EMOTIONS AND LEARNING IN A DEVELOPING ROBOT

R. MANZOTTI*, G. METTA*, G. SANDINI*
**DIST, University of Genova, via Opera Pia 13, Genova, Italy 16100, Italy*

ABSTRACT

The role of emotion has been underestimated in the field of robotics. We claim that emotions have a twofold aspect relevant to the building of a purposeful active robot: a cognitive aspect and a phenomenological one. We need to understand both these aspects. With regard with the first of them, it is possible to split it by at least two other relevant points. First, emotions could be the basis for binding between internal values and different external situations following the somatic marker theory. Second, emotions could play a crucial role, during development, both for taking difficult decisions whose effects are not immediately verifiable both for the creation of more complex behavioral functions. Thus emotions can be seen, from a cognitive point of view, as a reinforcement stimulus and in this respect, they can be modeled in an artificial being. Inasmuch, emotions can be seen as a medium for linking rewards and values to external situations. Besides, we would like to accept the division between feelings and emotions. Emotions are, in James' words, the body theatre in which several emotions are represented and feelings are the mental perception of them. We could say that feelings are the qualia of the external (even if bodily) events we could call emotions. We are using this model of emotions in the development of our project: Babibot. We stress the importance of emotions during development as endogenous teaching devices.

1. Emotions and Qualia

1.1. Cognitive and phenomenological side

Traditionally emotions have not been considered in the development of robots. The main reason has been the confusion between their cognitive and phenomenological aspect. It is important to clarify the difference.

Speaking in a purely cognitive fashion, we can see emotions as simple devices assigning to a particular situation a global value without having to analyze the details. Thus, emotions can be seen as reward variable able to represent wide collections of external situations. For example, if animals (or particular classes of them) have no conscious experiences they could still have unconscious emotions.

Phenomenological emotions, or feelings[1], can be seen as the conscious perception of cognitive unconscious emotions (Damasio 1994). They require a conscious being. Its existence (and necessary conditions) is still far from being obvious (Dennett 1991). Without entering the details of the consciousness debate, we only state that we accept the existence of consciousness as a real phenomenon waiting for a proper answer (Chalmers 1996, Searle 1992). Given this position,

[1] Somebody refer to the dualism unconscious emotion/conscious emotion as emotion/feeling, other as mood/emotion. We prefer the former choice.

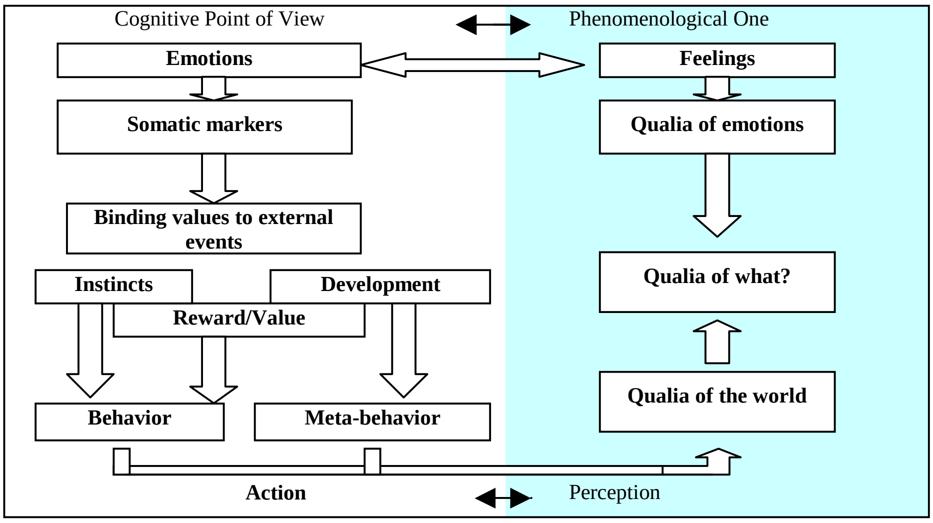

Fig. 1. Comparison between the cognitive side of emotion and their phenomenological side. Most, if not all, of the behaviouristic aspects can be treated in the cognitive side.

feelings have to be explained as well as other subjective entities. What we can hope to give in this context is just a cognitive framework for their phenomenological counterpart.

1.2. James' theatre

Emotions are thus useful devices to represent possible future rewards or punishments. The idea is that they can be embodied not just in the neural structure of the brain but that they could be physically part of the body itself. The body would be the theatre in which the emotions are represented (unconsciously). In this fashion, given a particular situation, we can associate to it a particular body response and, after, we can perceive it in order to plan our future actions (James 1890). In other words, the body became, by the use of emotions, a part of the cognitive architecture of an active being. This structure has several advantages. First the system perceives part of itself and is thus able to develop a sort of subjective personality. Secondarily the system can process the external information through two different channels: the sensory channel perceived directly by the neural system and the body itself with its physical interaction with the environment. For example if I have eaten too much it is probable that my mood will change and that, my disposition toward different kinds of physical activities will change too. The blood flux toward my stomach has directly modified the body variables and I perceive them as a modification in my mood.

More complex emotional responses can be obtained as a result of neural structure explicitly devoted to the activation of a particular emotional response. They can be activated by specific (and unconscious) stimula regarding whatever relevant aspect of the environment (visual expressions, dangerous situations, philogenetically selected stimula, etc.). Among these dedicated neural structure we can quote the amygdala, the cingulate cortex and the talamus (Lane *et alia* 1998, Le Doux 1996, Morris 1998, Adolphs 1998).

1.3. Feelings as qualia of what?

Emotions are thus a specific body response to a particular kind of stimulus. This response can be different with respect to the global state of the body. For example, the sight of a naked person of the opposite sex can elicit different responses in the same person due to its overall conditions. Eventually the brain, through several internal sensory channels (blood pressure, hearth rate, breathing rate, etc.), can perceive this body state. This perception involves the existence of the mental object that carries the meaning of the emotion: a feeling or qualia of an emotion.

Phenomenologically speaking qualia can be seen as the carriers of the meaning of perception. When we have a red sensation (we are in such state that it is something to be us seeing something red) we are in some kind of relation with the meaning of ‘red’ (intentional, identity or whatever). Then a particular qualia has a meaning that is in relation with the kind of information it is carrying. We stress the difference we see between the subjective meaning carried by qualia and the objective information carried by the relations between qualia. Speaking of feelings, as we have said, they are the conscious perception of emotion (that is a complex body state perceived through several channels). A particular feeling can thus be seen as the qualia of a particular emotion. If an emotion represents something (danger, joy, happiness, whatever), its qualia is the subjective meaning associated with that something. If a red qualia is the meaning of a red patch then a feeling, or qualia of an emotion, is the meaning of something more complex but equally real (Nagel 1974, Stubenberg 1998).

2. Emotion learning

2.1. Emotion and reinforcement

In our model emotions are associated with a particular body state. It is important to associate to particular situation the corresponding body state and, conversely, to associate to a particular body state its corresponding reward value. The external world is thus perceived from sensory channels, from separate specialized sensory channels eliciting specialized emotional structure (like amygdala) and from its direct effects on the body.

A hard-wired pain/pleasure structure is needed to give to the system the first feedback about the results of its actions. They can serve as teaching devices able to reinforce or inhibit a particular behavior. The state of the body (possibly after having being modified by an emotional response) together with the normal sensations is eventually perceived by the system which chooses the best action.

3. BabiBot project

3.1. The experimental setup

The goal of the present experiment is two fold: a) learn to control the arm motion in order to reach for objects in extra-personal space; b) discriminate between two classes of objects on the basis of a reinforcement signal (primary reinforcer). As we shall see, the robot learns both to move the head and the arm in a coordinate fashion and to decide when to initiate a reaching movement.

We designed the experiment so that the system may see two kinds of colored objects. We then associated each color with a distinctive primary reinforcement signal (positive or negative). Associations between color and reinforcer are kept constant during the experiment. To increase the difficulty of an otherwise simple associative task, we added also a set of non-relevant signals such as target position and size (computed in the image plane).

In this experiment, we used a four d.o.f. setup. Two d.o.f. belong to the head which control the direction of one camera. The remaining d.o.f. belong to the arm whose motion is thus constrained into a plane. Vision is based on a space-variant color camera [add ref]. It has photoreceptor elements arranged in order to resemble the distribution of those of the human retina. The visual processing is based on a color segmentation algorithm, which is able to locate objects because of their colors. It provides the subsequent learning stages with information about location, size and color of objects in the environment.

3.2. Learning motor control

Motor control is based on the so-called "force fields" approach first proposed by [add ref]. Following this paradigm, the positioning of the arm in space can be obtained by controlling a set of parameters. They represent the relative stiffness of antagonist muscle pairs. In this case, the goal of the learning algorithm is that of building a map between the target positional information (gathered by vision) and

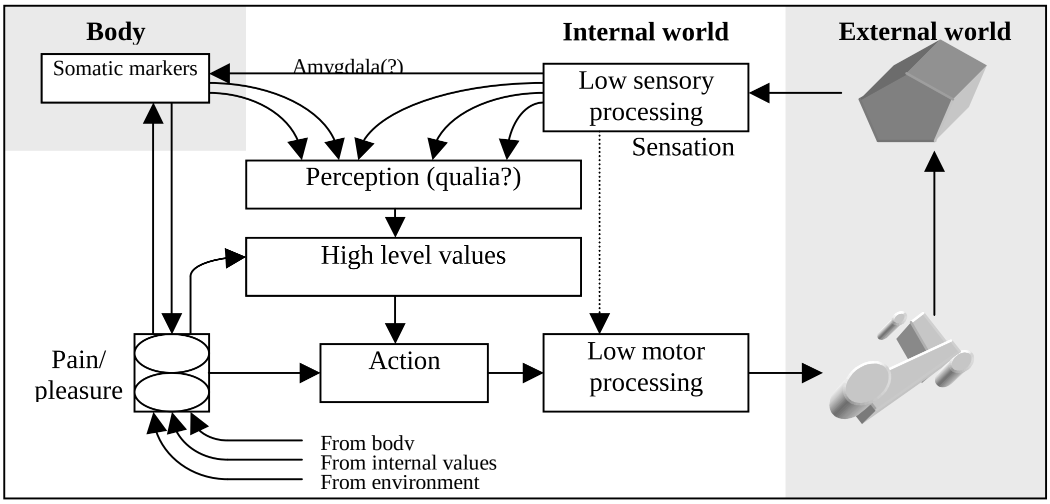

Fig. 2. Relation between the body, theatre for emotions; the internal world of perception; and the external, physical world.

the above mentioned control parameters. Indeed, a map can be assembled based on the continuous interaction between the robot and the environment. A supervised learning technique has been used in this case. It is worth mentioning the fact that as learning progresses the system smoothly switches from a purely reflexive behavior to a state where reflexes are used as primitives to perform coordinated reaching.

3.3. Learning when to initiate reaching

A 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 the two kinds of objects based on the reinforcement signal. It is worth noting that the relevant information (the color) is embedded into a set of sensory signals (mostly non-relevant). Despite this fact, the neural network is able to learn which subset is relevant for the task at hand.

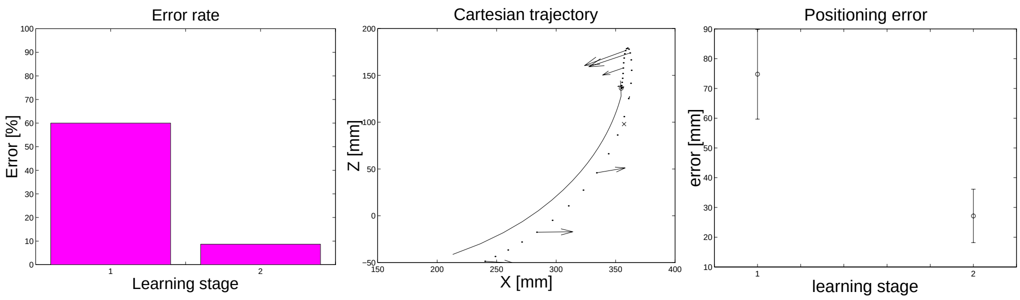

Fig. 3. To demonstrate the feasibility of the approach we present some experimental results. The system has been tested at different stages of learning. At these stages, we stopped the learning procedure and we carried on a test phase. Fig. 3.a. shows the success rate (percentage of correct actions) as function of the above mentioned learning stages. There is a clear trend toward the value of 100. Fig. 3.b. instead, plots the precision of the reaching actions. Again, the trend is toward an increase of the precision as the learning progresses. Finally, Fig. 3.c. shows an exemplar arm motion after the last learning stage.

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

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

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

D. Chalmers, (1996), *The Conscious Mind. In search of a fundamental theory.* New York: Oxford University Press

A. R. Damasio (1994), *"Descartes' Error: Emotion, Reason and the Human Brain",*

Dennett D. (1987), *The Intentional Stance*, Cambridge, Mass.: MIT Press

Dennett D. (1991), *Consciousness Explained*, Boston: Little, Brown and Company

Dennett D., *Brainstorms, Philosophical Essays on Mind and Psychology*. Brighton, Sussex: Harvester Press

Grosso E., G. Metta, A. Oddera and G. Sandini (1996) "Robust visual servoing in 3D reaching tasks". *IEEE Transactions on Robotics and Automation*, 12(8):732-742, October 1996.

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)

Kasniak A.W., S.L. Reminger, S.Z. Rapcsak, E.L. Glisky (1998), *Conscious experience and autonomic response to emotional stimuli following frontal lobe damage,* in Tucson III, Tucson

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

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

F. A. Mussa-Ivaldi and S. F. Giszter (1992) "Vector field approximation: a computational paradigm for motor control and learning" *Biological Cybernetics*, 67:491-500.

F. A. Mussa-Ivaldi, 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.

Nagel T. (1974) "What Is It Like to Be a Bat?" *The Philosophical review* 83. Pages 435-450

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

Searle J. (1992), *The Rediscovery of the Mind*, Cambridge: MIT Press.

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

L. Stubenberg (1998), *Consciousness and qualia*, Amsterdam: John Benjamins

E. T. Rolls (1996) "A Theory of Emotion and Consciousness, and its Application to Understanding the Neural Basis of Emotion" in *The Cognitive Neurosciences*, Bradford Book, 1091-1106