From Perception-Action loops to imitation processes
A bottom-up approach of learning by imitation

P. Gaussier, S. Moga, J.P. Banquet (*), M. Quoy
Groupe Neurocybernetique, ETIS / URA CNRS 2235,
6, Avenue du Ponceau, 95014 Cergy-Pontoise, France.
(*) CREARE, Neurosciences et modélisations,
9, quai St. Bernard, 75252, Paris cedex 13, France
e-mail: gaussier or moga@ensea.fr

Running head : From PerAc loops to imitation processes

Mailing address:
Sorin MOGA
Groupe Neurocybernetique ETIS / URA CNRS 2235
6, Avenue du Ponceau, 95014 Cergy-Pontoise cedex
France

To appear in "Applied Artificial Intelligence: An International Journal", no 7, vol 1

Abstract

This paper [1] proposes a neural architecture for a robot in order to learn how to imitate a sequence of movements performed by another robot or by a human. The main idea is that the imitation process does not need to be given to the system but can emerge from a mis-interpretation of the perceived situation at the level of a simple sensory-motor system. The robot controller is based on a PerAc (Perception-Action) architecture. This architecture allows an autonomous robot to learn by itself sensory-motor associations with a delayed reward. Here, we show how the same architecture can also be used by a "student" robot to learn to imitate another robot allowing the student robot to discover by itself solutions to a particular problem or to learn from another robot what to do. We discuss the difficulty linked to the segmentation of the actions to imitate. This imitation problem is demonstrated by a task of learning a sequence of movements and their precise timing. Another interesting aspect of this work is that the neural network used for sequence learning is directly inspired from a brain structure named the hippocampus and mainly involved in memory processes (Banquet *et al.* , 1997). We discuss the importance of imitation processes for the understanding of our high level cognitive abilities linked to self-recognition and to the recognition of the other as something similar to me.

1 Introduction

Till now, our work has been mainly focused on the design of a neural network architecture (named PerAc: Perception-Action) for the control of a visually guided autonomous robot. Our architecture is directly inspired by neurobiological models and tries to take into account animal behavior. Using this architecture, we have shown that a robot can learn to isolate a particular "object" in a visual scene and to associate this object with a motor behavior: reaching, avoiding by the left or the right... (Gaussier *et al.* , 1997a). We have further developed a probabilistic conditioning rule which allows our robot to learn sensory-motor associations according to a delayed reward. This algorithm has been successfully tested on real maze problems (Gaussier *et al.* , 1997c). During navigation in an open environment, we have shown that a robot can learn how to reach any position with a high precision only by using landmarks in the visual scene. However, "discovery" of interesting locations or "discovery" of the correct set of sensory-motor associations for a particular task is an NP complex problem. The training time becomes quickly too large when the size of the problem increases. A solution used to reduce complexity was proposed in our previous papers (Gaussier *et al.* , 1997d; Gaussier *et al.* , 1997a). It consists in keeping the same code from the sensory input to the motor output (egocentric coordinates) so as to reduce the quantity of data the system has to learn. With our PerAc architecture, the robot can "take for free" the environment continuity properties that are lost in symbolic systems. The addition of planning capabilities is also required to allow latent learning and to quickly find the shortest pathway solution to a particular problem (Revel *et al.* , 1998). However, the PerAc architecture does not help to solve problems which have an intrinsically high dimension.

Box 1: The PerAc architecture

The PerAc (Perception-Action) block has been proposed as an elementary generic brick of neuronal computation (Albus, 1991; Brooks, 1986; Hecht-Nielsen, 1987). Basically, it allows on-line learning of sensory-motor associations. A PerAc block is divided into two levels corresponding to the action and the perception data flows. The first level is a reflex mechanism which extracts basic information from the perceived input so as to directly and roughly control the actions. The second level performs situation recognition and allows learning of the associations between what is recognized in the perceptive flow and the chosen movement. This level permits to maintain the behavior provided by the reflex system or to avoid it when there are contradictions with the robot viability constraints. Thus direct "one to one" links between the reflex action proposal group and the action selection group allow making movements in the direction of features already known as "relevant". A PerAc block is a competitive network composed of 4 neural boxes as seen in Fig. 1.

Therefore imitation of already learned behaviors or subparts of a behavior not completely discovered is certainly one way to allow a population of animals or robots to learn and to find solutions by themselves. Indeed, imitation is often a fast means of learning in contrast to trial and error strategy (Galef, 1990). The autonomy of the robot requires that there is no intrusion in the robot brain during learning. The robot must be able to select which informations to store. Hence, the problem becomes how to initiate the communication mechanism between

[1] This paper is based on the working notes of the SIA symposium (Gaussier *et al.* , 1997b).

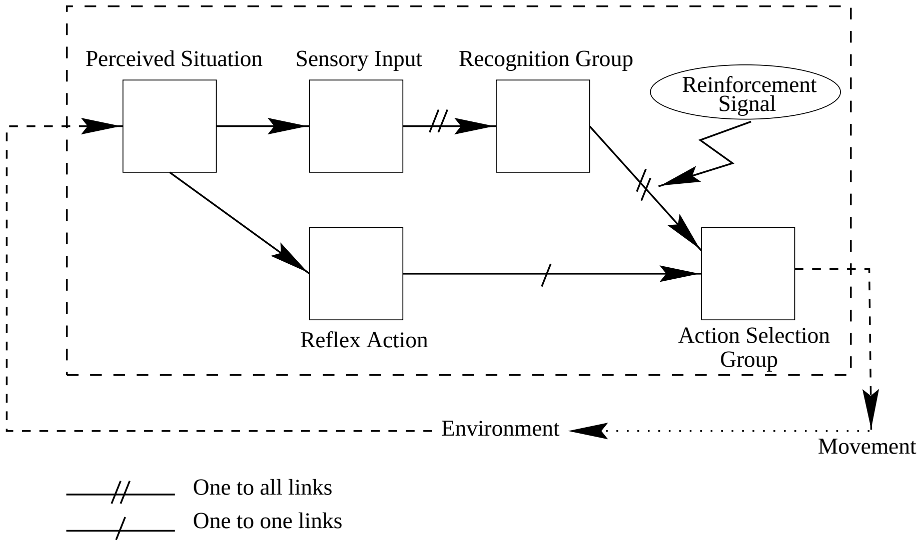


Figure 1: Schematic representation of the PerAc block. From the perceived situation, the reflex system extracts information to control directly the actions. Concurrently, the recognition system learns sensory input patterns and how to link them to the action by associative or reinforcement learning. The system adapts itself dynamically to the environment.

the student robot and the teacher. Imitation is also a good starting point to allow human-robot interactions. Learning by imitation is already used in a few projects of Artificial Intelligence (AI). In (Dautenhahn, 1995), heterogeneous mobile robots go by a "hilly landscape" attaching themselves to the teacher robot and imitating its trajectory. In (Hayes & Demiris, 1994; Demiris & Hayes, 1996) a robot agent follows a teacher through a maze, detecting significant teacher actions and associating them to the environment. In (Kuniyoshi, 1994; Berthouze *et al.* , 1996) a robot agent watches a human performing a simple assembly task and it reproduces the sequence of actions. A short review of these works is presented in (Bakker & Kuniyoshi, 1996). By contrast to other works in this area (Berthouze *et al.* , 1996; Matarić, 1995; Hayes & Demiris, 1994), our main concern is to create a generic neural network architecture that allows on-line and unsupervised learning by imitation mechanisms. Another constraint is that the student architecture must be the same as the teacher N.N. architecture (a teacher can become a student at anytime). Indeed, the long term goal, is to create robots able to discover solutions by them self and/or to imitate other robots to complete their own learned behavioral repertoire. The constraints and ideas used to design our N.N. architectures are inspired by biological and psychological models of mammal brains. We hope to validate those models or to propose new ones and at the same time to propose a new robot controller architectures.

In this paper , we propose a neural architecture for imitation based on visual information and we show how to use it to teach the robot to perform a particular sequence of movements (to make a zigzag trajectory, a square ...). First, we provide a brief summary of our previous work on robot collective behavior for clustering tasks and we show how we can deduce from these results the scope of our imitation process. Second, we present a tracking mechanism used as a bootstrap for our imitation task. The shortcomings of the first tested solutions are pointed out. An active filtering mechanism to allow the "teacher" robot movement segmentation is proposed. Finally, a robust N.N. architecture to learn temporal sequences is explained and applied to our robot imitation task. At last, we discuss our viewpoint on the imitation problem and present the future developments of our project.

2 Emergent properties of robot/environment interactions and imitation

In (Gaussier & Zrehen, 1994), we realized a robotic experiment inspired by Deneubourg's work on sorting and clustering tasks performed by ants (Deneubourg *et al.* , 1990). As in Deneubourg's model, our robots did not take into account the interactions between the agents. All the robots acted as if they were alone. Nevertheless, lumping pieces of wood together created obstacles and divided the environment in areas separating the robots. The robots then became specialized in the clustering of a particular area. The most interesting point in this work is certainly not the group behavior since with more than three (or four) robots the achievement of the task took longer than with two robots! But these robots were not explicitly programmed to build large clusters. Indeed, the only robot instructions were: to avoid obstacles like in Braittenberg vehicles; to take an object if the robot was not already moving another object; and to leave an object beside the other object in the opposite case. With this kind of instructions the robot can build stacks of 2 objects but there is no reason to build stacks with more than 2 objects. Indeed there is the same probability to take an object from a stack or to put it in another stack. In average the size of the stacks should not vary. However in the real experiment (see Fig. 2) stacks are created which implies a bias in the probabilities of taking and putting down an object. That bias is linked to a perception problem. To separate obstacles from objects, a decision on the number of Infra Red sensors saturated on our Khepera robots was used. The consequence of putting an object nearby another object is that they appear as an obstacle and not as two objects. Their significance for the robot has changed. The intrusion of this Gestalt effect (the whole is bigger than the sum of its parts) has raised for us the question of how robots can learn by themselves to take into account those emergent properties of the robot/environment interactions and how to use them in learning by imitation.

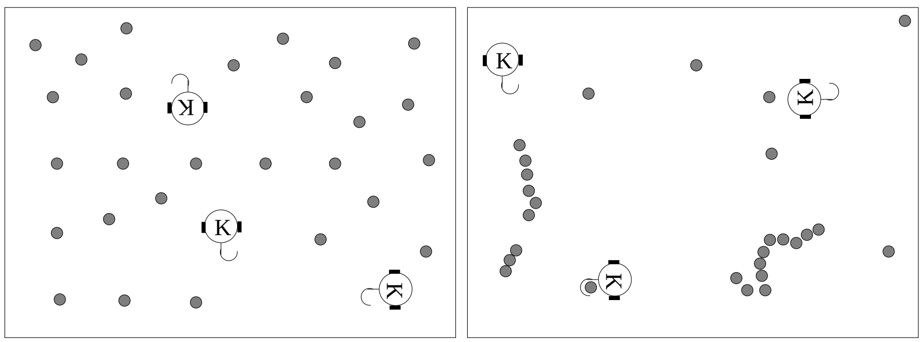


Figure 2: Overview of 'ant-like' clustering and sorting experiment. Left: in the initial position, the pieces of wood are uniformly distributed on the environment. Right: after a time, the robots builds stacks with more that 2 objects.

For the imitation behavior, we start with the same kind of assumption as for the previous clustering problem: imitation is triggered by a perception error. For instance, an imitation behavior between two robotic arms controlled by vision could be explained as follows: a robot arm learns the visuo-motor coordination between its camera and its hand. It creates a correspondence between a given hand position in the visual scene and the angular positions of the different joins. Then, if the robot looks somewhere else and sees (see Fig. 3) another arm in its visual field, it will perform the same movement as the second arm because it will try to reduce the differences between the representations it supposes to have of its arm (visual and motor representations). Finally, if the arm movements induce a reduction of an internal drive (associated to the satisfaction of a particular motivation), a positive reinforcement is triggered. The movement sequence will be stored and associated with the internal drive. Later, if the value associated with the internal drive changes too much from its optimal value, the sequence of movements will be triggered and an observer will consider the student robot has learned by imitation the behavior of the teacher robot.

If we return to our mobile robot imitation problem, it is complex to imagine a program that allows a robot to learn to visualize what another robot is doing (it is something that primates and perhaps other mammals succeed in doing (Heyes *et al.* , 1992) or not (Tomasello, 1990), but in a first place, we would like to see what kind of imitation mechanism could be performed on a robot that cannot have a complex internal representation of the world). However, it is simple to allow the student robot to follow (or track) the teacher (human or robot) as a way to avoid a difference between perception and action. Like in the robot arm example, the student tries to reduce the difference of speed between the information of the visual flow and the information about the motor wheel speed - homeostasis principle (Berthoz, 1997), see Fig. 4. The teacher movements in the visual field of the student induce the changes in the perceived optical flow of the student. Using the proprioception informations, the student homeostatic system tries to cancel this changes by modifying his wheel speed; a tracking behavior emerge. The modifications of the robot position in the environment change the proprioception informations and "modulate" the visual flow informations. This mechanism can be assimilate to the reflex mechanism in the PerAc architecture (see Fig. 4).

First and foremost, the associations between the visual flow informations and the motor output commands are supposed learned. The rule which allows a robot to learn sensory-motor associations using a PerAc architecture

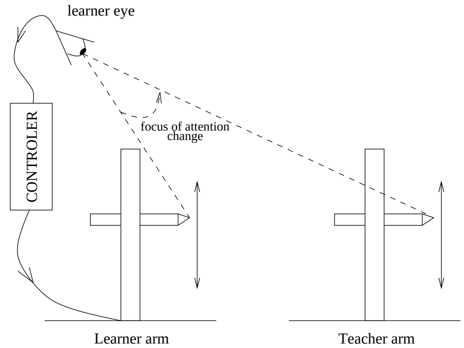


Figure 3: The student robot has already learned the correspondence between its arm internal representations and its positions in its visual field. If the student focuses its attention on the teacher arm, it will reproduce the teacher's movement just because it will perceive a difference between the proprioceptiv and the 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! An external observer will then deduce the learner robot is imitating the teacher.

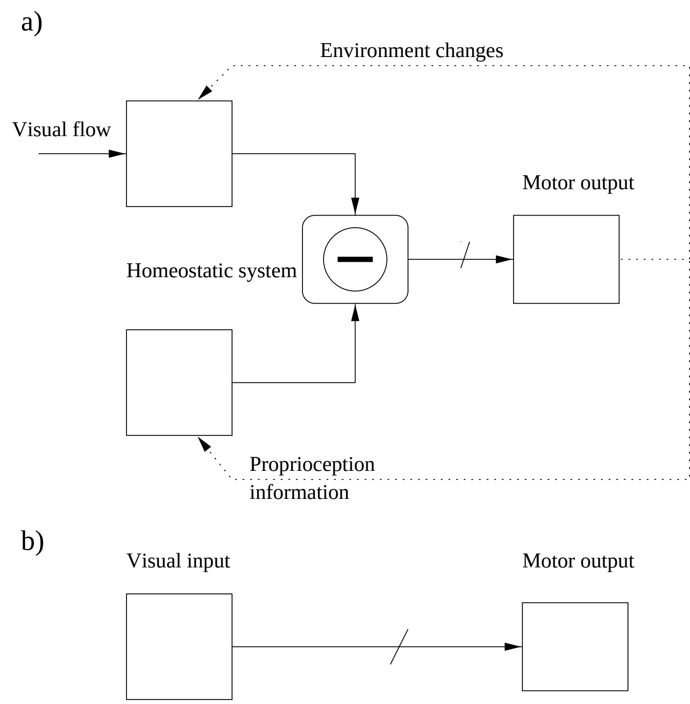


Figure 4: The implementation of the tracking mechanism. a) Using the difference between perception (visual flow) and action (proprioception) a reflex motor command is proposed. b) The mechanism can be assimilated to the reflex mechanism linking two maps of neurons. Different populations of neurons are used to code the pairs of direction and intensity of movements.

is showed in (Gaussier *et al.* , 1997c). In fact, in all the experiments shown below there is only one teacher (i.e. only one moving object). Also, we will suppose that a behavior or a skill is characterized by a sequence of movements (time series). This restrictive points of view will be sufficient to emphasize interesting problems mainly linked to the segmentation of the "atomic" actions the student robot has to learn and imitate.

3 Robot tracking mechanism

At the beginning of our work, we tried a very simple and powerful movement tracking mechanism which consists in going towards the direction of the center of a moving area. Fig. 5 presents a typical example of a tracking situation used to allow the learner robot to reproduce the movement sequence (changes in body orientation) of the teacher robot.

A frame-grabber is used to take an image sequence. A time integrated image is built by the temporal

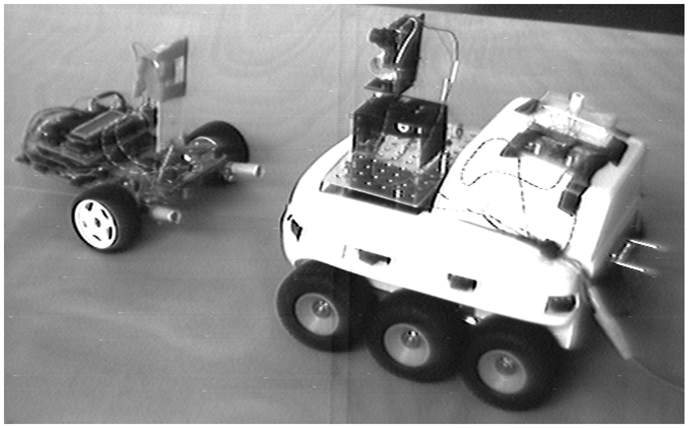
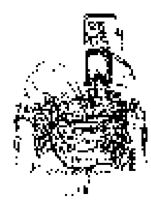


Figure 5: Overview of a tracking sequence. On the left, the teacher robot (a MIT handy-board base), on the right, the learner robot (KOALA from KTeam SA, Switzerland).

integration of a few (5 in our experiment) images from the CCD camera. A " movement image" is then directly computed by thresholding the pixel-by-pixel difference between 2 time integrated images of the above sequence (Fig. 6). In our architecture the information is coded on maps composed of analog formal neurons (the input images are directly copied on maps of neurons). The perceived movement direction is the value on the x axis of the neuron centered on the most activated area of movement and the speed of the proposed movement is deduced according to the projection of the same neuron on the Y axes (see Fig. 7).

a)



b)

Figure 6: a) A CCD image used by our robot in its tracking mechanism. b) The difference between two time integrated images gives informations about where the moving object is.

If the movement is just near the robot (in the lower part of the "movement image") then the robot speed will be negative and the robot will try to avoid the collision as shown in Fig. 8. In our experiment, we suppose, the robot has already learned to go backwards or forwards according to respectively the perceived expansion or contractions of the optical flow. Accordingly , when an expansion point appears on its left, the robot has learned to turn right. This behavior is frozen in the N.N. and considered as a reflex mechanism for the learning of our imitation tasks. This hard-wired mechanism allows the maintenance the same distance between student and teacher during the tracking task. In our PerAc architecture (see Fig. 15), it is represented by unconditional connections between two neural maps representing the robot movement proposal (Movement Input - MI) and the robot effective movement respectively (expressed in the same egocentric coordinates). This last map of neurons is a Winner Take All (WTA) called Motor Output (MO).

3.1 Shortcomings of the simple tracking mechanism

In order to learn a sequence, the student robot just has to detect and to learn the transitions of its body orientation and to be able to reproduce them.

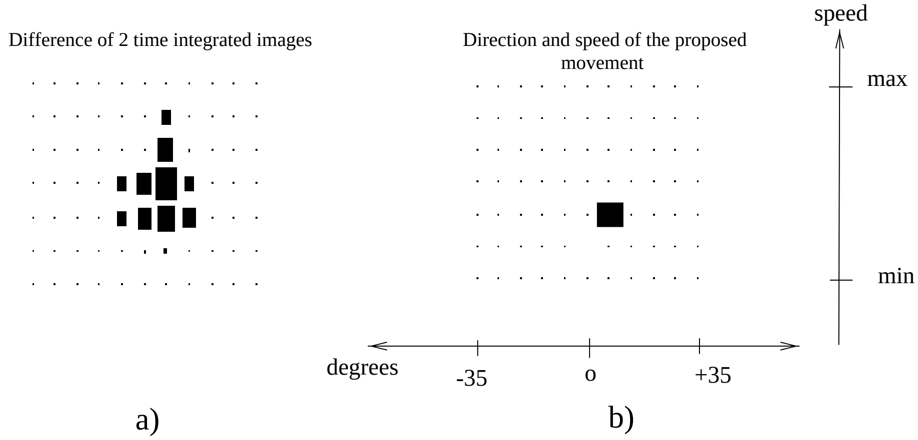


Figure 7: a) Perception of the teacher movement. The difference between two time integrated images (Fig. 6) is copied on maps of neurons. b) The result of the movement image is sub-sampled and the center of the maximum activity area is used to control the robot movement. A speed and an angle movement is proposed.

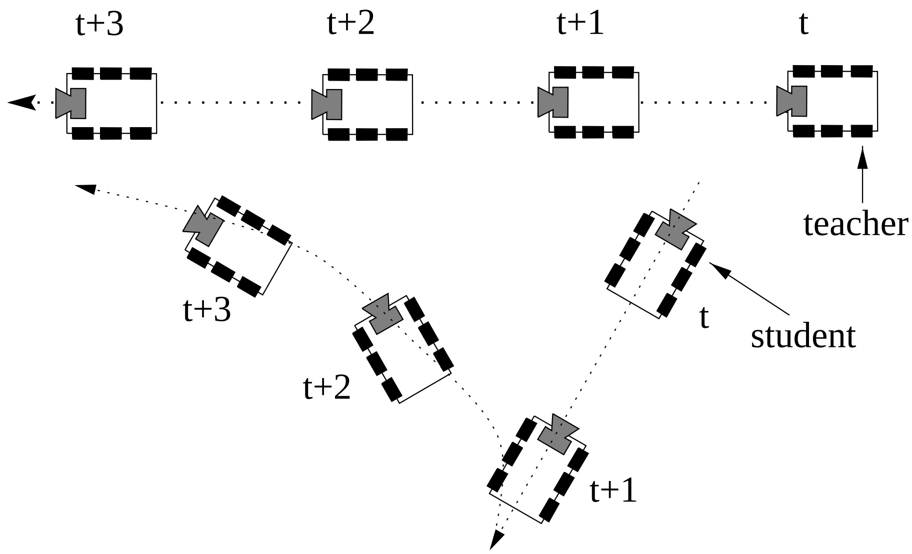


Figure 8: The reflex mechanism allows at the same time to track a moving object or to avoid collision according to the type of optical flow (dilatation/contraction) and to the apparent distance of the moving target.

The information (the perceived images and the motor commands) are represented according to egocentric coordinates. An arbitrary reference orientation is maintained in order to allow learning of the robot orientation transitions. The absolute orientation is maintained by a magnetic compass but the robot odometry information can also be used. The choice of the referential azimuth is performed, according to student orientation, at the beginning of each new sequence learning and the robot uses this reference while it learns/plays this sequence. Returning to a biological analogy, it is proved that we can dance a waltz by beginning with any body orientation!

The previously described tracking mechanism allows the student robot to correctly follow a teacher but its moves are quite noisy and unusable for sequence learning (see the robot orientation curve Fig. 9). The "noise" is mainly due to the difficulty to localize the center of the moving object. When the teacher orientation changes, its visual surface also changes and the center of the area can move very quickly without any direct link with the teacher trajectory. Accordingly, movement extraction problems (linked to thresholds problems) can change the shape of the moving object and also induce problems. Moreover, when the teacher is a human, the student can sometimes perceive the two legs as two different objects. Hence, the student robot track alternatively the two legs and a sinusoidal movement is superposed on the correct trajectory.

"Noise" is also due to the fact that the student takes shortcuts of the trajectory of the teacher (see Fig. 20) and does not provide a clear transition from one orientation to the other (the set of successive rotations to

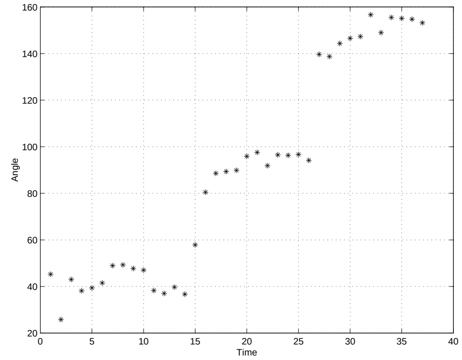


Figure 9: Curve representing the student robot orientation through time. The orientation is expressed in degrees. The 0 degree orientation corresponds to the north direction. Data are obtained from tracking a teacher robot performing a "zigzag" trajectory (see Fig. 20).

follow the teacher robot is not repetitive).

The use of a classical filtering algorithm allows suppressing of parts of this noise but induces problems in estimating the moment of the robot-body orientation transitions to be learned. The precise detection of robot orientation transitions becomes impossible. The problem could possibly be solved with a shorter step of the elementary robot movement (to obtain more points for the robot orientation curve) but then the teacher has to be constrained at such a slow speed that it becomes unusable in any real experiment. Also, increasing the room dimensions to allow longer trajectories is not an acceptable solution since the time interval between transitions increases and does not allow the teacher to repeat the trajectory with the same timing (human beings are not very precise to reproduce long time intervals).

The main problem is to filter the movement "noise" linked to the variability of the movement perception (target form perception, CCD camera limitations) and to the limitations of the possible movements (mechanical constraints). Hence, the orientations of the teacher must be integrated through time. An orientation change is taken into account only if the variation of the integrated absolute orientation of the robot is high enough.

There is a dilemma between the trajectory smoothing and the precise detection of the body orientation changes. A linear and passive filtering does not appear to be a good way to allow a precise segmentation of the sequence of movement.

3.2 The movement segmentation problem: An active filtering mechanism

Because of those trajectory filtering problems, we first decided that the main movements to be imitated are movements in a straight line. But then, if the robot is allowed to perform only straight movements it very often loses the teacher (the teacher goes out the visual field of the learner camera when it performs a fast and important change in its movement orientation). To avoid this new problem, we allow the robot camera (robot head) to rotate thanks to a servo-motor (control in azimuth). In this new system, the head tries to pursuit the teacher at any time by centering it in its visual field. The robot body turns only if the teacher movement is observed in the same angle for a given time interval.

Hence, the filtering is performed on the head movement and not only on the perceived image or the robot movements. This new architecture allows the robot to perform a clear trajectory, composed by long line segments (see Fig. 10). The head direction is the imaginary axis corresponding to the center of the visual field of the robot. The head angle is the angle between the head direction and the body direction.

The details of the neural architecture are presented on Fig. 11. The motor input block (MI) provides the information on perceived movements as an angle relative to the head direction. This angle is translated by the switching block (SW) in an angle relative to a pseudo fixed referential system. The translation mechanism of the switching block correspond to the $\Sigma\Pi$ neuron modeling ((Durbin & Rumelhart, 1989), (Koch & Ullman, 1985)). The SW block uses the absolute robot position, via an electronic compass (C). The head orientation (ϕ) and the body orientation are known at any time by the robot.

The time needed to integrate enough activity linked to the head orientation and to trigger a real robot rotation allows learning of the delay prediction between two rotation instants very accurately. Indeed, because the student robot is approximately at 60cm to 90cm from the teacher it would be a problem if it turned exactly

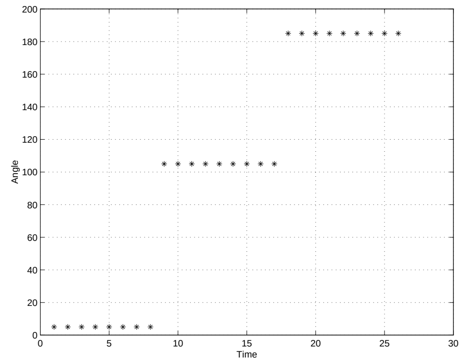


Figure 10: Curve representing the learner robot orientation through time. The orientation is expressed in degrees. The 0 degree orientation corresponds to the north direction. Data are obtained from tracking a teacher robot performing a "zigzag" trajectory (see Fig. 20). This curve show the performances of the denoising mechanism.

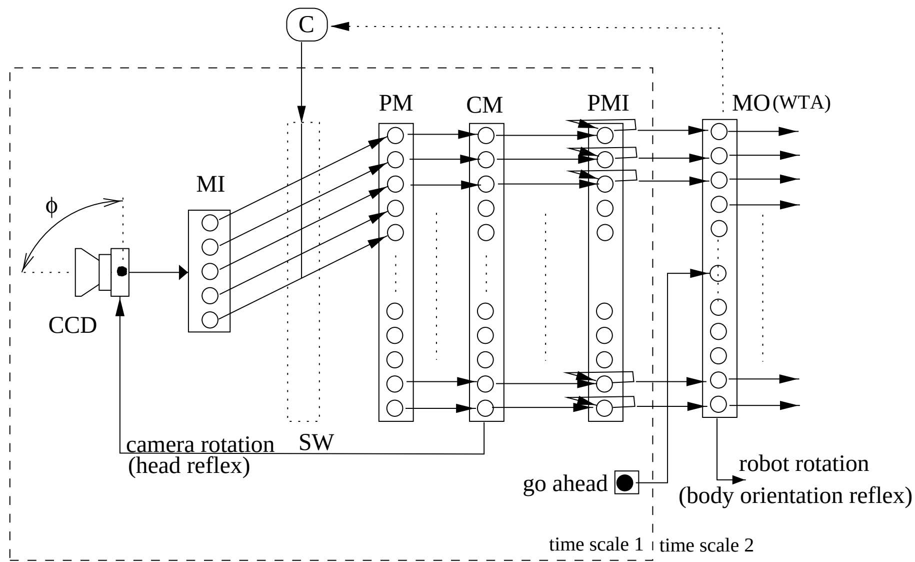


Figure 11: Sketch of the complete reflex mechanism able to control at the same time the head and the body direction in the robot following task. The motor output is expressed in the same coordinates as the retinal information (egocentric coordinates). CCD-CCD camera, MI-motor input (horizontal projection of the perceived movement), MO-motor output (Winner Take All group), SW-switching block (angle translation), C-electronic compass, PM-projection map (perceived movement in the egocentric referential system), CM-camera map (enable the camera rotation), PMI-perceived movement integration.

at the same time as the teacher. The time to detect the new orientation of the teacher allows the student to maintain the same distance between itself and the teacher. Also, it allows to obtain a correct measure of the time interval between two successive rotations of the teacher (just by measuring the interval between the learner rotations).

A first reflex (the head reflex in Fig. 11) is used to enable the fast alignment of the head direction with the direction of the perceived movement. The output of the block which performs the head reflex is also used as an input for a second reflex controlling the alignment of the body direction and of the perceived movement direction (the body orientation reflex). The time scale of the body reflex is quite different from the time scale of the head reflex. The head reflex is performed at each time "step". Time integration enables the robot to change its body orientation only if the perceived movements were detected in the same region of the visual field for some time. Fig. 12 shows a tracking situation in which there is no change in the robot body direction because the target does not always turn in the same direction (noise filtering).

Fig. 13 shows the opposite case. At t_0 time, the direction of the perceived movement varies and is not sufficiently stable to generate the change in the body direction. At time t_1, the teacher movement appears always approximately in the same direction. The temporal integration of the motor input activity in that

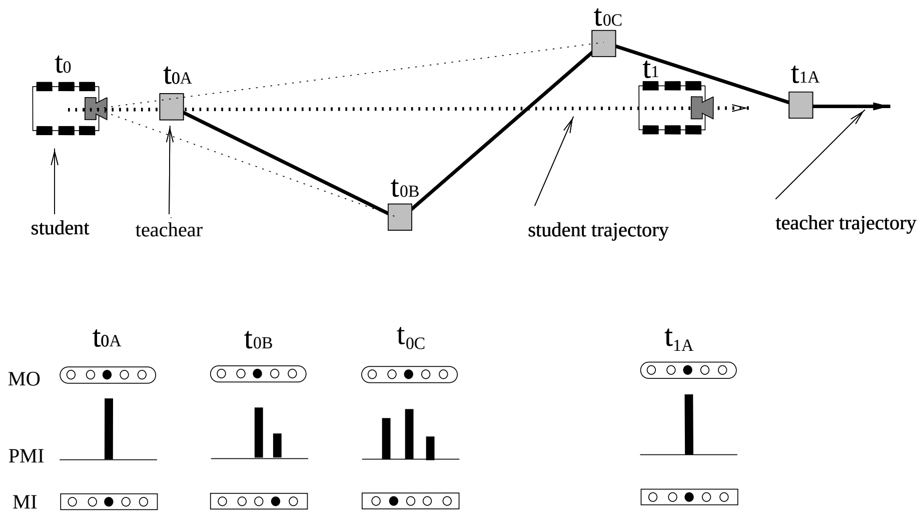


Figure 12: Overview of the tracking situation without changes of the robot body direction.

direction is then high enough to trigger a body rotation of the learner robot in the direction of the teacher robot.

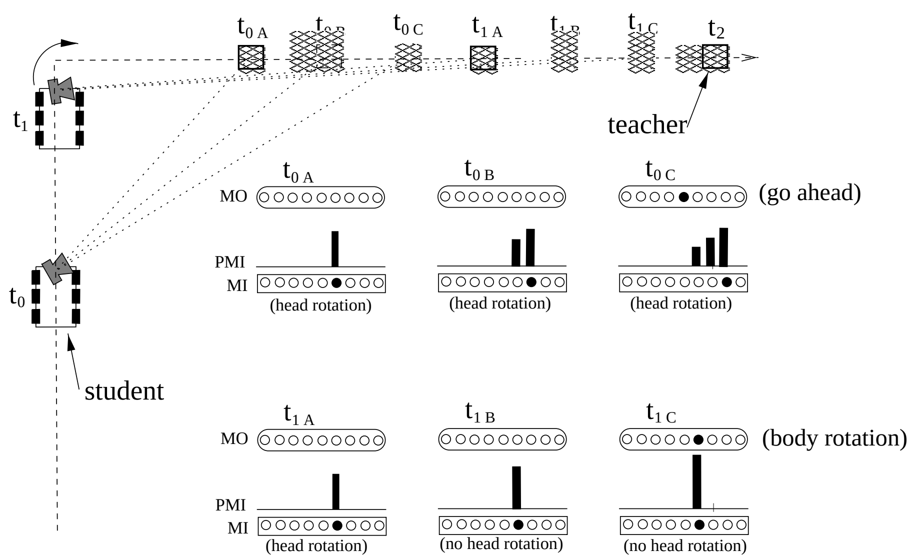


Figure 13: Filtering mechanism used to suppress the "noise" of perceived movements. At time t_0, the direction of the perceived movement varies (t_{0A}, t_{0B} and t_{0C} head's rotation) and is not sufficiently stable to generate the change of the body direction (see PMI activation). At time t_1, the teacher movements appear always approximately in the same direction (no head rotation) and the student performs a body rotation. MI- Motor Input, PMI - Perceived Movement Integration (through time), MO - Motor Output.

Fig. 14 shows the effect on a typical situation in which the student robot succeeds in following correctly the teacher (it turns at the correct place - after the correct delay) even if the teacher movements appear sometimes in the wrong direction.

4 A robust N.N. architecture for learning temporal sequences

Our robot does not voluntarily try to learn directly to imitate a human or another robot (a home-made simple robot - see Fig. 5). It learns only to reproduce its own sequence of actions primarily induced by the tracking reflex behavior. It learns to predict its own next movement and can use this information to detect novelty (Denham & Boitano, 1996) (situations in which its predictions are wrong).

The movement changes characterized by OFF-ON transitions (Time Derivative TD group) of MO neurons

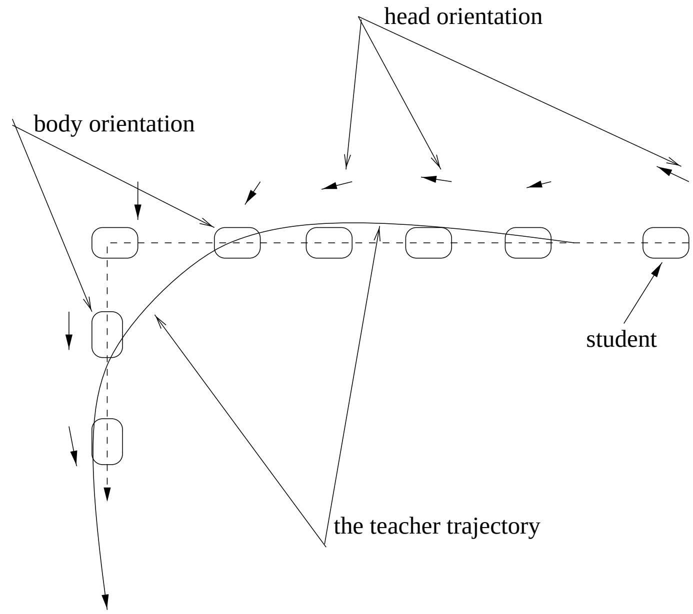


Figure 14: Overview of the typical tracking situation. The student trajectory is showed by a dotted line and the teacher trajectory is showed by the full line.

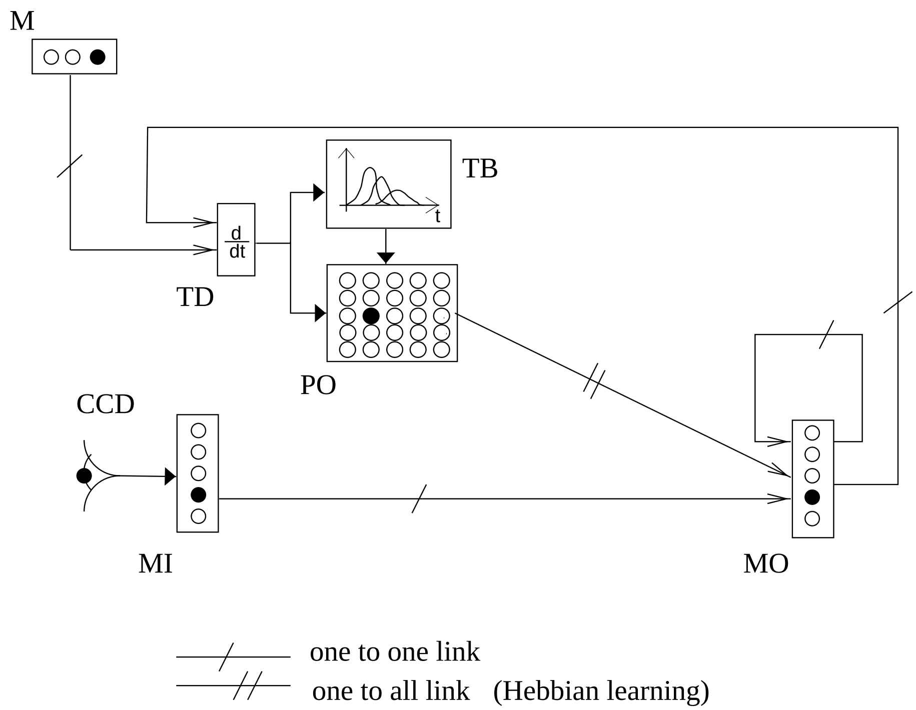


Figure 15: Overview of the PerAc architecture for robot following and sequence of movement learning. CCD - CCD camera, M - Motivations, MI - Movement Input, MO - Motor Output, TD - Time Derivator, TB - time battery, PO - Prediction Output

are used as input information for a bank of spectral neurons (Grossberg & Merrill, 1992; Bullock *et al.* , 1994) (TB in Fig. 15). Time filter batteries (TB) act as delay neurons endowed with different time constants. As such they perform a spectral decomposition of the signal that will allow the neurons in the Prediction Output group (PO) to register transition patterns between two events in the sequence (Banquet *et al.* , 1997). An input to a specific battery of TB granules performs both a reset of any eventual residual activity in this battery, and an initialization of the spectral timing activity within the group of cells of the battery.

The mechanism predicting new events is inspired from the functions of two brain structures involved in memory and time learning: the cerebellum and the hippocampus (the hippocampus is involved in the transient (short, medium or long term) storage of events and sequences (see (Banquet *et al.* , 1997) for more neurobiological references) and the cerebellum learns tuning and patterning of motor skills like ballistic trajectories (Bullock *et al.* , 1994)). Here, the main interest of this kind of N.N. code is to allow a compact and robust coding of time. Short intervals (less than 2 seconds) are stored with a high precision while long intervals (several seconds) are stored with a less precision. Monitoring a full range of time intervals according to a Weber law is of high importance for the trial to trial incremental learning since the teacher will never repeat exactly the same timing for long movements. The time battery used for learning looks like a wavelet basis and allows time scale problems to be avoided.

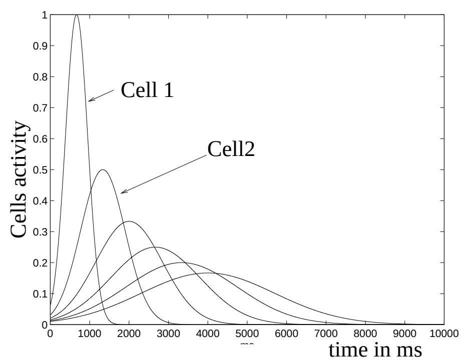


Figure 16: Time activity of a group of cells that allows to measure time in TB.

There are 15 cell elements in a battery. Time activity of 5 batteries of cells is presented in Fig. 16. The activation law of a TB cell is presented in Eq. (1).

$$Act_{j,l}^{TB}(t) = \frac{1}{m_j} \cdot \exp - \frac{(t - m_j)^2}{2 \cdot \sigma_j} \tag{1}$$

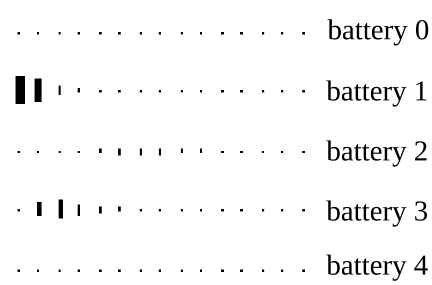


Figure 17: Time activity of TB after a sequence 2, 3, 1.

where j is associated to the choice of a particular time constant and l to the label of the recognized input pattern that triggers on its associated time battery. m_j is the value of the time constant and σ_j its associated standard deviation.

The PO group receives information from TD through both the direct and indirect pathways (Fig. 15). It learns the intervals between two events of the sequence in the strength of the connections between TB and PO (proximal connections). A new input from TD (distal part of the dendrite) triggers the learning of its association with the trace of the previous input in TD. The learning of the transition between the new and the previous input on TD is encoded by a PO neuron. Its associated neuron in PO learns to predict the transition according to the trace of the previous item presented on the neurons of TB. Then, PO will encode transition patterns between successive events in the sequence.

The potential of a PO neuron is the sum of distal and proximal activities. The potential of a PO neuron (i,j) is computed as in Eq. (2).

$$Pot_{i,j}^{PO} = \sum_{l} W_{po(i,j)}^{tb(j,l)} \cdot Act_{j,l}^{TB} + W_{po(i,j)}^{td(i)} \cdot Act_{j}^{TD} \tag{2}$$

$Act_{j,l}^{TB}$ is the activity of the l-cell of the j battery of TB. $W_{po(i,j)}^{tb(j,l)}$ is the strength of the link between the TB(j,l) neuron and the PO(i,j) neuron. One neuron of the PO is linked to all neurons of a battery of the TB. Act_{j}^{TD} is the corresponding TD neuron of a PO(i,j) neuron and $W_{po(i,j)}$ the strength of link between them. The variation through time of a PO neuron is shown in Fig. 18.

A PO neuron only fires when its potential reaches its maximum value (the sign of the derivative of $Pot_{i,j}^{PO}$ change from positive to negative). The firing condition is computed using Eqs. (3)-(4). Fig. 18 shows the potential of a PO neuron and the moment it fires. The activation (firing) of a PO neuron corresponds to a prediction of a new pattern.

$$Act_{i,j}^{PO} = f_{PO}\left(Pot_{i,j}^{PO}\right) \tag{3}$$

$$f_{PO}\left(x\left(t\right)\right) = \begin{cases} 1 & if\ \frac{dx(t)}{dt} < 0\ and\ \frac{dx(t-1)}{dt} > 0 \\ 0 & otherwise \end{cases} \tag{4}$$

The strength between a PO neuron and a TB battery is modified using the Eq. 5.

$$W_{po(i,j)}^{tb(j,l)} = \begin{cases} \frac{Act_{j,l}^{TB}}{\sum_{j,l}\left(Act_{j,l}^{TB}\right)^2} & if\ Act_{j}^{TD} \neq 0 \\ unmodified & otherwise \end{cases} \tag{5}$$

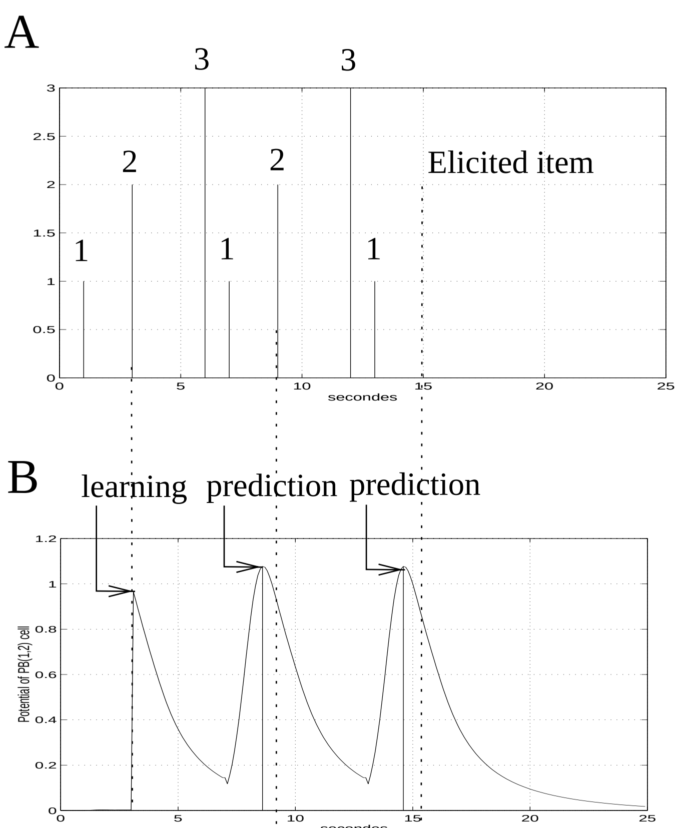


Figure 18: A. Time series of the presented patterns. B. Activity of a neuron associated to the prediction of pattern 2 by pattern 1.

Finally, the PO group is linked to the MO group via a one-to-all link. Hence, after a first sequence of actions, a motor neuron will be activated by the reflex input and will also receive information from the transition prediction group (PO). A simple conditioning rule then allows the activated neuron to react the next time the action is predicted even if the reflex does not provide information. Moreover, performing that action provides information to the event prediction mechanism that will reinforce the sequence. Thus replaying the sequence allows it to be maintained in memory. The learning mechanism needs the presentation of only one or two complete sequences of movements to learn to predict the changes of movement (the system learns the timing) and to replay the correct sequence. If the sequence of movements induces a positive reward then the past predicted transitions and their associated movements are reinforced. This way, the robot learns to imitate the behavior of the other robot. It succeeds to reproduce learned sequences of movements according to the activated motivation. The detailed connectivity of the prediction part of the network is presented in Fig. 19.

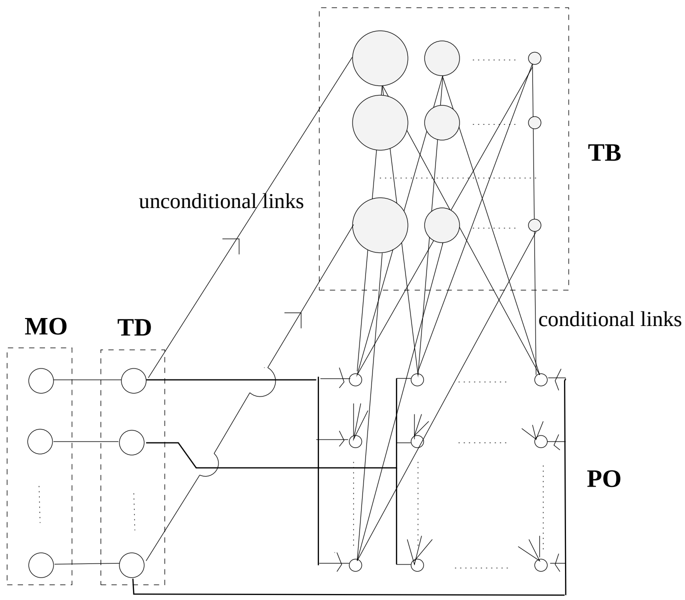


Figure 19: Detailed connectivity of the event prediction network. The circle size in TB is associated to the time constants (m_j) of the neurons.

The prediction capabilities of the event prediction network (Fig. 19), was tested for a large set of inter events interval (from 150 ms to 6 s). The response of the net is always in accordance with the Weber law (the prediction precision is directly dependent of the time interval between 2 events).

The link between our N.N. architecture and an hippocampal model can be established if one considers that TD is associated to the entorhinal cortex, TB to the Dentate Gyrus and PO to the CA3 region (for more details see our work on neurobiological modeling of the hippocampus (Banquet *et al.* , 1996)). The hippocampus is a brain structure strongly involved in mammals imitation processes. One of the signs of autism in infants or young children is socio-emotional disturbances (e.g. a reduction of initiation in social contacts). Imitation is supposed to be an important mechanism in child development, especially in development of individual contacts and social relationship. The absence of mimicry or other forms of imitation could be viewed as the signs of autism. Besides, the emergence of autistic-like behavior was observed in monkeys with damage to hippocampal formation (Bauman & Kemper, 1997) and we find here one possible link between imitation and the hippocampus.

5 Application of our sequence learning architecture to sequence imitation

The imitation of any sequence learning is triggered by an internal motivation or an emotion. At the beginning of any new sequence, we consider that the orientation of the student is at “0” degree. After learning a new sequence, the robot can perform this sequence with any orientation of its body (example of the waltz steps). The sequence is performed in response to an internal trigger signal (motivation, emotion).

We now describe step by step the activation and the learning in the different components of the system when the teacher performs a zigzag trajectory. For sake of clarity, we suppose that Time Derivation block (TD in Fig. 19) receives the sequence of elementary patterns “go ahead at this orientation”, “turn right 90 degree and go ahead” , “back to initial orientation and go ahead”, ... from the motor output block (MO). This kind of sequence is similar to the sequences of patterns received at the various time intervals that we studied in the previous section. To learn this simple trajectory (the zigzag) , we will consider the student robot will have to learn the timing of transitions between the 2 orientations, i.e. the transition from the direction 1 to the direction 2 and from the direction 2 to the direction 1. To simplify, we will speak about patterns 1 and 2.

While the activation of pattern 1 is maintained in the motor output (MO) block, the TD corresponding unit performs a time derivation to time battery (TB) and the output is sent to prediction output (PO). The activated pattern from TD to TB resets any residual activity in the corresponding battery (a row in the TB "matrix") of TB neurons. At the same time, it triggers an activation of the different components of this battery. Due to the absence of any significant previous pattern through the indirect TB pathway (we are at the beginning of the sequence), there is no significant learning related to pattern 1 , i.e the activated pattern from TD to PO has no effect. This sole TB input is not strong enough by itself to trigger learning at the PO level. Learning will only result from the conjunction of this TB input with a phasic transient signal from TD. In our case this direct input will now be pattern 2, the next pattern of our sequence. In the same manner, the presentation of pattern 2 generates an activation of the corresponding TD cell. This activation resets the corresponding TB battery and triggers learning of 1 to 2 transition. Thus the system learns the time interval between the pattern 1 and 2 in the weights of TB to PO connections. After the end of the sequence first presentation, i.e. when event 1 is presented again, the transition from 2 to 1 will be learned, that is the detection of the cyclic appearance of a sequence. At this point of the repetition of the sequence, the PO predictive capacity will become operational. The sequence 1,2 is now learned and can be performed without the teacher. The pattern 1 will trigger the 1-to-2 transition, the pattern 2 will trigger the 2 to 1 transition and so on.

As the presentation of events 1 and 2 allowed the latent learning of the 1-2 transition, similarly, the presentation of a new event "3" could provide learning the 2-3 transition, in the case of a 1, 2, 3 or more complex sequences.

In the robotic experiment, the student robot succeeds in learning of different skills (or sequences) with correct timing. All these sequences are mimicked correctly. There is no restriction on sequence length or number of sequences, other that the ones imposed by the total storage capacity of the neural net and the fact that the same event must not appear twice in the sequence. The timing of a sequence is learned very well and the corresponding trajectory is correctly performed. This sequence learning architecture has been implemented in our mobile robot in real time. In Fig. 20, a), we show the teacher and the student pursuit trajectory, i.e. the student trajectory during the tracking phase. The pursuit trajectory is not exactly the teacher's trajectory. The movements sampling and the physical limitations of CCD camera can be sources of this kind of "noise". This noise is removed using the time integration and the "denoising" mechanisms discussed above. The learned trajectory, i.e. the trajectory performed by the student after the learning period, is very close to the teacher's trajectory. The errors are smaller than 10 % in orientation and/or distance (see Fig. 20, b)). The on-line learning aspect of our model allows the correction of imprecise or mis-estimated timing in the sequence of movements: when the teacher again performs the same trajectory, the learned trajectory will be the sum of the new and the last learned trajectories.

In our real time experiment, the robot imitates 3 movement sequences: a square, a zig-zag and a complex movement (see Fig. 20). The choice of behavior is modulated by an internal variable. This internal variable can be viewed as an emotional or motivational state of the student robot. For example, if the student robot performs a behavior (A) and after a time period it does not receive any reward, it can decide it is tired of doing the same thing and decide to change his behavior.

For instance, the internal variable which modulates the behavior of the robot is directly wired to the Infra Red sensors of the robot. Saturating one of this sensors, we can modify the state of an internal variable and trigger one of the learned behaviors. Also, the interaction with the robot, via the Infra Red sensors can be viewed as a primitive communication form.

6 Conclusions and perspectives

We have shown that a N.N. approach of learning through imitation can be employed on a real autonomous robot and give interesting results. In the beginning, we have used our architecture to study robot/robot interactions. In these series of experiments, the purpose was to investigate how a robot (the student) can learn to perform the different trajectories. The student is equipped with the ability to imitate, i.e. to follow, to learn and to predict. We use another robot for teaching purpose (the teacher). The notion of "teacher" should be understood as a "trainer".

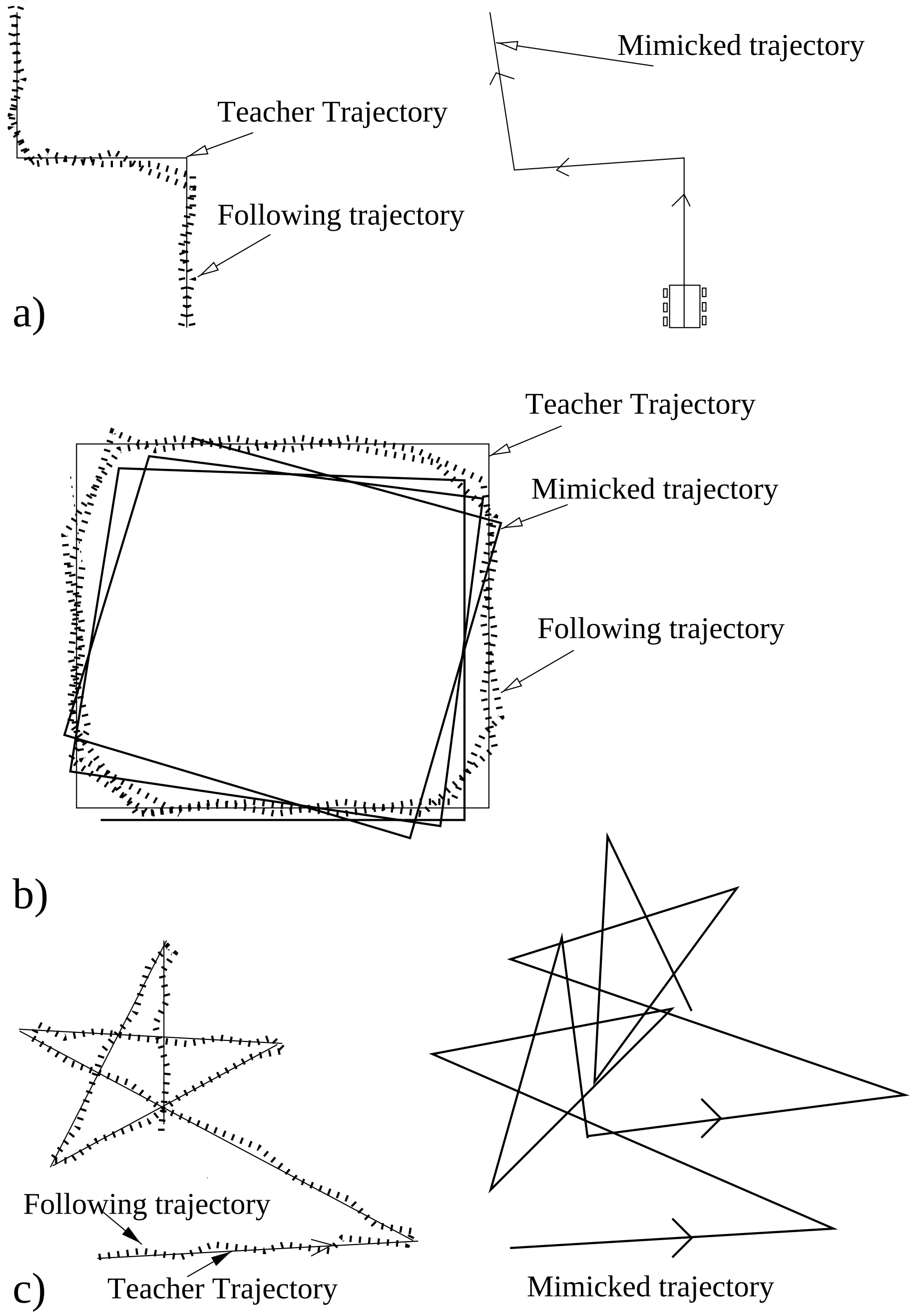


Figure 20: 3 experimental trajectories learned in the same experiment by our robot. The access to the replay of a partial sequence is triggered by a user touch on a particular part of the robot body (simple communication mechanism).

One of the first definitions of imitation is provided by (Thorndike, 1898): imitation is “learning to do an act from seeing it done”. After that, a lot of definitions was be done (see (Whiten & Ham, 1992) for a review). It is interesting to notice that ”imitation” seems to be very dependant of the sensorial modality. For instance, there is an important contrast between bird vocal imitation and non vocal imitation used to learn object manipulation . Whatever, in all those cases, the subject attempts to reinstate only the outcome of the model action (Whiten *et al.* , 1996) which explains why ethologists prefer to talk about ”valence transformation” or ”Emulation” (Wood, 1989). Lot of ”imitation” experiments can be simply explained by a stimulus enhancement : ”B learns from A to what (object or location) to orient behavior” (Whiten & Ham, 1992). With the help of stimulus enhancement, an animal B might quickly learn behavior similar to A’s by trial and error or may even apply matching techniques already in its repertoire to the correct location, quickly solving the problem and giving the impression of imitation (Whiten *et al.* , 1996). Real imitation is defined by Whiten as the fact: "B learns from A some part of the form of the behavior". The presented work is clearly about imitation but it will be very important to allow our robot to used stimulus enhancement too. In N.N. framework, it could be seen as a simple priming effect. the robot would have a higher probability to focus on that object and to take it. As in our imitation system, the enhancement of the probability to focus on the right object need the capability to detect it as something different from the other object. So we can imagine that both imitation and stimulus enhancement relies on the same functional derivation mechanism or novelty detection mechanism It means a time derivation of a behavior trajectory for the imitation (to find the key points of the trajectory) and time derivation of objects features for stimulus enhancement.

According Whiten's taxonomy, the imitation in animals can be divided in tree categories: non-social processes (mimicry arise without social interactions of student and teacher), social influence (the student is influenced by the teacher but does not learn any part of mimicry from the teacher) and social learning (the student learns some part of the mimicry from the teacher). It is clear that learning using social interactions, like imitation, suppose the pre-existence of the behavioral categories. According Whiten's classification of animal imitation, we "use" the mimicry with correspond to the first-level of imitation defined by Mitchell ((Mitchell, 1987) and (Dautenhahn, 1995)). Due to tracking reflex mechanism there aren't the social interactions between the student and the teacher.

But, during our experiments, we realized that our system can be more efficient with a human teacher. In the learning process, we observed that there are in fact bidirectional interactions between the teacher and the student: obviously the student follows the teacher and learns to associate a specific trajectory with an internal state but at the same time the teacher cannot avoid adapting his behavior in order to help the student to perform the correct trajectory. The teacher looks at the robot trajectory and increases or reduces his speed to allow the student robot to pass exactly through the different correct edges of the trajectory. The timing of the action activation of the student is then more precise and the trajectories shape is better because the rotation orders are also more precise.

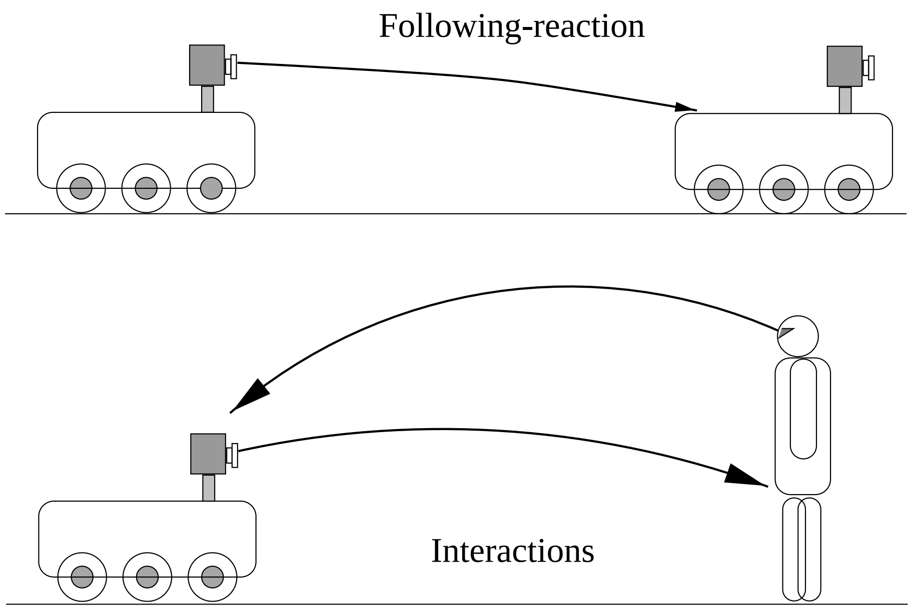


Figure 21: Natural interaction mechanisms between human and robot allows better performances than just an imitation through a "spying" process. The difference between level-1 (a)) and level-2 (b)) of imitation. The performances are better in b) because the teacher is aware of the learner.

Obviously the trajectory learned by the student after one trial remains rough. When the teacher performs the same trajectory several times, the student can refine the reproduced trajectory using the difference between its prediction and the perceived movement. This interactive process between the teacher and the student shows the close link between imitation and communication problems. This corresponds to the level two of Mitchell's definition of imitation.

Usually two different forms of learning by imitation are distinguished: learning through imitation and learning to imitate. We have already situated our work clearly with regard to the first case. Learning to imitate can be divided in two aspects: when to imitate and how to imitate. Taking a decision about when to imitate, can be negotiated through the internal variables of our robot architecture , (i.e. links to primary motivations). For instance, the robot can decide by itself to imitate because no other task has been triggered. Concerning the problem of how to imitate, our present system only knows a single way to imitate (it follows only the teacher - low level reflexes). It would be very interesting to provide it with other ways to imitate. A relevant problem should be to find how to select a particular imitation mechanism according to the external situation or according to the state of particular internal motivations. For instance, we could imagine learning the places where the teacher has decided to turn rather that learning the sequence of movement transitions. In the first case, the student would have to return to the different learned locations in the room (reproduce the path in absolute coordinates). In the second case, the student robot reproduces a relative trajectory according to its starting point.

This work is at its very beginning but its interest is in its use of the PerAc architecture which allows to reuse systems already developed in our lab. Tools using learning by imitation seem to be able to solve a wide variety of problems. Nevertheless, a lot of fundamental problems must be further investigated. First, how to decide the moving object is "like" the robot and so must be imitated? In our approach, everything could be imitated and after a while, if no reinforcement signal is received, the robot should learn the object is not interesting and should not be imitated (because its visual shape will be associated to an avoidance of the imitation process). The presence of more than one moving object in the learners visual field also supposes the existence of an attentional focusing mechanism able to discriminate the visual moving objects in order to determine who is the teacher. In a near future, we plan to add an already developed shape recognition and discrimination mechanism (Gaussier *et al.* , 1997d) in order to continue to track the target, even in the absence of a movement, or to avoid tracking an object learned to be uninteresting (its tracking is not associated at the end with a reinforcement signal). The reflex mechanism used here corresponds to the Action pathway in the PerAc architecture.

This approach could explain how a robot can learn to recognize somebody of its sort (conspecific) (Dautenhahn, 1995) and perhaps we will be able to generalize that to the learning of the consciousness of the robot itself (Dennett, 1991): this is my arm because I can predict what it will do (reward). My internal schemes involving my arm remain stable. Our N.N. architecture could aslo be a good starting point to allow robot to learn to communicate. It would be necessary to add an innate sociability reflex (and reward) to communicate that could induce the emergence of a communication and of social relationship (Piaget, 1936; Wallon, 1945; Salvador, 1997).

Another philosophical question linked to this kind of experiment is the following: In our case, the robot does not know it is imitating. Do we need to realize robots that really understand they are imitating? What modification is needed in the proposed architecture? Learning social relationships requires adding the possibility of learning "together" and not only one from an other. Techniques for managing this kind of problems seem to be available but a part from a few experimentations (Dautenhahn, 1995), real size applications managed in a bottom-up approach have yet to be considered.

A great challenge for the future could be to succeed in merging together learning by imitation and social interaction capabilities (learning to work together to solve a particular task). The main problem is certainly the lack of enough characteristic examples of robot cooperation to understand the different features required to build a theoretical framework useful for the design of a globally efficient control architecture to solve those kind of problems.

References

Albus, J.S. 1991. Outline for a Theory of Intelligence. *IEEE trans. on syst. and cybern.*, **21**(3), 473–509.

Bakker, P., & Kuniyoshi, Y. 1996. Robot See, Robot Do : An Overview of Robot Imitation. *In: AISB Workshop on Learning in Robots and Animals.*

Banquet, J.P., Contreras-Vidal, J.L., Gaussier, P., & Burnod, Y. 1996. *Fundamentals of neural network modelling for neuropsychologists.* Boston: MIT Press. Chap. The cortical-hippocampal system as a multirange temporal processor: A neural model.

Banquet, J.P., Gaussier, P., Dreher, J.C., Joulain, C., & Revel, A. 1997. *Cognitive Science Perspectives on Personality and Emotion.* Elsevier Science BV Amsterdam. Chap. Space-Time, Order and Hierarchy in Fronto-Hippocamal System: A Neural Basis of Personality, pages 123–189.

Bauman, M., & Kemper, T. 1997. *The neurobiology of autism.* Baltimore, USA: The John Hopkins University Press.

Berthouze, L., Bakker, P., & Kuniyoshi, Y. 1996. *Learning of Oculo-Motor Control: a Prelude to Robotic Imitation.* IROS, Osaka, Japan.

Berthoz, A. 1997. *Le sens du mouvement.* Paris: Odile Jacob.

Brooks, Rodney A. 1986. A Robust Layered Control System for a Mobile Robot. *IEEE Journal of Robotics and Automation*, **R.A. 2**(1), 14–23.

Bullock, D., Fiala, J.C., & Grossberg, S. 1994. A Neural Model of Timed Response Learning in the Cerebellum. *Neural Networks*, **7**(6/7), 1101–1114.

Dautenhahn, K. 1995. Getting to Know Each Other - Artificial Social Intelligence for Autonomous Robots. *Robotics and Autonomous System*, **16**(2-4), 333–356.

Demiris, John, & Hayes, Gillian. 1996 (July). Imitative Learning Mechanisms in Robots and Humans. *In: Proceedings of the 5th European Workshop on Learning Robots.* 1996, Bari, Italy.

Deneubourg, J.L., Goss, S., N.Franks, Sendova-Franks, A., Detrain, C., & Chrétien, L. 1990. The Dynamics of collective Sorting: Robot-Like Ants and Ant-Like Robots. *Pages 356–363 of:* Meyer, J-A., & Wilson, S.W. (eds), *From Animals to Animats: Proc. First Int. Conference on Simulation of Adaptive Behavior.*

Denham, M.J., & Boitano, J. 1996. *A model of the interaction between prefrontal cortex, septum and the hippocampal system in the learning and recal of goal-directed sensory-motor behaviours.* Tech. rept. NRG-96-01. University of Plymouth - School of computing.

Dennett, D.C. 1991. *Consciousness Explained.* Boston, Massachusetts: Brown.

Durbin, R., & Rumelhart, D.E. 1989. Product Units: A Computationally Powerfull and Biologically Plausible Extension to Backpropagation Networks. *Neural Computation*, **1**, 133–142.

Galef, B.G. 1990.

Gaussier, P., & Zrehen, S. 1994. Avoiding the World Model Trap: An Acting Robot Does Not Need to Be So Smart! *Journal of Robotics and Computer-Integrated Manufacturing*, **11**(4), 279–286.

Gaussier, P., Joulain, C., Moga, S., Quoy, M., & Revel, A. 1997a (July). Autonomous Robot Learning: What Can We Take for Free ? *Pages SS1–SS6 of: International Symposium on Industrial Electronics - ISIE'97.* IEEE, Gumarães, Portugal.

Gaussier, P., Moga, S., Banquet, J.P., & Quoy, M. 1997b. From Perception-Action loops to imitation processes: A bottom-up approach of learning by imitation. *Pages 49–54 of: Socially Intelligent Agents.* AAAI fall symposium, Boston.

Gaussier, P., Revel, A., Joulain, C., & Zrehen, S. 1997c. Living in a partially structured environment: How to bypass the limitation of classical reinforcement techniques. *Robotics and Autonomous Systems*, 225–250.

Gaussier, P., Joulain, C., Zrehen, S., Banquet, J.P., & Revel, A. 1997d (September). Visual Navigation in an open environement without map. *In: International Conference on Intelligent Robots and Systems - IROS'97.* IEEE/RSJ, Grenoble, France.

Grossberg, S., & Merrill, W.L. 1992. A neural network model of adaptively timed reinforcement learning and hippocampal dynamics. *Cognitive Brain Research*, **1**, 3–38.

Hayes, G.M., & Demiris, J. 1994. A Robot Controller Using Learning by Imitation. *Pages 198–204 of: Proceedings of the 2nd International Symposium on Intelligent Robotic Systems.*

Hecht-Nielsen, R. 1987. Counterpropagation Networks. *Applied Optics*, **26**(23), 4979–4984.

Heyes, C.M., Dawson, G.R., & Nokes, T. 1992. Imitation in rats: initial responding and transfer evidence from a bidirectional control procedure. *Quaterly Journal of Experimental Psychology*, **Section B: Comparative and Physiological Psychology**(45B), 229–240.

Koch, C., & Ullman, S. 1985. Shifts in selective visual attention : towards the underlying neural circuitry. *Human neurobiology*, **4**, 219–227.

Kuniyoshi, Y. 1994. *The Science of Imitation - Towards Physically and Socially Grounded Intelligence -.* Special Issue TR-94001. Real World Computing Project Joint Symposium, Tsukuba-shi, Ibaraki-ken.

Matarić, M.J. 1995. Issues and Approaches in the Design of Collective Autonomous Agents. *Robotics and Autonomous System*, **16**(2-4), 321–331.

Mitchell, R. W. 1987. A comparative-developmental approach to understanding imitation. Perspectives in Ethology.

Piaget, J. 1936. *La naissance de l'intelligence chez l'enfant*. Geneve: Delachaux et Niestle Editions.

Revel, A., P.Gaussier, Lepretre, S., & Banquet, J.P. 1998 (August). Planification versus sensory-motor conditioning: what are the issues ? *In: Accepted to SAB 98*. SAB, Zurich.

Salvador, L. 1997. Assimilation, Imitation and the Elementary Social Fact. Toward a Definition of Social Interactions. *Pages 115–117 of: Socially Intelligent Agents*. AAAI fall symposium, Boston.

Thorndike, E. L. 1898. Animal Intelligence: An experimental study of the associative process in animals. *Psychology Rev. Monogr.*, **2(8)**, 551–553.

Tomasello, M. 1990. *"Langage" and Intelligence in Monkeys and Apes*. Cambridge University Press. Chap. Cultural transmission in the tool use and communicatory signaling of chimpanzees ?, pages 274–311.

Wallon, H. 1945. *Les origines de la pensée*. Paris: Presses Universitaires de France.

Whiten, A., & Ham, R. 1992. On the nature and Evolution of Imitation in the Animal Kingdom: Reappraisal of a Century of Research. *Pages 239–283 of:* Slater, P.J.B., J.S, Rosenblatt, Beer, C., & Milinski, M. (eds), *Advances in the study of behavior*. San Diego, CA: Academic Press.

Whiten, A., Custance, D., Gomez, J.C., Teixider, P., & Bard, K.A. 1996. Imitative learning of Artificial Fruit Processing in Children and Chimpanzees. *Journal of comparative psychology*, **110**(1), 3–14.

Wood, D. 1989. *Interaction in human development*. Hillsdale NJ: Lawrence Erlbaum Associates. Chap. Social interaction as tutoring, pages 59–80.