THE ROLE OF EMOTIONS IN SENSORI-MOTOR LEARNING

R. MANZOTTI*, G. METTA*, G. SANDINI*
**LIRA-Lab, DIST, University of Genova,*
Via Opera Pia 13, 16145 Genova, Italy
Email: pasa@dist.unige.it

ABSTRACT

The aim of the present paper is to show that emotions, which play an important role in development and learning both in humans and animals, could be embedded in an artificial system. Besides, we emphasize the importance of emotions during learning and development as endogenous teaching devices.

Speaking from a biological point of view, emotions can be seen as "states produced by reinforcing stimuli". Some neural structures like the amygdala, the orbital cortex, and the cingulate cortex as well as other subcortical areas are involved in the processing of such signals. Among them, the amygdala seems to play a crucial role being implicated in the learning of associations between stimuli and reinforcers.

To illustrate our point we will present an implementation of a plausible model on a real head-arm robot, which learns both how to control movement and when to move toward a given target.

We stress the biological parallelism at each stage of our implementation. For instance the motion control paradigm is based on the so-called "force field approach" proposed in the literature as a model of the spinal cord functions. At a higher level a network using Hebbian weights updating creates the associations between relevant stimuli and reinforcers.

1. Introduction

The aim of the present paper is to demonstrate that "artificial emotions" could be embedded in the action-perception-learning cycle of an autonomous robot. In the field of neuroscience, it has been realized that the neural correlates of emotions can be investigated with the same tools used, for example, to study motor control, memory or other high-level brain functions [A. Damasio 1994, J. Le Doux 1996]. Through these studies, the important role of emotions as endogenous teaching devices during development and learning (among others cognitive processes) has been demonstrated and the identification of the brain areas involved in emotional responses has been possible (see J. Le Doux 1994, for a review).

Among them the amygdala plays a crucial role. This structure receives projections from the peripheral sensory systems both through the thalamus and through the high-order sensory cortices many of them receiving feedback signals from the amygdala (allowing the tuning of low-level sensory processes). The amygdala projects also to a number of dedicated structures eliciting, unconsciously, stereotyped behavior such as freezing or running. These structures include the autonomic and the endocrine responses (see also Figure 1).

Besides, amygdala is the locus where learning of association between stimuli and reinforcement signals (primary emotional signals) take place. It is worth noting that this effect seems to happen on the synapses from neurons receiving sensory signals that project onto neurons receiving inputs from primary reinforces (i.e. smell, noxious stimuli, etc). Synaptic changes appear to follow an Hebbian-like rule [D.O. Hebb 1949].

The role of amygdala in learning seems to be limited to the generation of new bindings between stimuli and reflex responses. More flexible behaviors can be elicited by the activation of some cortical and sub-cortical areas (e.g. the orbital cortex, the hypothalamus) and may be explained by the James-Lange-Damasio hypothesis of somatic markers [A. Damasio 1994].

James and Lange first recognized that emotions are embodied not just in the neural structure of the brain but also they may be physically part of the body itself. The body is the theatre in which emotions are represented (unconsciously) and where a specific situations can be associated to a particular body response such as heart rate or blood pressure changes (James 1890, Damasio 1994). Following this theory, the body becomes, through emotions, a processing element of the cognitive architecture of an active being. This structure has several advantages. One of them is the fact that the system is able to develop a sort of subjective personality (different behaviors in similar external situations). The representation of the body's emotional state is embedded in the body itself, which after all, is its own best representation (Brooks 1991).
Starting from these considerations, we propose a model that includes "artificial emotions" as teaching devices. Such model has been then tested in a simple "go/don't go" robotic task. In particular the robot has to learn how to reach with the arm visually identified targets and, at the

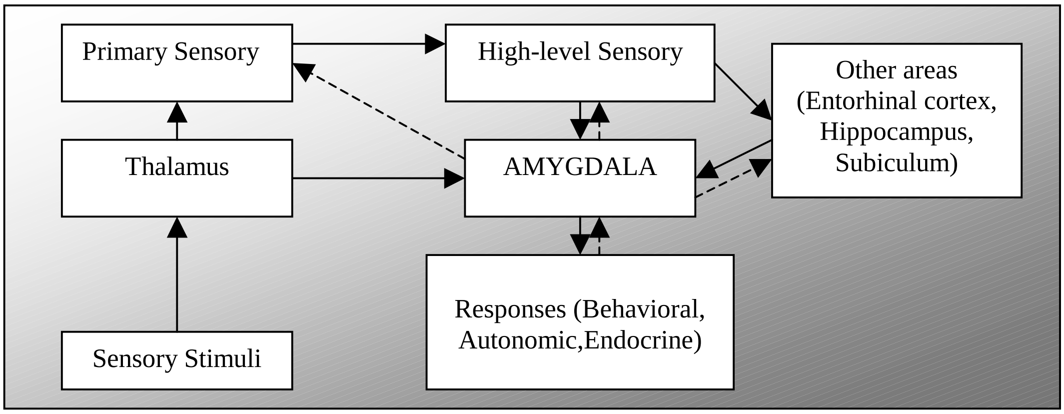

Fig 1. A simplified sketch of amygdala connections. It shows the interconnecting pathways between various brain areas and the amygdala. As can be seen the amygdala receives inputs from different levels of sensory processing and projects to structure devoted to the generation of responses (solid lines). There are also projections (dotted lines), which allow the amygdala to control how sensory processing is carried on.

same time, to discriminate between "good" and "bad" targets. This experimental situation is similar to what happens to babies when they learn to reach for a visual target and, simultaneously, learn that touching some targets (e.g. a flame) is painful and motion toward them has to be inhibited. In section 2 we present the specific experiment we realized with our robotic set-up together with the overall structure of the system. Sections 3 and 3.1 deal with the motion control paradigm, which is based on the so-called "force field approach" first proposed by Mussa-Ivaldi and Bizzi (Mussa-Ivaldi et.al. 1993). In sections 3.2 and 3.3 the algorithm used to learn the control parameters of the arm is described. Section 3.4 describes the associative learning between stimuli and reinforcement signals and its implementation on the robot. Finally, section 3.5 presents some experimental results.

2. Emotion and reinforcement

As we said above, emotions influence learning, perhaps, in order to achieve better performance. In human beings the emotional basis of learning is all but straightforward. On the contrary, in our experiment the implementation of emotional states and feelings is still rather simple and it is essentially related to the generation of a pleasure/pain sensory feedback to reinforce/inhibit specific sensori-motor behaviors. In spite of the rather simple emotional content of the experiment it is important to stress that the overall system is acting and learning sensori-motor coordination also on the basis of internally generated body signal explicitly coding an *emotional* parameter.

Figure 2 shows the overall architecture of the robot system. The interface with the external world is, for the time, based on visual data processed through two parallel pathways. One used to build the visuo-motor mapping required to reach the visual target with the arm, the other processing the emotional value of what is happening. This value (coded with respect to the pain/pleasure scale) is assigned on the basis of information coming from the environment (a noxious signal or

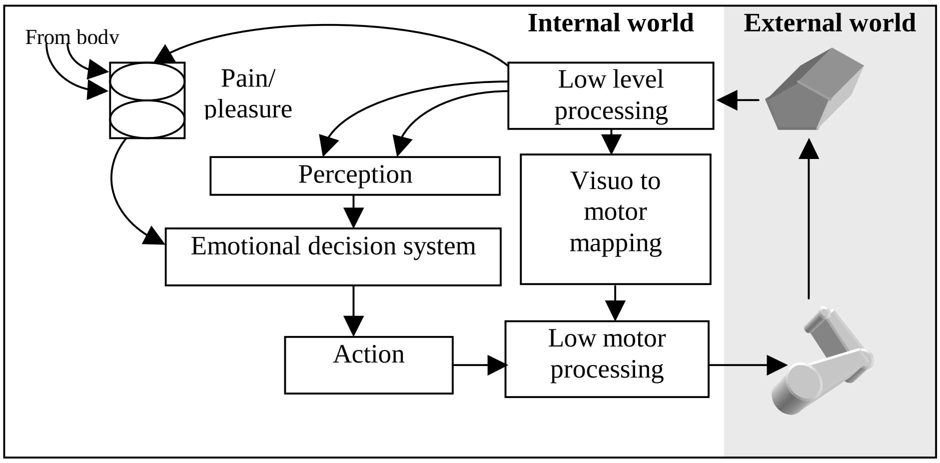

Fig. 2. The overall architecture of the robot system; information flowing through sensory channels (e.g. vision, proprioception) is processed in parallel by the "emotional subsystem" and by a specific visuo-motor coordination subsystem. Both subsystems are learning from the environment-robot interaction although the former is also taking into account emotional information.

a "teacher") as well as from internal memory (e.g. a dark spot can be arbitrarily considered as something to be avoided). These events are all mapped into the pain/pleasure structure that acts on the higher order decision network as a reinforcement signal. This latter decision network receives its input from the current state of pain/pleasure system as well as from the higher order perception network. Based on such information it decides the next action to perform which, in our case, is either to enable or to inhibit the arm's reaching action. The specific experiment we describe here concerns the learning of arm-reaching control movements toward *good* visual targets (and the inhibition of such reaching toward *bad* targets). In the framework of the present paper, therefore, learning has both a physical and an emotional component. We use only 4 of the 10 degrees of freedom of our robot: two of them controlling the direction of gaze of a color camera and two controlling the motion of the arm (which is, therefore, constrained to move on a plane). The distinction between *good* and *bad* objects is based on color. More specifically reaching toward green objects is reinforced while the experimenter (acting as an external teacher) punishes reaching movements toward red ones. To increase the difficulty of an otherwise simple associative task, visual processing, which is based on a space-variant color camera (Sandini 1980), provides, besides color information, also a set of non-relevant signals such as target position and size (computed in the image plane).

3. The experiment

If we consider an artificial system engaged in a reaching task, the number of motor degrees of freedom (d.o.f.) that have to be controlled in parallel can be as high as ten or more. A

fairly standard approach to this kind of problem is to characterize the plant as much as possible. After the identification procedure a general purpose or a customized control law can be applied (Yoshikawa 1990). In this case the problem is simplified by an accurate design but it might require a substantial effort during the design phase in order to be applied.
An alternative solution, which we believe is more efficient, is based on direct motor primitives representing multi-joint synergies. In this case a single command may produce complex multi-joint coordinated movements without the voluntary control of each individual d.o.f. The overall learning procedure is defined by the following points:

- the motor primitives;
- the mechanism of sensori-motor mapping;
- the developmental rules;
- the influence of reinforcement signals on learning.

3.1. Motor primitives

As far as the coding of motor primitives is concerned, one possible procedure is the so-called force fields approach originally proposed by Mussa-Ivaldi and Bizzi (Mussa-Ivaldi 1992, Mussa-Ivaldi 1993). According to this theory the neuro-muscular control of a limb can be mathematically described considering that each joint is controlled by a set of actuators with spring-like properties. Each actuator is fully modeled by a torque field, such as:

$$\bar{\tau}(\bar{q},a)$$

where $\bar{q}$ is the vector of generalized coordinate, a is the activation value and $\bar{\tau}$ is the generalized torque field. Actuators, as used in our experiments, are described by the following equation:

$$\tau = -a\kappa(q - q_0)$$

where a is the activation value (which modulates the overall stiffness κ) and q_0 the resting configuration. From the mechanical point of view a multi-joints structure controlled by a set of spring-like actuators is a passive system. As a consequence it has a stable Equilibrium Point (EP) in its state space. The EP can be thought of as the point toward which the arm is moving at each instant of time and a limb trajectory can be represented by a sequence of EPs. Concerning the motor primitives, each of them can be represented by a structure, which activates a single or a group of actuators. Thus, the following torque field can describe each primitive:

$$\bar{T}_j = C_j \sum_i I_{ij} \bar{\tau}_i$$

where $\bar{\tau}_i$ is the i^{th} actuator field, C_j the activation value and:

$$I_{ij} = \begin{cases} 1 \text{ if the } j^{th} \text{ controller activates the } i^{th} \text{ actuator} \\ 0 \text{ otherwise} \end{cases}$$

The feasibility of this schema comes from the fact that any position of the arm workspace can be reached by combining linearly a small number of primitives represented through torque fields (which are called *basis fields*). The task of the controller is thus that of generating the activation values C_j. Specification of C_j determines consequently a new EP for the system. Following this model the total torque applied to the system is:

$$\bar{T} = \sum_j C_j \bar{T}_j$$

It is worth noting that if the torque corresponding to the previous equation were applied to the arm it would eventually move to the respective EP (at rest). An example of a force field for a two d.o.f. manipulator is shown in Figure 3 along with the corresponding resting position of the arm (Figure 4).

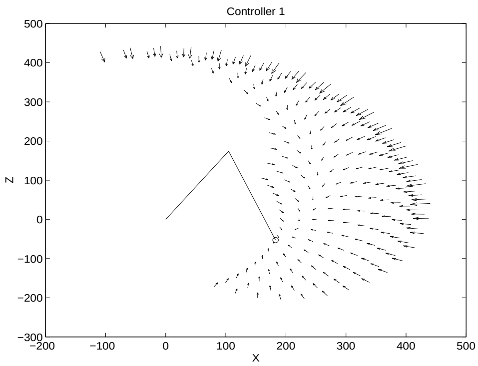

Fig. 3. Force field applied to the robot arm corresponding to the activation of one of the controllers.

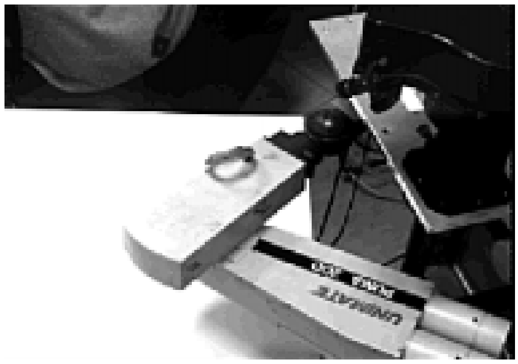
Fig. 4. Exemplar resting position of the arm corresponding to application of the force field showed in Figure 3.

3.2. Motor-motor map

The situation, however, becomes more complicated when visually driven, goal-directed movements (such as reaching a point in space) are considered. In this case the trajectory of the arm has to be controlled (or initiated) on the basis of visual information. The solution we propose is based on the use of a direct mapping between the eye-head motor plant and the arm motor plant. One premise we make is that the position of the eye's fixation point coincides (at least at some stage of the control process) with the object to be reached. In other words, the reaching of an object starts by looking at it. Under this assumption, the fixation point can be seen as the *end-effector* of the eye-head system. Its position in space with respect to the shoulder is uniquely determined by the motor commands controlling the position of the head with respect to the torso and that of the eyes with respect to the head. Assuming the eye-head plant is controlled using force fields, the position in space of the fixation point can be coded using motor commands (or alternatively the head configuration vector $\bar{q}$) and, at least in principle, the arm's force fields can be obtained by a transformation of the eye-head's force fields. We call this approach *motor-motor coordination*, because the coordinated action is obtained by mapping motor commands onto motor commands.

3.3. Development of a motor-motor map

Formally, the motor-motor map can be seen as a function $\bar{C} = f(\bar{q})$, which converts values from head position to arm activation. Under these hypotheses, each time step i of the proposed learning algorithm can be described as:

1. A proper stimulus appears in the field of view;
2. By fixating the visual target the robot also initiates arm motion by computing the arm activation vector C in the following way:

   $$\hat{f}_i(q) + n$$

   The term n describes a zero-mean uniform noise component introduced to simulate errors in the arm control. $\hat{f}_i$ is the estimate of f at the i^{th} iteration.
3. The vector C is used by the arm controller, which moves the arm toward the new EP.

4. At this point the arm is as close as possible to the target so that the system can re-direct the gaze to its own hand.
5. As a result of the previous step a new pair $(\overline{q}, \overline{C})$ is available which is used to update the map by computing the value $\hat{f}_{i+1}(\overline{q})$ in the following way:

$$\hat{f}_{i+1}(\overline{q}) = \hat{f}_i(\overline{q})\frac{n-1}{n} + \frac{\overline{C}}{n}$$

   where n is the number of visits of the cell corresponding to $\overline{q}$.
6. The arm then returns to a fixed resting position.

It is worth mentioning that the map must be initialized in a meaningful way in order to allow the initiation of movements. In our experiment the robot utilizes three initial reflexes which are manually tuned and stored into the map from the beginning. In practice these reflexes are similar to the asymmetric tonic neck reflex found in newborns and serve the important goal of keeping the arm in the field of view thus eliminating the need to search for the arm end-point.

3.4 Learning emotional value

Besides the learning of the head-arm coordination map, the system receives also a reinforcement signal from the experimenter in order to select a particular class of objects. A hebbian neural network carries on the higher level decision. It receives sensory and reinforcement inputs and generates the decision (i.e. reaching/no reaching) as output. The network learns to distinguish between bad and good objects on the basis of the reinforcement signal given by the teacher at the end of each motor action. It is worth noting that the relevant information (color) is embedded into a set of non-relevant sensory signals. Despite this fact, the neural network is able to learn which subset is relevant for the task at hand.

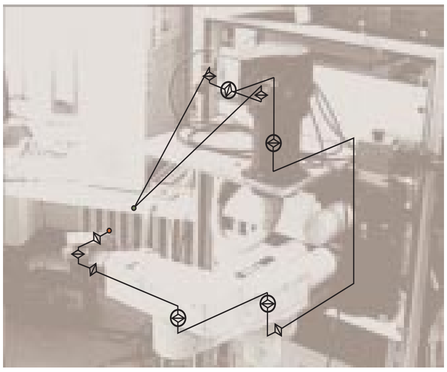

Fig. 5. The experimental setup. The robot has ten d.o.f. (marked by small quadrilaterals). The solid lines identify links. The last two links emanating from the cameras are "virtual" and their intersection point represents the "end-effector" of the head subsystem.

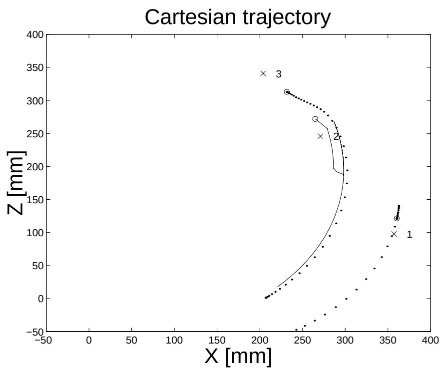

Fig. 6. Trajectories of the hand. The picture shows three typical arm trajectories. Abscissa and ordinates are respectively the X and Z coordinate. The X-Z plane is the plane where the arm motion is constrained. Small "x" signs identifies the target, "o" represents the end-effector final position.

3.5 Experimental Results

Figure 5 shows the experimental setup described in section 3.1. In order to assess the performance of the system the arm trajectories when reaching toward some fixed targets was recorded and the positioning error at the end of each trial measured. Figure 6 shows three exemplar trajectories. The “x” signs represents the target position, the “o” are the end-effector final positions. It is worth noting that before the training the robot can only use the initial reflexes as described in section 3.3, the precision is very low, approximately 10-15 cm. After learning, precision increases to about 3 cm on average, which is the maximum achievable with the present implementation. At the same time, the system learns also to discriminate among two different classes of targets. With regard to the latter point, it exhibits an increase in success rate (the percentage of correct decisions) from 50% (random decision) to about 90%. It takes only 40 trials to learn how to take the correct decision and roughly 100 to build the visuo-motor map.

4. Conclusions

The system presented in this paper is still far from being a close model of a learning system where emotional data are explicitly considered. However we believe that the implementation is a first step in that direction. Initially the robot is controlled by three reflexes: the arm is extended to the right whenever the eyes are looking to the right. Similarly, when the eyes look straight ahead the arm is extended in front and it is extended toward the left if the gaze is directed to the left. The appearance of a target in the field of view initiates the overall procedure: the gaze is directed toward the target and consequently the arm is extended in the proper direction. If the experiment were limited to this behavior the system would eventually learn to reach successfully a visual target. What is added by the emotional part of the system is the possibility for the system to learn that not all targets are the same and, consequently, to inhibit the reaching action in the presence of “bad” targets. In this sense it is the first appearance of a voluntary control over an otherwise exclusively reflexive behavior. In other words the system learns how to inhibit selectively the reflexive actions embedded initially, a step which is considered essential also for a correct development of sensori-motor coordination in humans.

Acknowledgements

Research activity described here is supported by the Italian Ministry of Research and University (MURST grant No. 9709112600) and by the EU TMR grants VIRGO (FMRX- CT96-0049) and SMART-II

References

Adolphs R., D. Tranel and A. Damasio (1998) “The human amygdala in social judgement” *Nature*, Vol 393:470-474.

Brentano F., (1973), *La psicologia dal punto di vista empirico*, (trad. Giovanni Gurisatti , orig. F. Brentano (1873), *Psychologie vom empirischen Standpunkt*), Laterza, Roma.

Brooks R., 1991, “Intelligence without representation”, *Artificial intelligence* 47, pp.139-159

Damasio A.R. (1994), “*Descartes’ Error: Emotion, Reason and the Human Brain*”, Putnam's Sons.

Hebb D.O. (1949) “The Organization of behavior”, Wiley, New York.

Hyman S.E. (1998) “A new image for fear and emotion” *Nature*, Vol 393:417-419.

James W. (1890), *The principles of Psychology*, Dover, NewYork (1950)

Ledoux J.E. (1996), “In Search of an emotional system in the Brain: Leaping from fear to Emotion and Consciousness”, in *The Cognitive Neurosciences*, M. Gazzaniga ed., Bradford Book, 1049-1061

Morris J.S., A. Ohman and R. J. Dolan Ditzinger, (1998) “Conscious and unconscious emotional learning in the human amygdala”, *Nature*, Vol 393:467-470.

Murphy, S.T. and Zajonc, R.B. (1993). Affect, cognition and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, 64, 723-739

Mussa-Ivaldi F. A., S. F. Giszter and E. Bizzi (1993) “Convergent force fields organized in the frog's spinal cord.” *The Journal of Neuroscience*, 13(2):467-491.

Sandini G. and V. Tagliasco (1980) “An anthropomorphic retina-like structure for scene analysis”. In: *CGIP*, 14 No.3:365-372.

Sandini G., Metta G. and Konczak J. (1997) “Human Sensori-motor Development and Artificial Systems”. In: *AIR&IHAS '97*, Japan.

Shannon C. and W. Weaver (1963), *The Mathematical Theory of Communication* (Urbana, Chicago: Univ. of Illinois Press).

Yoshikawa T. (1990), *Foundations of Robotics, Analysis and Control*, MIT Press.