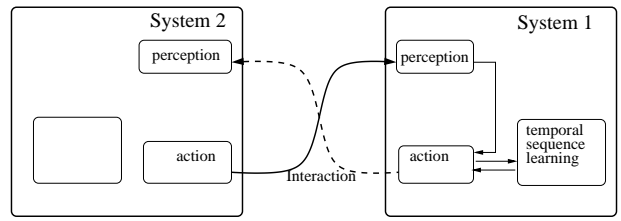


Fig. 7. Interconnection of two systems. System 1 and 2 have the same architecture. Each system has learned perception-action associations. The two systems must produce outputs (the same sequence of motor outputs for example) at the same time.

second one, and vice versa (this simulates perfect perceptions of the other’s action). Both systems have learned separately the same sequence of actions (for example transitions 1→2, 2→3, 3→1 are learned, allowing the production of the action sequence 1,2,3,1,2,.. and so on).

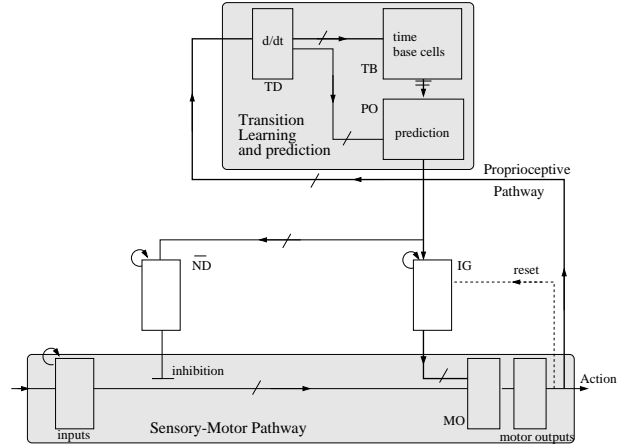


Fig. 8. Introduction of new elements to allow synchronization between agents. “Non novelty detection” (ND) and Integration (IG) groups are used to control the internal dynamic of the system.

In a machine-machine loop (Fig. 7), the direct pathway connecting perception to action of both systems plays an important role. It is critical for the learning process, but also interferes with the simple production of a given sequence. Thus, after learning, this sensory-motor pathway must be inhibited (Fig. 8), in order to allow an independent and complete production of the sequence without perturbations. The inhibition is realized by connecting “non novelty detection” neurons (\overline{ND} , Fig. 9) to the input group. Once activated, these neurons remain potentialized (due to self feedback, Eq. 6) inducing permanent inhibition of the corresponding input neuron. PO neurons act as “decision cells”, triggering the inhibition, when the system has completed its learning phase. The activity of the PO cells tells if an event was learned or not. The potential of the \overline{ND} cells is computed as follows (only one PO cell fires at a time):

$$\frac{dPot_i^{\overline{ND}}}{dt} = -\alpha_{\overline{ND}} \cdot Pot_i^{\overline{ND}} + \sum_l W_{\overline{ND}(i)}^{PO(i,l)} \cdot Act_{i,l}^{PO} \quad (6)$$

$\alpha_{\overline{ND}}$ is the value of the recurrent connection (see Ap-

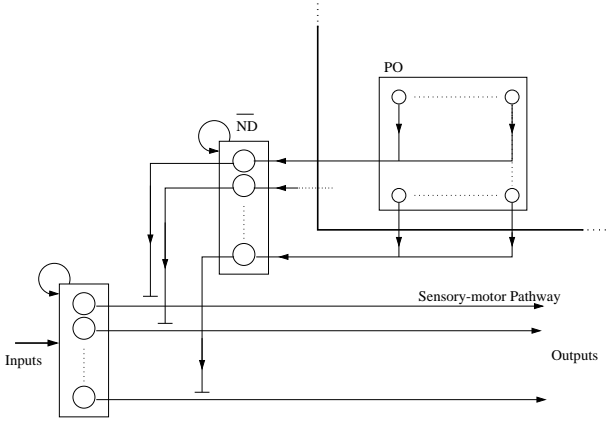


Fig. 9. Mechanism allowing to inhibit the perception-action pathway. The triggering information comes from the corresponding PO neurons, firing if the corresponding event is implicated in the production of a sequence. Recurrent connections on \overline{ND} neurons then maintain the inhibition on the inputs.

pendix B for numerical value).

$$Act_i^{\overline{ND}} = f_{\overline{ND}}(Pot_i^{\overline{ND}}) \quad (7)$$

with:

$$f_{\overline{ND}}(x(t)) = \begin{cases} 1 & \text{if } x(t) > T_{\overline{ND}} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$T_{\overline{ND}}$ is a threshold, which sets the number of times the system can produce a sequence before the activation of \overline{ND} cells and therefore the inhibition of the perception. The weights of the connections between \overline{ND} cells and the input neurons will affect the value of the inhibition. A complete inhibition ensures an independent production: the “transition learning and prediction” network drives the system. However it also cuts the system from the interaction. It is no longer possible to modulate the production according to the perceptions. So, the inhibition must only be partial (the weights of the connections between \overline{ND} and input neurons are less than 1): the input signal must not be strong enough to induce motor reaction by itself, it must be nevertheless present in order to modulate the timing of the motor production.

B. Modulation of the motor production timing

By “modulating” we mean changing the timing of the sequence, in order to obtain synchronous production. This capability to synchronize could be a basis for more complex sensory-motor interactions. In order to produce simultaneous actions, each system has to modulate its production speed in the manner of a phase locking loop mechanism (PLL). Accelerating or slowing the production is enough to obtain synchronization. We choose to enhance our architecture in order to provide our system the benefit of its perception to accelerate its production (slowing would have been much more complex to implement, because it needs to add a negative energy to the system). If the system predicts earlier the motor activations, it will be able to “know” in advance the next action to perform, and an incoming perception of this precise action will allow to trigger it earlier.

The following equation is used for an incremental learning of the TB-PO connections (instead of the “one-shoot learning” of equation 2):

$$\frac{dW_{PO(i,j)}^{TB(j,l)}}{dt} = 2\mu \cdot (Act_i^{TD} - Pot_{i,j}^{PO}) \cdot Act_{j,l}^{TB} \quad (9)$$

μ is the learning rate of the system. This equation (a classical conditioning rule [25]) induces a less accurate but earlier prediction of the motor event, if the learning process is stopped after 2 or 3 presentations of the sequence. The integration of the PO output is done by a group of neurons called IG (Fig. 10) according the following equations:

$$\frac{dPot_i^{IG}}{dt} = \alpha_{IG} \cdot Pot_i^{IG} + \sum_l W_{IG(i)}^{PO(i,l)} \cdot Act_{i,l}^{PO} - reset \quad (10)$$

$$reset = \sum_l W_{IG(i)}^{MO(i)} \cdot Act_{i,l}^{MO} \quad (11)$$

α_{IG} depends of the weight of the recurrent connection (see Appendix B for numerical value). $reset$ is a strong inhibition value allowing motor activation to reset the IG group. Pot_i^{IG} is always positive and increases since prediction is emitted. It is then sent to the Motor Output (MO) group of neurons.

$$Act_i^{IG}(t) = f_{IG}(Pot_i^{IG}(t)) \quad (12)$$

$$f_{IG}(x(t)) = \begin{cases} 1 & \text{if } x(t) > T_{IG} \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

A given Motor Output (MO) neuron triggers the activation

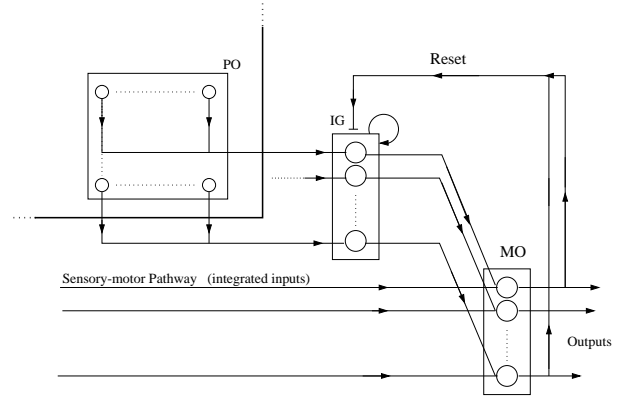


Fig. 10. The Motor Output (MO) group triggers the motor output when integrated prediction overshoots a given threshold. If an input information (also integrated) happens during the integration of prediction, the summing potential of MO reaches the threshold early: the system accelerates.

of a given action. The potential of MO neurons is computed as follows:

$$Pot_i^{MO} = \sum_j (W_{i,j}^{IG} \cdot Act_j^{IG} + W_{i,j}^{Per} \cdot Act_{i,j}^{Per}) \quad (14)$$

It simply sums two incoming informations:

- A signal from IG neurons, that can trigger activation, whenever it reaches a fixed threshold. Alone, the system is able to produce a learned sequence.

$$Act_{i,j}^{MO} = f_{MO}(Pot_{i,j}^{MO}) \quad (15)$$

$$f_{MO}(x(t)) = \begin{cases} 1 & \text{if } x(t) > T_{MO} \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

- The integrated (but partially inhibited) perception signal from Input neurons (Per), whose value isn't high enough to overshoot the threshold.

$$\frac{Pot_i^{Per}}{dt} = -\alpha_{Per} \cdot Pot_i^{Per} + \theta - \sum_l W_{Per(i)}^{ND(i)} \cdot Act_i^{ND} \quad (17)$$

α_{Per} is the value of the recurrent connection (the potentiation is decreasing after activation), θ is the binary activity of the other system's MO neuron (direct connection)¹.

$$Act_i^{Per} = [Pot_i^{Per}]^+ \quad (18)$$

In normal functioning, the PO prediction is integrated by the IG neurons. Activity of IG neurons increases due to self feedback and triggers the action when the threshold of MO neuron is exceeded (Fig. 11). If a perception event occurs, between the proposal of an action prediction and the triggering of the action, it will result in an acceleration of the system. Perception thus results in an earlier triggering of the chosen action (overshoot of the motor threshold). Finally the connections between IG and MO groups change according a conditioning rule (a LMS equation, [25]). It modulates the efficiency of IG activity on MO potential according to the right timing of the sequence during the learning phase:

$$\frac{dW_{MO(i,j)}^{IG(j,l)}}{dt} = 2\mu \cdot (Act_i^{Per} - Pot_{i,j}^{MO}) \cdot Act_{j,l}^{IG} \quad (19)$$

where μ is the learning rate of the system. Between two systems producing the same sequence, the effect of connecting action to perception induces a step-by-step adjustment of the sequence production until synchronization is achieved.

C. Results of simulated experiments

Simulations were performed according to two experimental protocols:

1. A first series of experiments tested the synchronization of our architecture with a simple generator of time fixed sequences (Fig. 12, upper). Once the sequence is learned, we tested the system's production simultaneously with the generator. The system was triggered at random instants, while the generator produced its sequence recurrently.
2. In a second series of experiments, two copies of our architecture were used (Fig. 12, lower). Both had learned the same sequence (from a fixed sequence generator). Both systems were switched at random instants. Each architecture was a separate application, designed and simulated using Leto and Prometheus software involving PVM².

¹ $[x]^+ = x$ if $x > 0$, 0 otherwise

²The Parallel Virtual Machine software creates a virtual machine from a set of computers.

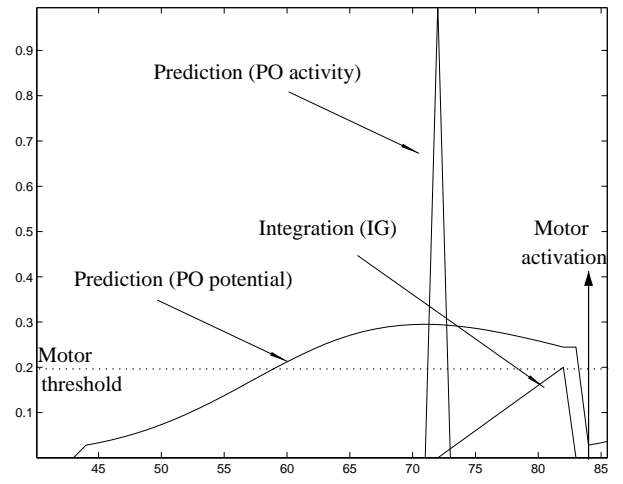


Fig. 11. Activity of PO and IG cells during the production of a sequence: PO cell fires for a transition. The maximum of its potential generates an activation spike. This spike will be integrated (IG), until the MO neurons reach the threshold, triggering the motor activation.

For both experiments, synchronization is an attractor of the interactions. The conditions and speed of convergence toward this attractor are directly dependent of the detailed equations. Synchronization is here an example of a stable state, which can be obtained by controlling parameters and equations of the systems forming parts of the whole dynamic.

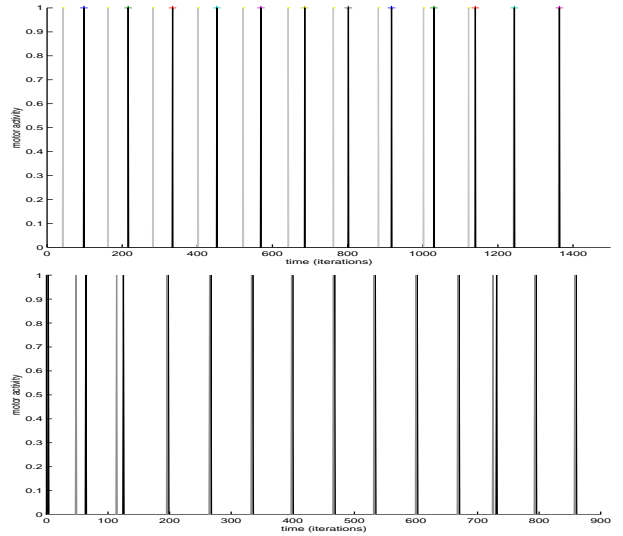


Fig. 12. Interaction between two systems: Upper panel graph shows the synchronization phase of our system (grey) on a fixed sequence (black). Lower panel graph shows the very quick synchronization between two identical systems (grey and black). A small perturbation occurred near iteration 730, due to the temporary loss of messages between stations (via PVM). Perfect synchronization is quickly recovered.

In the following section, we will use the fundamental properties of interactive systems such as synchronization to build a non explicit reward signal from the rhythm of the interaction. We will show this network can easily be derived from the “transition learning and prediction” neural

network used in the previous experiment.

V. LEARNING NEW SENSORY-MOTOR ASSOCIATIONS WITHOUT EXPLICIT REINFORCEMENT

A. Description of the problem

A simplified version of our neural network can be adapted to the prediction of the rhythm exchange in a student-teacher interaction. The current application concerns rhythm prediction, used as an internal reinforcement signal in order to permit autonomous learning. In a simple association problem (Fig. 13), sensory input has to be associated with a given motor response. Here, we consider the simplest application, where each input and output is only coded by a single neuron. As shown in figure 13, we have 3 input neurons and 3 output neurons. The input neurons could represent the detection of a perceived movement, and the output neurons the execution of possible motor responses. Via interactions, the system has to learn to propose the correct (*i.e.* similar) response according to the stimulus input. For example, if input 1 is activated, the system has to produce the output A. If input 2 is presented, then output B should be the response, and so on.

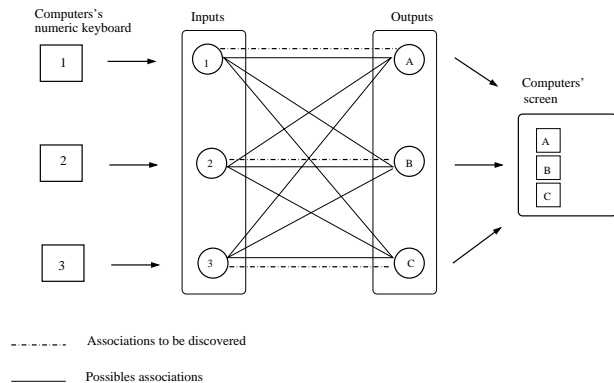


Fig. 13. A simple associative network with one to all connections. The human teacher activates the input neurons by pressing the computer’s numerical keys. The network proposes “motor” responses, displayed on the computer’s screen. Bold links are those to be learned during the game (the rules to be discovered by the system).

The task can be viewed as a game: on one side, the human participant (*i.e.* the teacher) expects the associations to be performed. By activating one of the keys of the numerical keyboard of a computer, he directly activates the correspondent input of the network (we call this action a “key pression” or “key activation”). On the other side, the system, detecting an activated input, has to choose and associate it to one of the three possible responses. During the game, there is no way for the teacher to give the learner any explicit reward (no kind of meaningful “explicit key”). The system will have to discover by itself the reinforcement information in the data flow, which is only a temporal series of inputs. In the following subsections we show that learning rules based on a simple correlation mechanism [8] is not enough. However, the prediction mechanism described previously can be used to build an internal reward efficient

enough to control a simple reinforcement learning mechanism.

B. Can a simple correlation rule do the task ?

A working hypothesis can be the use of a Hebbian learning rule. Associations between correlated active input and output could be learned according to the following rule:

$$\frac{dW}{dt} = -\lambda W - \lambda' WI + \mu SI \quad (20)$$

where: λ is a term of constant passive decay, λ' is a term of active decay (decrease of the associations between the activated input and the non-activated outputs-forgetting), μ is the learning rate ($\mu \ll \lambda'$). I stands for the activity of the input neuron ($I \in [0, 1]$), S the value of the output neuron ($S \in [0, 1]$), W the value of the weight between I and S (simulations where made with weights where initialized at random low values).

Although this rule permits the learning of sensory-motor correlations, it constrains the experiment on one main point: here the teacher has to adapt his way of playing to the system propositions to permit correct correlations to be learned. The teacher has to wait, drive and teach the learner (see previous similar works [6], [7]). Emphasizing non-explicit reinforcement, we consider that this rule generates too much asymmetry between the teacher and the learner. Indeed, with this rule, the “game” turns into a dynamics where: the system proposes a motor output, and then, the teacher answers with the right input, thus reinforcing the weight of the corresponding link, and weakening the weight of the non-correlated input (active forgetting). Instead of this explicit and immediate reinforcement (the correlation) depending on the teacher, we think that our system could build a non explicit reward information from the dynamics of the interaction while testing different associations. Then, it could learn the correct associations. We use a PCR learning rule (Probabilistic Conditioning Rule) [13], which has been experimented on delayed conditional learning, where the system had to test and evaluate the credibility of possible rules, before learning them. This rule computes a credibility value associated to each connection. The credibility is used to switch the weight according to a reward value (details of the algorithm in Appendix). If we want the reward to be non-explicit, we should rather see the experiment as: the teacher is producing a motor output (*i.e.* the system input) and expects responses from the other part to continue the interaction process. The system responds and then evaluates the temporal regularity in the teacher’s behavior. If the rhythm has changed, there is a good probability that the associated input-output weights of the network have changed. To go further, the teacher will be considered as another player whose reactions change during the game according to how quickly the system learns. If correct associations are learned, the teacher tends to continue to play at nearly the same rhythm, but if a wrong response occurs, there is a strong probability that the teacher makes a break to “say something” or express his disagreement. He may also

change the rhythm (as an extreme situation, the teacher could give up the game). In such a game the rhythm of the key activation is the only information that can be used to adjust the learning. If the student is able to correctly predict the rhythm, there is no need for it to understand the meaning of the words pronounced when a disagreement occurs. A break in the rhythm is sufficient to provide an error signal. This model only requires a non-verbal context and low-level cognitive capabilities.

C. Predicting the rhythm as a reward signal

If the rhythm prediction is correct, the system can maintain its learning of the current presented association. If the prediction is wrong, it may set the association as non-credible, and decrease the corresponding weight.

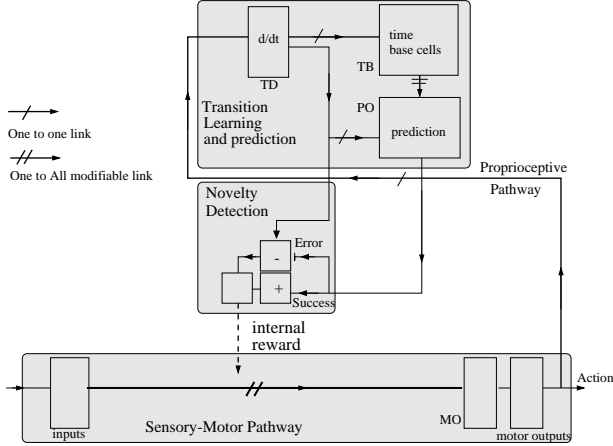


Fig. 14. Presentation of the system used to learn and predict delays between the teacher’s game. The result of the comparison between the prediction (PO cell activity) and the effective input (derived in TD) is directly connected to the reward evaluation modulating the learning of associative connections between inputs and outputs.

In order to maintain the rhythm prediction, we again use the “transition learning and prediction” system depicted in Fig. 4. As shown in Fig. 14 the three components TD, TB, PO are used to learn the time delay between two keys activations. TD “pulses” each time a key is pressed, initializing the activity of TB. During the interval between two key activations, TB cells are firing according to Eq. 1, as shown in Fig. 5. When the next key is pressed, PO is activated, the connections (between PO and TB) corresponding to the currently activated cell of TB are reinforced (Eq. 2) and the delay is learned. The system is then able to predict the instant of the next key pressure. We use the prediction information of PO to do so: the analogical activity of a single PO cell (Eq. 4), and the binary prediction information (Eq. 5). Figure 15 shows how the system maintains the two information during an experiment with a human teacher.

The activity of the PO cell is then summed with the output of TD, to compare the prediction with the real event. If we take the PO activity (Fig. 15, bottom) as an inhibitory activity summed to the event, we directly obtain an error message, which is proportional to the prediction mismatch:

Fig. 15. Example of a 20 second experimental scenario. Upper graph: Each pulse indicates when the teacher presses a key (only the information of time is given). A constant rhythm is maintained for the first five strikes (time period = A), then is broken (according to a bad system’s response for example). Middle graph: Pulses show the system’s binary prediction of the teacher’s play. Each pulse happened just before the teacher’s key activations. The star means a prediction error due to the break. Bottom graph: activity of the PO cell. Maxima of activity trigger pulses of prediction. After the rhythm break, the new delay, B, is learned.

$$Error = [Act^{TD} - Pot^{PO}]^+ \quad (21)$$

By multiplying the prediction information by the TD output, we obtain a binary success value:

$$Success = Act^{TD} \cdot Act^{PO} \quad (22)$$

where Act^{TD} Act^{PO} are binary values $\in 0, 1$ (a value of 1 means the matching between the system’s prediction and the teacher’s rhythm). Error and success information are then compared to modulate directly the learning of the associative network:

$$Reward = \begin{cases} r & \text{if } Success > \varepsilon_1 \\ -error \times r & \text{if } Error > \varepsilon_2 \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

where r is a high value, enough to induce an instantaneous weight modification (see appendix B for numerical values). Figure 16 shows a complete trial, with the input train of the human key activations. For each activation, the system gives a response (randomly at start). Then, it learns from the other’s behavior. Interruptions in the human actions are visible, and the middle graph shows the success/error information extracted by the predicting system.

Experiments lasted about 4 minutes. The network always succeeds in learning the three correct associations. Most of the time, an experiment starts with wrong associations due to the random weights initialized on the network

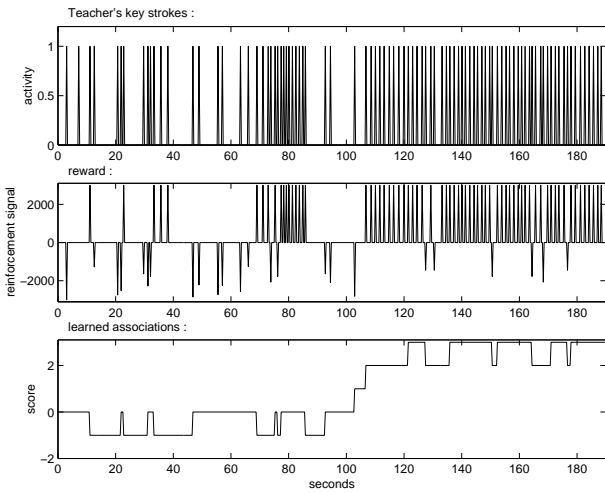


Fig. 16. The learning process during a whole experiment (about 200 seconds). Upper panel graph: action frequency of the human player. Middle panel graph: variation of the reinforcement signal according to the rhythm prediction. Positives values are caused by a successful prediction of the time interval between two activations, while negative ones are due to prediction error. These reinforcement variations act on the update of the associative weight. Lower panel graph: Evaluation of the learning level. A well learned association gives 1 point, an unlearned one gives 0 point while a bad learned one loses 1 point.

links. This explains the player’s numerous breaks (Fig. 16, top). When the first association is discovered, the learning of the others is facilitated by the decreasing number of possible associations: in this non-verbal communication perspective, a dynamic learning is possible. Here, the system is able to detect novelty in the teacher’s behavior, and to use this novelty as an implicit reward signal for learning. Novelty is defined as a non-expected break in the rhythm, perceived at the system’s inputs. This could help to “locate” important events or, to learn properties on objects or surrounding entities.

VI. CONCLUSION

The main contribution of this paper is a novel approach for building autonomous systems able to imitate actions of increasing complexity, and to distinguish between imitation of novel and previously learned actions. Our approach challenges the classical view, which focuses on how a robot can imitate and capture the meaning of a movement, and which delineates the learning phases by specific signals. To achieve our goals, we postulate the imitation capabilities of our system are based on an homeostatic principle. The variations of the perceptive flow are interpreted as a signal of error. In an effort to reduce this signal, the system produces an imitative behavior. We believe that this simple principle can be useful to build systems that are really able to process in an autonomous way different levels of imitation, such as following a path, reproducing the movement of an arm, or imitating complex actions like opening a door, cooking, or building a complex machinery. When the system has to perform the immediate imitation of an already learned behavior, uncoupling perception and ac-

tion appeared to be decisive in order to prevent perceptive interferences. Although the current assumption underlying robotic works on imitation is that distinctive capabilities should characterize teacher and imitator systems, our model shows the interest to implement systems with similar basic imitative capabilities. Indeed, if both simulated systems come to perform a similar action, this induces a phase lock on this action, which will facilitate further introduction of turn-taking and learning from each other.

In this work, the collaboration with a psychologist has been crucial, since we have tried to understand some of the fundamental properties of young infant involved in imitation games and interaction situations. We have discovered that immediate imitation is much more complex than previously thought and that its double role in learning and communication was of a very high importance for robot learning. This collaborative work inspired us the idea that an autonomous system could generate by itself its own reward signal for learning. It also inspired us the exploitation of the temporal properties (prediction, synchronization) of our model for interaction purposes. Of course, the simulations presented in this paper are quite simple, in comparison with the complex cognitives faculties of the young infants. But we think that basic skills such as synchronization, rhythm detection are a possible step toward the understanding of social capabilities (it could be the kind of information babies use to detect the correspondence of their behavior with their care-givers).

Moreover we can notice our Neural Network architecture is based on previous works on hippocampus modeling [5]. In these models, we claim hippocampus is not just a working memory or at the opposite a cognitive map [23] but a structure where transitions between states or places are stored and predicted. This model could explain the role of the hippocampus during conditioning learning [27] and the recent puzzling results in spatial navigation [29]. And of course, it provides the fundamental elements we need for learning by imitation and learning by interaction. Note that, this structure is only necessary in the learning phase and not in a simple imitation or reproduction phase (in a more detailed model the definitive learning of the sequences should be performed in other structures). Finally, we are interested in the idea that imitation capabilities could be progressively build-up from very simple sensory-motor schemes, since it would promote important advances in man-machine interface and robot learning. Moreover, a robot could be considered as a good heuristic tool to propose new behavioral therapies given the present imitative capabilities in children with autism, who are supposed to face problems with human models.

Our future works will focus on real size robotic experiments using mobile robots with mechanical arms. We will have to understand how to add to our neural model structures allowing to learn categories of actions at the program level. We hope the developed architecture will be able to exhibit different phases of developments that we will be able to compare with babies development. Hence, we will perhaps be able to help in the understanding of mental

development problems like children with autism: is autism linked to a problem of theory of mind, to the sensory-motor level, to the management of novelty detection, or even to the capability to mobilize or express internal states? This kind of questioning is perhaps unfamiliar for engineers since we know we are really far away from building non autistic robots. But, it is clear that all the new results in psychology and neuro-imagery will be of high interest to improve current robot controllers. At the same time, robotics experiments appear more and more as a new way to perform synthetic simulations of psychological and neurobiological models and are promised to an important development in the field of cognitive sciences.

VII. ACKNOWLEDGMENTS

This work is part of the French program ‘‘Cognitic’’ (Action Concertee Incitative Cognitique, COG 156), involving neuro-imaging, modeling, behavioral, and clinical studies. The author gratefully acknowledge J.L. Contreras-Vidal for improving the English of the paper, Guido Bugman, Arnaud Revel, Mathias Quoy, Inbar Fijalkow, and Kerstin Dautenhahn for their interesting comment about this work.

APPENDIX

I. THE PCR ALGORITHM

The Probabilistic Conditioning Rule is useful for conditioned learning with a delayed reward (a complete description of the algorithm can be found in [13]). Binary weights are associated with a probability p_{ij} , which measures the certainty of the weight W_{ij} . If $Rnd > p_{ij}$ and $\bar{I} \cdot \bar{O} \neq 0$:

$$\begin{aligned} W_{ij} &= 1 - W_{ij} \\ p_{ij} &= 1 - p_{ij} \end{aligned} \quad (24)$$

When the *reward* signal varies, probabilities are updated and weights can be modified according to the certainty value.

If $\left| \frac{dReward}{dt} \right| > \xi$:

$$\frac{dp_{ij}}{dt} = \alpha_{pcr} \cdot \frac{dReward}{dt} \cdot C_{ij} \cdot ((2 \cdot W_{ij}) - 1) \quad (25)$$

If there is no reinforcement variation neither the probabilities nor the weights are modified but information about the correlation between the input and the output of the weight go on being stored.

$$\overline{X_j}(t+1) = \frac{\tau \overline{X_j}(t) + X_j(t)}{\tau + 1} \quad (26)$$

$\overline{I_{ij}}, \overline{O_{ij}}, \overline{IO_{ij}}$ are updated with Eq. 26 and $C_{i,j} = \frac{\overline{IO_{ij}}}{\sqrt{\overline{I_i} \overline{O_j}}}$. α_{pcr} is the delayed conditioning learning rate, ξ is a constant fixed by the experimenter and Rnd is a random value in $[0, 1]$.

II. EXPERIMENTAL PARAMETERS

The μ parameter is the global learning rate used in the simulator: $\mu = 0.2$

Parameters for the section IV simulations:

$$\alpha_{\overline{ND}} = 0.05, T_{\overline{ND}} = 0.6$$

$$\alpha_{IG} = 0.05, T_{IG} = 0.01 W_{IG}^{PO(i,l)} = -100$$

$$\alpha_{Per} = 0.2, W_{Per}^{\overline{ND}(i)} = 0.8$$

$$T_{MO} = 0.9,$$

Parameters for the section V simulations:

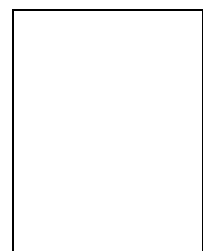
$$\alpha_{pcr} = 0.3, \xi = 0.4$$

$$\lambda = 0.01, \varepsilon_1 = 0.2, \varepsilon_2 = 0.9, r = 2000$$

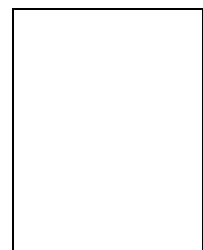
REFERENCES

- [1] P. Andry, S. Moga, P. Gaussier, A. Revel, and J. Nadel. Imitation : learning and communication. In J. A. Meyer, A. Berthoz, D. Floreano, H. Roitblat, and S. Wilson, editors, *Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior SAB'2000*, pages 353–362, Paris, 2000. The MIT Press.
- [2] W. R. Ashby. *Design for a brain*. London: Chapman and Hall, 1960.
- [3] A. Bandura. *Psychological modeling: conflicting theories*. Chicago: Aldine-Atherton, 1971.
- [4] J.P. Banquet, P. Gaussier, J.L. Contreras-Vidal, and Y. Burnod. The cortical-hippocampal system as a multirange temporal processor: A neural model. In R. Park and D. Levin, editors, *Fundamentals of neural network modeling for neuropsychologists*, Boston, 1998. MIT Press.
- [5] J.P. Banquet, P. Gaussier, J.C. Dreher, C. Joulain, and A. Revel. Space-time, order and hierarchy in fronto-hippocampal system : A neural basis of personality. In G. Matthews, editor, *Cognitive Science Perspectives on Personality and Emotion*, pages 123–189. Elsevier Science BV Amsterdam, 1997.
- [6] A. Billard, K. Dautenhahn, and G. Hayes. Experiments on human-robot communication with robots, an imitative learning and communicating robot. *Proceedings of ‘‘Socially Situated intelligence’’ Workshop, part of the Fifth International Conference on Simulation of Adaptive Behavior 98, SAB 98*, August 1998.
- [7] A. Billard and G. Hayes. Transmitting communication skills through imitation in autonomous robots. *Proceedings of Sixth European Workshop on Learning Robots, EWL97*, July 1997.
- [8] A. Billard and G. Hayes. Learning to communicate through imitation in autonomous robots. In *Proceedings of 7th International Conference on Artificial Neural Networks, ICANN97*, pages 763–768, October 1997.
- [9] G. Butterworth. Neonatal imitation: existence, mechanisms and motives. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 63–88. Cambridge: Cambridge University Press, 1999.
- [10] G. Cheng and Y. Kuniyoshi. Complex continuous meaningful humanoid interaction: A multi sensory-cue based approach. In *Proceedings of IEEE International Conference on Robotics and Automation (ICRA 2000)*, pages 2235–2242, April 2000.
- [11] K. Dautenhahn. Getting to know each other - artificial social intelligence for autonomous robots. *Robotics and Autonomous System*, 16(2-4):333–356, December 1995.
- [12] P. Gaussier, S. Moga, J.P. Banquet, and M. Quoy. From perception-action loops to imitation processes: A bottom-up approach of learning by imitation. *Applied Artificial Intelligence*, 7-8(12):701–727, 1998.
- [13] P. Gaussier, A. Revel, C. Joulain, and S. Zrehen. Living in a partially structured environment: How to bypass the limitation of classical reinforcement techniques. *Robotics and Autonomous Systems*, 20:225–250, 1997.
- [14] P. Guillaume. *L’imitation chez l’enfant*. Paris: Alcan., 1925.
- [15] G. Kugiumutzakis. Genesis and development of early infant mimesis to facial and vocal models. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 36–59. Cambridge: Cambridge University Press, 1999.
- [16] O. Maratos. The origin and development of imitation in the first sixth month of life. Paper presented at the British Psychological Society Annual Meeting, Liverpool, April 1973.
- [17] A. Meltzoff and M. K. Moore. Imitation of facial and manual gestures by humans neonates. *Science*, 198:75–82, 1977.
- [18] A. Meltzoff and M. K. Moore. Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6, 1997.

- [19] A. Meltzoff and M. K. Moore. Persons and representation: why infants imitation is important for theories of human development. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 9–35. Cambridge: Cambridge University Press, 1999.
- [20] J. Nadel. The functional use of imitation in preverbal infants and nonverbal children with autism. In A. Meltzoff and W. Prinz, editors, (*in press*). *The Imitative Mind: Development, Evolution and Brain Bases*. Cambridge: Cambridge University Press, 2000.
- [21] J. Nadel, C. Guerini, A. Peze, and C. Rivet. The evolving nature of imitation as a format for communication. In J. Nadel and G. Butterworth, editors, *Imitation in Infancy*, pages 209–234. Cambridge: Cambridge University Press, 1999.
- [22] C. L. Nehaniv and K. Dautenhahn. Mapping between dissimilar bodies: Affordances and the algebraic foundations of imitation. In J. Demiris and A. Birk, editors, *Proceedings of Seventh European Workshop on Learning Robots 1998 EWLR98*, pages 64–72. World Scientific Press, July 1998.
- [23] J. O’Keefe. The hippocampal cognitive map and navigational strategies. In J. Paillard, editor, *Brain and Space*, pages 273–295. Oxford University Press, 1991.
- [24] J. Piaget. *La formation du symbole chez l’enfant*. Delachaux et Niestle Editions, Geneve-Paris, 1945. English translation: Play, Dreams and imitation in childhood (1952).
- [25] R. A. Rescola and A. R. Wagner. *Classical Conditioning II : Current Research and Theory*, chapter A Theory of Pavlovian Conditioning: Variations in the effectiveness of reinforcement and non reinforcement. Appleton-Century-Crofts, New York, 1972.
- [26] G. Rizzolatti. From mirror neurons to imitation: Facts and speculations. In A. Meltzoff and W. Prinz, editors, (*in press*). *The Imitative Mind: Development, Evolution and Brain Bases*. Cambridge: Cambridge University Press, 2000.
- [27] R.F. Thompson. Neural mechanism of classical conditioning in mammals. *Phil. Trans. R. Soc. Lond. B*, 329:161–170, 1990.
- [28] C. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8(3):279–292, 1992.
- [29] I.Q. Whishaw and L.E. Jarrard. Evidence for extrahippocampal involvement in place learning and hippocampal involvement in path integration. *Hippocampus*, 6:513–524, 1996.
- [30] A. Whiten and R. Ham. On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. In P.J.B. Slater, J.S. Rosenblatt, C. Beer, and M. Milinski, editors, *Advances in the study of behavior*, pages 239–283, San Diego, CA, 1992. Academic Press.

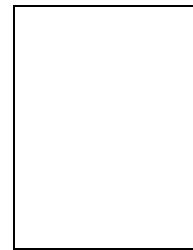


Pierre Andry is born in 1975 in Annecy, France. He received a M.S degree in Artificial Intelligence from the of Paris VI University (Pierre et Marie Curie) in 1999. He is actually a Ph.D. student in ETIS lab, from the Cergy Pontoise University and the ENSEA. He’s main interests concern imitation and interaction processes in the field of autonomous systems.

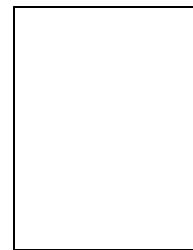


Philippe Gaussier Philippe Gaussier received the M.S. degree in electronic from Aix-Marseille University in 1989. In 1992, he received a Ph.D. degree in computer science from the University of Paris XI (Orsay) for a work on the modeling and simulation of a visual system inspired by mammals vision. From 1992 to 1994, he conducted research in Neural Network (NN) applications and in control of autonomous mobile robots at the Swiss Federal Institute of Technology. He has edited a special issue of Robotics and Autonomous System Journal on "moving the frontier between robotics and biology". He is now Professor at the Cergy-Pontoise University in France and leads the neurocybernetic team of the Image and Signal processing Lab (ETIS). His research interests are focused on one hand on the modeling of the cognitive

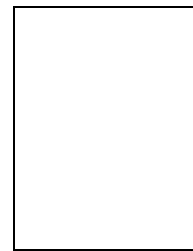
mechanisms involved in visual perception, motivated navigation and action selection, and on the other hand on the study of the dynamical interactions between individuals (imitation capabilities, collective intelligence, social interactions...).



Sorin Moga was born on December 14, 1971 in Targu-Mures, Romania. He received the Engineering degree in electronics and telecommunication from the Universitatea Politehnica, Timisoara, Romania in 1995 and a master degree in image and signal processing from the University of Cergy Pontoise in 1996. In 2000, he received a Ph.D from the University of Cergy Pontoise, France. His research interests focus on the robotics and the learning by imitation in robotic fields.



Jean Paul Banquet holds a MD with a residency in Neuropsychiatry and a Ph.D. in Applied Mathematics from Paris VI, 1981. During two three-year periods, he was appointed Full-bright Research Fellow at Stanley Cobb Laboratories for Neurophysiological Research (Harvard Medical School), and at the center for Cognitive and Neural Systems chaired by Professor Grossberg at Boston University. After doing electrophysiological research on memory and learning, he is presently working at an INSERM unit: Neuroscience et Modelisation at Pierre et Marie Curie, where he conducts research on neural network modelling of the hippocampus, basal ganglia, and their relation with the cortex.



Jacqueline Nadel is a PhD researcher in the area of developmental psychology and psychopathology. Her main contributions concern the functional use of imitation and detection of imitation in young children and children with autism. She is the co-editor of *Imitation in infancy*, 1999, Cambridge University Press, the first book to bring together the extensive modern evidence for innate imitation in babies.

mechanisms involved in visual perception, motivated navigation and action selection, and on the other hand on the study of the dynamical interactions between individuals (imitation capabilities, collective intelligence, social interactions...).