VERSITA

Emotions as a dynamical system: the interplay between the meta-control and communication function of emotions

C. Hasson, P. Gaussier[*]**, S. Boucenna**

ETIS, CNRS UMR 8051, ENSEA, Cergy Pontoise University 2 Av Adolphe Chauvin, 95300 Pontoise, France

Abstract

Classical models of emotions consider either the communicational aspect of emotions (for instance the emotions conveyed by the facial expressions) or the second order control necessary for survival purpose when the autonomy of the system is an issue. Here, we show the interdependence of communication and meta-control aspects of emotion. We propose the idea that emotions must be understood as a dynamical system linking two controllers: one devoted to social interactions (i.e. communication aspects) and another one devoted to the interactions within the physical world (i.e metacontrol of a more classical controller). Illustrations will be provided from applications involving navigation among different goal places according to different internal drives or object grasping and avoidance.

1. Introduction

Models and computational architectures of emotional systems can be divided into two different families: on one side, the models devoted to autonomous learning and meta-control and on the other side, the models devoted to the expression of an emotional state or more generally to the control of the expressiveness or interactivity (mainly developed for the purpose of Man Machine Interfaces). One can question the interest of building dedicated models for each aspect of the problem since the global dynamics (and the emergent properties) of an autonomous system interacting with humans may introduce constraints that could invalidate them [21]. For instance, if an animal flies, the other animals in the neighborhood may imitate its behavior inducing a contagion effect allowing the whole group to escape from the danger. In this case, the danger is perceived directly from the analysis of the congeners behavior. Hence, it is important to take into account the communicative function of emotions and the circular interactions [74] that may appear between the emotions as a meta-control system and the emotions as a communication system.

To test our emotional architecture, we propose to use an autonomous navigation paradigm since it is easy to measure the goal achievement and the interest of introducing an emotional system[1]. Figure 1 shows the robot and its environment. Following a behavioral approach [6, 17] in the frame of an Animat system [60], the robot must maintain a set of artificial physiological variables within safe limits to ensure its survival. Thus, the robot must look for different resources to fulfill its various needs (i.e. the robot must have goals that depend on its motivations [57, 58]). However, sustaining a durably efficient behavior in a dynamic

and complex environment remains a difficult task [8, 49]. The varying nature of the environment as well as the robot's own imperfections will lead to situations where the learned behaviors are not sufficient and might be end up with deadlocks [23]. Lacking the ability to monitor their behavior, robots get no satisfaction from productive actions and no frustration from vain ones. This is why most robots exhibit a very counterproductive rigidity when facing unforeseen situations.

Figure 1. The robot in its environment (5m x 5m). A color detector is placed under the robot. Colored squares on the ground represent simulated resources.

The paper is organized as follows. In section 2, different models of emotions are reviewed in order to define the different components and variables present in our robot emotional system. The integration within a single dynamical model of both aspects of emotions (metacontrol and communication) and its interest for robotic systems will be discussed. We defend a constructivist approach [59, 85] to try to capture the minimal features allowing bootstrapping emergent behaviors (study of the

[*]E-mail: gaussier@ensea.fr
[1] This paper is a synthesis of different results presented in IROS and KEER 2010 conferences [22, 44]

co-development between the sensory-motor capabilities and the emotional capabilities on long term interactions). In section 3, we consider a mobile robot that must satisfy two different drives (i.e. simulation of water and food requirement) and uses two different kinds of strategies based either on visual or odometric information to reach a given place. Water and food resources will be symbolized by small pieces of paper of different colors (only visible by an ad hoc sensor put below the robot). The robot will be able to choose a given goal (the nearest place) or an alternative solution when the desired resource is unreachable. A simple self-monitoring system will be introduced to compute a "frustration" level when one goal cannot be satisfied. This self-evaluation will be used in a meta-controller to inhibit the current strategy, goal or drive and allow the robot to behave in an autonomous way. Finally, in section 4, this metacontrol mechanism is coupled with some simple interaction capabilities. The display of the robot internal state will be used to allow the robot to recognize the human facial expression first and next to modify the robot behavior according to human partner expression (building of a social referencing). We will conclude on the importance of emotions as a dynamical system (see fig. 2) where an emotional state is an area in the physical and social dimensions of the system.
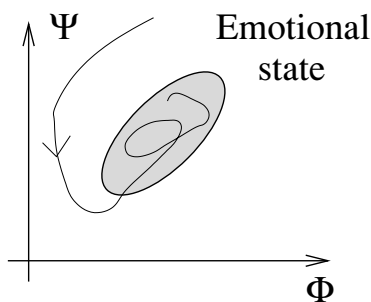

Figure 2. *Illustration of what could be an emotional state as a dynamical phenomenon linking physical ϕ and social interactions ψ.*

2. Drives, Self monitoring and emotions

Although many researchers agree that emotions involve "physiological arousal, expressive behaviors, and conscious experience" [64], or that emotions are important for survival [26, 28, 54, 55], there is clearly no agreements on the underlying mechanisms. For instance, from James and Lange theory [47, 53], which considers emotions as direct consequences of physiological modifications in reaction to the interactions with the environment (peripheralist theory: the emotional state is the recognition of a given physiological state) to the Cannon-Bard theory [10, 24], which supports that emotion is the result of a brain processing (centralist theory: physiological changes are the results of the triggering in the brain of a given emotional state), there is a wide spectrum of models, mostly dedicated to address only one aspect of emotions. For instance, if we focus on emotion expression then the opposition will be between discrete models of emotions (FACS: Facial Action Coding System [31]) versus dimensional/continuous models of emotions that suppose any emotion may by expresses as a point in a low dimensional space. Controlling the value of the different parameters would allow a continuous move from one expression to another [77]. These models are very appealing for engineers designing avatars or expressive robots since the models provide a way to control, in an independent way, the mood (from negative to positive values) and its intensity for instance.

Yet, there is no agreement on what must represent the different axis (usually three or four dimensions) to obtain a really coherent model. Sometimes, problems can arise when moving continuously from one expression to another: the agent displays some non-existent expression or provides the feeling that it was not the natural way to move from one expression to the next. At the opposite, the FACS supports the idea of specific activation patterns are independent from social culture [46] and is mostly used for facial expression analysis and to account for any muscular configuration of the face.

At the neurobiological levels, the models focus more on a particular emotion. For instance, the brain circuitry of pain [54] is much more detailed than the circuitry of happiness. This is certainly due to the fact that pain signals are managed by particular neurons and their associated neuronal circuitry. An important point here is the fact that emotions can be related to extrinsic signals such as pain or pleasure while other ones seem to rely much more on intrinsic variables like novelty for surprise. In this case, there is no specific input related to the surprise. The surprise can only be characterized by the inability to predict the current system state from the previous states. Hence, it appears that even what Ekman and Izard considered as basic emotions are perhaps the results of a complex process involving both a combination of extrinsic and intrinsic variables. Today emotional models, combining appraisal theory [7, 80] and arousal theory [79] offer a means to combine both physiological and cognitive components of emotions and show that one of the difficulty is also a vocabulary problem. If we consider the affects as hardwired or preprogrammed biological mechanisms that can be either positive (interest, excitation, satisfaction, joy) or neutral (surprise, novelty) or negative (hunger, fear, shame, disgust...) then we see the same vocabulary can be used to describe those affects and the emotions.

Starting from the neurobiological substrate of the visceral brain [73] (with the regulation loop connecting the thalamus, the hypothalamus, the hippocampus and the cingular cortex), we would like to understand how basic emotions [75, 83] can emerge and become complex cognitive processes involving planning and inhibition of action [26, 42]. From this literature[1–3, 12, 18, 27, 45, 70–72, 87], we know that a lot of structures are involved even for the "basic" emotions. Yet, physical and social interactions are certainly not governed by independent controllers and must share some common substructures.

Following the animat approach, we start from a minimal homeostatic regulator simulating physiological variables like hydration or glucose levels. These variables levels constantly decrease as the robot consumes its internal resources. It follows that collecting a simulated resource (i.e. detecting a needed resource) results in an increase of the corresponding resource level. However, the robot's survival is only possible if it collects periodically the resources it needs so their levels do not decrease below a given critical threshold (simulated death). A low-level drive system reacts to the physiological state perception. For instance, as food level gets low, the hunger drive gets high. This physiological and drive system is what gives a goal to the robot. A distinction is made between the inner drives, drives as they are computed directly from the physiological variables levels, and integrated drives, temporal integration of the inner drives. The integrated drives offer the possibility to modulate drives according to higher order source of information without manipulating the actual physiological state of the system. The most active drive dictates the robot's behavior (competition mechanism). When a needed resource is detected, the corresponding physiological variable level increases (following the above equations) and the temporal integration of the corresponding drive is reset to 0. Figure 3 describes this system.

In our case, an ad hoc sensor under the robot detects the presence of a resource (red and blue pieces of papers pasted on the floor to symbolize food or water) and induces the simulation of an eating
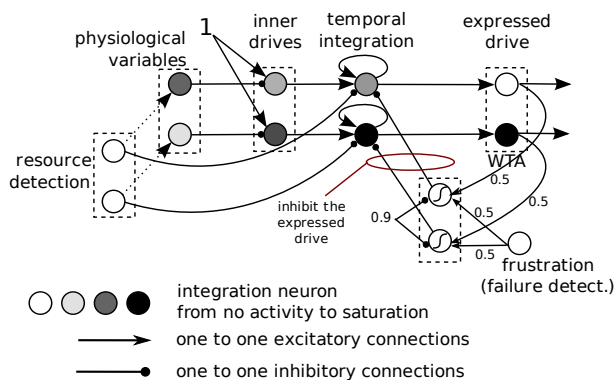
Figure 3. *Low level drive system : inner drives are computed from the physiological variables levels, integrated drives are signals that can be manipulated without affecting the inner states of the system and the expressed drive is the most active integrated drive. Frustration inhibition is described in section IV.*

or drinking behavior. The height of the camera used for the visual navigation guaranties the pieces of papers on the floor are not visible from the robot ensuring the need for the robot to learn a visual strategy to come back to the different sources (no direct goal detection is possible from a remote place). According to the desired task, it should be possible to change those basic sensors in order to guaranty the algorithm can adapt to a large variety of applications.

Now, if we want to deal with both emotions as a meta-controller and as a communication tool, one way to formalize the interactions between two agents is to differentiate two virtual channels even if they rely on the same physical channels (see fig. 4). The first one corresponds to the physical interactions with the environment such as object manipulation or fighting another animal. The second one concerns the social interactions and more specifically in our case the emotional interactions (detecting the fear associated to the flight behavior from the visual or auditory stimuli for instance).
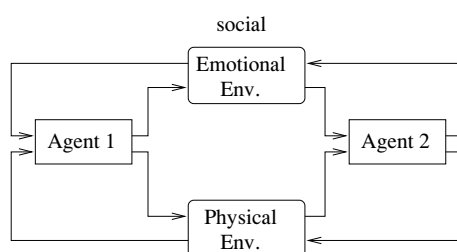

Figure 4. *Non verbal interactions between 2 agents can be either related to physical interactions (for instance object manipulation, displacements...) or to some social/emotional exchanges (for instance facial expression or body language)*

The interaction in itself has an emotional value. The appraisal of a given situation [56] can either be related to the evaluation of the physical interaction (capability to predict the interest of the current state or action) or to the evaluation of the emotional interaction. The evaluation of a physical interaction is an input of the social/emotional system: produc-

tion of a sound linked to the danger, the distress... or display of facial expression or other morphological modification that can be correctly interpreted by the other agents. Conversely, the evaluation of the social interaction can be used to modulate the parameters of the controller devoted to the physical interactions: fear perception can modulate the responsiveness to external stimuli and reduce the reaction time for instance. At another level, a human adult doing a still-face [65, 84] has a very negative effect on infants. Other studies show that desynchronized "interactions" are also associated with a negative feeling while online interactions have a positive value (a reward). In the case of mother-baby face to face interactions, a double video system has been used to control the mother-baby interactions [63, 66]. The results show that the introduction of a temporal delay disrupts the baby interest in her mother. Contingency is essential to maintain the interaction and imitation games between young infants have more a hedonistic value than a learning value. A lot of examples show that children enjoy imitating each other doing already known actions. The pleasure seems to be linked to the fact of being imitated by the other and doing some unaffordant (or unusual) behavior. From all theses different facts, we can conclude that the rhythm and the synchrony are important elements for the interaction [4, 5]. At the opposite end to reinforcement/punishment learning which is well studied in the frame of the learning theory, very few works focus on how the analysis of the interaction by the agent itself can be used to build an internal reinforcement. We believe that using the interaction as a way to self-generate a reinforcement/punishment signal is an interesting paradigm for online learning in a cooperative situation. This should allow building robots that could develop new skills in an open-ended perspective.

Our preliminary conclusion is that physical and social interactions are certainly not governed by independent controllers and must share some substructures. It can be interesting to try to complete some of the simplest existing models of both aspects of emotions in order to test the possibility to build a simple global model. In the following, we propose a very simple integrated model allowing focusing on how to obtain at least one global coherent dynamics taking into account the main neurobiological and psychological data available. For an autonomous robot, we suppose that a pain signal can be related to ad hoc receptors sensitive to the lack of resources (lack of food or water) and the collision with obstacles while pleasure can be associated to the refueling of necessary resources. Yet, to avoid the trap of using only ad hoc physiological signals, non-modal emotional signals must be introduced.

3. Self monitoring and meta control for navigation: a model of frustration

In this section, we first summarize two novelty detection mechanisms that can bootstrap surprise feeling and then we propose a method to measure the system frustration when it fails to reach a given goal. In previous works, we used two different mechanisms for novelty detection. First, in each sensory modality, novelty can be seen as a recognition threshold. If a given pattern is different enough from previously learned or stored pattern then vigilance can be increased and allow the learning of the new situation (see for instance Carpenter and Grossberg's vigilance parameter in their ART model [25]). Next, novelty can be a precise configuration of local categories (patterning). In more complex cases, the states can be already known but their sequences or timing may vary. The inability to predict the timing of sensory-motor events can be used to detect novelty and to modulate learning also[9, 41, 43] in order to increase the system efficiency.

Yet, as a given task or drive is concerned, states can be correctly recognized but the behavior can fail because of some deadlocks or dynam-

ical environment changes. To regulate the robot behavior in case of persistent failures, we propose a generic frustration mechanism based on the evaluation/monitoring of an unlimited number of signals, either drives or goals or even strategies. The robot's navigation abilities are based on a bio-inspired learning system : the PerAc architecture [36]. This architecture allows the robot to learn the conditioning of an action by a sensory input in order to define a dynamical perception state. More precisely, the robot's navigation system is derived from a model of the rat hippocampus [48]. It consists of a simulated neural network able to learn to characterize (and thus recognize) different "places" of its environment using place cells i.e. neurons that code information about the location of visual cues of the environment from of a specific place in that environment [36, 39]. The activity of the different place cells depends on the level of the associated visual cues recognition (landmarks) and of their location (azimuth). A place cell will then be more and more active as the robot gets closer to its learning location[2]. The area where a given places cell is the most active is called its place field. A conditioning neural network enables the learning of the association[38, 40, 89] between a place field and an action. In parallel, path integration is computed from odometric information [32, 62] (return vector computing). Both navigation strategies are coupled to a low level motivational system (using the simulated physiology as input) in order to perform a survival task (for more details see the appendix).

To test the architecture, different perturbations have been successfully introduced: unreachable resources, no more visual navigation (all lights in the environment turned off), wrong hodometric information after the robot has been "kidnapped" and placed in another place, ... All these perturbations might get the robot trapped in a deadlock situation. Here, we focus on two particular cases. The first case is the self-discovery of the failure of a rule (or sensory-motor association) in a new situation and the learning of a new context (and the inhibition of the problematic rule). The discovery of the solution would induce the learning of a new context and a new sensory-motor rule allowing satisfying a motivation. The key point is the capability to monitor the evolution of the satisfaction or unsatisfaction of some "drives" or "motivations" in order to control the learning. The second case concerns the problem of the satisfaction of conflicting goals. In the case of a single autonomous agent, we have shown the possibility on the long term to discover and learn a solution based on the building and use of a cognitive map [34]. Yet, if the survival of the agent implies to find quickly a solution or if two agents compete for the same resources [20] there is a need for a fast mechanism able to modulate the behavior in order to find quickly a stable solution for both agent (dynamics with a bifurcation point allowing each agent to choose a different solution). In both cases, the measure of the appraisal of the situation activates an appropriate facial expression on a robot head.

3.1. Frustration measure and meta-control

Using two different sources of information (vision and proprioception), the robot has access to two different ways of monitoring its goal distance. From the proprioception, the robot can monitor the fields used for path integration. Each field can be seen as a working memory and holds the information needed to represent the return vector to its corresponding goal i.e. its direction (position of the maximum activity in the field) and distance (value of the maximum activity). As the robot gets closer to the goal, the maximum activity of the corresponding path integration field gets lower while the corresponding place cell activity gets

higher. From vision, the robot can learn which place cell corresponds to the goal and then monitor its activity. When a drive is active (e.g. when it is hungry), until food is found, the robot might assume that everything is all right as long as its predicted distance to the food decreases. But if goal distance $G(t)$ does not decrease, the robot behavior is inefficient. And if this inefficiency is lasting this means the robot is caught in a deadlock and becomes frustrated. A binary frustration decision F(t) :

$$F(t) = \begin{cases} 1 & if \ [f(t)]^+ > T \\ 0 & otherwise \end{cases} \qquad (1)$$

can be achieved from a frustration level $f(t)$:

$$f(t) = [f(t - \Delta t) + R(t) - P(t) + \epsilon - r - F(t - \Delta t)]^+ \qquad (2)$$

computed as the temporal integration of the instantaneous progress $P(t) = \frac{1}{\tau}[G(t) - G(t - \Delta t)]^+$ and the instantaneous regress $R(t) = \frac{1}{\tau}[G(t - \Delta t) - G(t)]^+$ with Δt is the duration of each calculation timestep and the and and with $[x]^+$ equals x if $x > 0$ and equals 0 otherwise and τ is a time constant ($\tau = \Delta t$). ϵ is a small constant and r is a reset signal that equals 1 when the goal is satisfied (when the needed resource is detected) and 0 otherwise. The threshold T (figure 5) defines the robot tolerance to frustration. This mechanism differs from a simple timeout because frustration is increased by the number of failures and not directly by the elapsed time. According to this view, solving a long problem should not be frustrating as long as progress can be perceived. Furthermore, the frustration increase is not necessarily regular since it relies on how much goal proximity approximation varies. Detection of this failure situation gives the robot a way to escape from inefficient repetitive behavior.
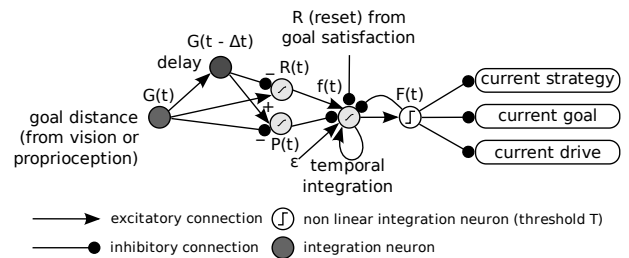

Figure 5. *Frustration mechanism: non linear integration of the goal distance derivative over time. When the goal distance comes from vision, its temporal difference is computed the opposite way as when goal distance comes from proprioception (goal place cell activity increases while integration field maximum activity decreases). A small constant input added to this integration insures that although goal distance is constant, frustration might arise. Above a definite threshold T, the active strategy, goal or drive is inhibited.*

The simplest way to escape a deadlock is to use failure detection to inhibit the underlying behavior. But there are many ways to alter the robot behavior. Failure detection might inhibit the currently used navigation strategy e.g. switching from path integration to visual navigation. But it can equally inhibit the active goal to look for another similar goal. Failure detection can also inhibit the active drive e.g. switching from hunger to thirst. An example of this inhibition is shown in figure 3 but the same kind of inhibition allows to switch from the active strategy or goal. Figure 5 shows the neural network used to detect failure situations and the way it can regulate the robot behavior.

---

[2] The details of the visual navigation architecture are described in the appendix

3.2. Experiments involving a meta-control

In the first experiment[3], the effect of the frustration regulation is tested on the drives. The visual navigation strategy is used in an environment containing one of each resource (colored square on the ground). After having learned to reach the two resources, the robot alternates between them according to its drive system. If an obstacle is put on one of the resources, the robot cannot access it. According to its drive system, the winning drive gets stronger with time and the robot should be stuck between going to the resource and avoiding the obstacle. When the frustration system is introduced, the robot gets more and more "frustrated" and inhibits the active drive allowing the robot to escape the deadlock to satisfy its other drive. Figure 6 shows the robot trajectories as well as its internal drive, failure detection and frustration signals.
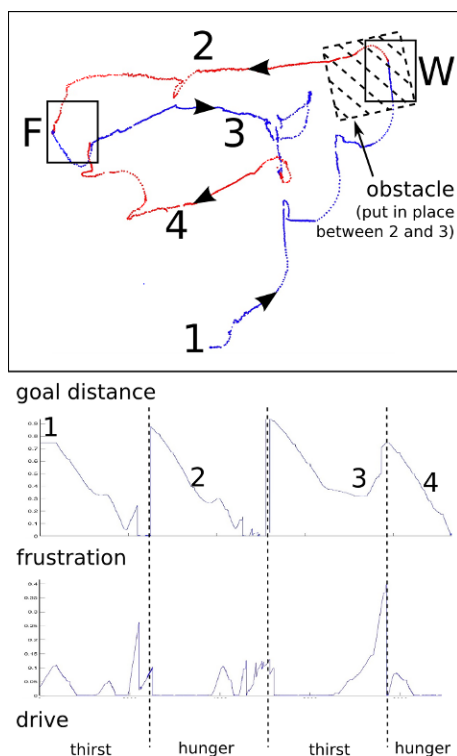
level. The proprioceptive navigation strategy is used in an environment containing two of each resource (2 goals for each drive). After having learned to reach the four resources, the robot alternates between the two closest goal places according to its drive system (determine the active drive) and its motor working memory [44] (determine the closest goal). Similar to the first experiment, an obstacle is placed under one of the resources the robot regularly use. The inhibition of the active goal allows the robot to escape from the deadlock to look for the other resource corresponding to the active drive. Figure 7 shows the robot trajectories as well as its internal goals, failure detection and frustration signals.
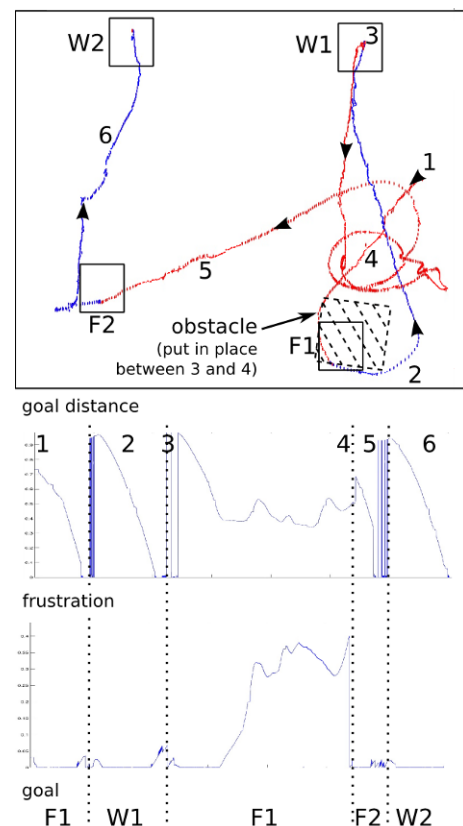

Figure 6. Up: robot trajectories with frustration of the active drive (visual navigation). F stands for food and W for water. Down: goal distance, failure detection and drive signals. In 1, the robot starts the experiment with thirst as the most active drive. In 2, the robot satisfies its thirst and hunger becomes the active drive. In 3, an obstacle obstructs the water resource. When enough failure detection has been integrated, a frustration inhibition is sent to the active drive (thirst) and in 4, the robot goes back the food location.

Figure 7. Up: robot trajectories with frustration of the active goal (proprioceptive navigation). F1 and F2 are the two food resources and W1 and W2 are the two water resources. Down: goal distance, frustration and goals signals. In 1, the robot starts the experiment with hunger as the active drive and its goal is F1, the closest food location. In 2, after reaching F1, the robot is now thirsty. It is reaching W1, the closest water location. In 3, the robot finds water and is now hungry. It is heading toward F1, the closest food location. F1 is now obstructed with an obstacle. In 4, when enough failure detection has been integrated, a frustration inhibition is sent to the active goal F1. In 5, the robot heads toward F2, the new closest goal satisfying the active drive. And in 6, the robot heads toward W2, the new closest water location.

In the second experiment, frustration regulation is tested on the goal

In the third experiment, frustration regulation is applied to strategy selection. Both path integration and visual navigation are used in an environment containing one of each resource. After having learned to reach the two resources with each strategy, the robot uses the proprio-

---

[3] Trajectories in all experiments are recorded from an onboard tracking device that is not used by the robot.

ceptive strategy to alternate between each resource. Next, the robot is "kidnapped" and placed in a different place of its environment. Because this movement cannot be integrated by the proprioceptive strategy, the return vectors all become erroneous. The robot then converges toward a wrong location. Inhibition of the active strategy allows the robot to switch from its proprioceptive to the visual navigation strategy that is robust to this kind of perturbation. Similarly, proprioceptive navigation is a good way to navigate in the dark thus offering a good alternative to visual navigation. Figure 8 shows the robot trajectories as well as its internal strategies, failure detection and frustration signals.

goal distance

frustration
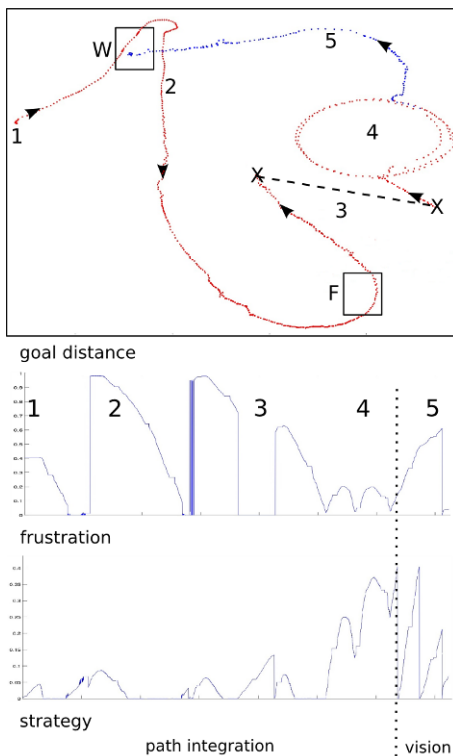
strategy

path integration   vision

Figure 8.   Up: robot trajectories with frustration of the active strategy. Down: goal distance, frustration and strategy signals. In 1, the robot starts the experiment with path integration. The active drive is thirst. In 2, after having satisfied its thirst, hunger becomes the active drive and the robot heads toward the food location. In 3, after having satisfied its hunger, the robot is thirsty and head toward the water location but the robot is "kidnapped" along the way and put somewhere else. This makes its proprioceptive strategy wrong. In 4, the robot follows its path integration until enough failure detection has been integrated. A frustration inhibition is sent to path integration strategy and in 5, the robot switch to visual navigation.

To summarize, the robot can learn evaluations of the distance to its goal from its different perceptions. Behavior effectiveness is viewed in terms of reduction of the goal distance. Accumulation over time of the inability to reduce goal distance (and reach satisfaction) gives rise to an inhibition potential that can be directed on different parts of the robot control architecture : the used strategy, the active goal or the active drive. This generic inhibition mechanism and the behavioral change it causes can be viewed as an emotional regulation: i.e. the frustration. Using a meta-control regulatory mechanism, the robot adapts its behavior to changing conditions rather than getting stuck in a deadlock situation

where its learning is not sufficient. It is clear the empirical frustration regulation mechanism described here could be easily refined. In order to allow failure detection to be robust to noise on the goal distance prediction (mainly concerning vision), we intend to use a statistical version of the proposed equation in future works. Yet, it is sufficient for monitoring the progress and allowing an efficient way to react to changing conditions of a dynamical environment.

The frustration associated to the robot strategies, goals or drives (through classical conditioning) can be seen as a prediction of the robot success or failure for this particular strategy, goal or drive and can then be used to select them accordingly. Our model thus bears strong similarities with TD lambda [29, 82] and the possibilities of hedonist neurons [52]. Our frustration regulation can also be compared to the novelty detection and curiosity mechanisms described by [69]. While curiosity regulates the robot behavior in order to stay in a state of learning progress, frustration regulates the behavior in order to stay out of failure states.

4.   Emotional interactions and social referencing

In the previous section, we illustrated how some internal prediction mechanisms can be used for a self-monitoring and for modifying the robot behavior. They can be seen as basic emotional mechanisms even if one can discuss the reality of these signals for the system or more exactly the capability of the robot to perceive these emotional states. It is clear in the previous architecture that nothing has been introduced to allow the categorization and the recognition of an emotional state. Yet these mechanisms are sufficient at least to trigger some reflex expressive behavior. Here, we focus on what can bring the expressiveness or the communicative function of emotions. We discuss two complementary aspects. First, how can a robot or a baby learn to recognize in an autonomous way the facial expression of a human caregiver? Second, we show that a new object or place oriented behaviors can be learnt thanks to the emotional interactions. They allow to close the loop between the meta-control and the communicative function of emotion in a triadic system involving one robot, one human and one object or place allowing to bootstrap some kind of low level social referencing.

Figure 9.   Experimental set-up: a robotic head that learns facial expression recognition and a mobile robot able of autonomous visual navigation tasks learning. The room size is 7m x 7m, but the robot's movements are restricted to an area of 3m x 3m (to allow a good video tracking of the robot's trajectories).

Our experiments rely on two major systems : an emotional facial expressions interaction system that gives a robotic head the ability to learn

to recognize and mimic emotional facial expressions and a navigation system that gives a mobile robot the ability to learn navigation tasks such as path following or multiple resources satisfaction problems (see fig. 9).

4.1. Learning to recognize facial expression

To understand better how the coupling between the cognitive and emotional capabilities co-develop, first we tried to model how babies can learn to recognize facial expressions of their parents without having a teaching signal allowing them to associate for instance an "happy face" with their own internal emotional state of happiness [37]. In a robotic viewpoint the question becomes us to close how a robotic system (fig. 9), able to exhibit a set of emotional expressions, can learn autonomously to associate expressions with those of others. Here, "autonomously" will refer to the ability to learn without the use of any external supervision. A robot with this property could therefore be able to associate its expressions with those of others, linking intuitively its behaviors with the responses of the others. Using the cognitive system algebra [35], we showed that a simple sensory-motor architecture (figure 10) using a classical conditioning paradigm could solve the task if we suppose that the baby produces first facial expressions according to his/her internal emotional state and that next the parents imitate the facial expression of their baby allowing in return the baby to associate these expressions with his/her internal state [35]. Moreover, psychological experiments [30, 67] have shown that humans reproduce involuntary a facial expression when observing it and trying to recognize it. Interestingly, this facial response has also been observed in presence of our robotic head. This low-level resonance to the facial expression of the other can be considered as a natural bootstrap for the baby learning ("empathy" from the parents). Because the agent representing the baby must not be explicitly supervised, a simple solution is to suppose the agent representing the parent is nothing more than a mirror. We obtain an architecture allowing the robot to learn the "internal state"-"facial expression" associations. We also showed that, learning autonomously to recognize a face could be really more complex than to recognize a facial expression. We proposed an architecture (figure 10) using the rhythm of the interaction to allow first a robust learning of the facial expression without a face tracking [15], and second, to stop the learning when the visual stimuli (facial expression or absence of face) are not synchronized with the robot facial expression.

We have experimentally verified that a robot can learn to recognize the facial expressions of a human without any supervised learning. Basically, the robot produces facial expressions according to its own internal state and associates each perceived stimulus with this state. After some time, the robot is able to learn if the objects in its environment are not correlated with any of its emotional states. Conversely, if a human shows some empathy to the robot he/she may produce correlated facial expressions (see [67]) that will be recognized and associated to the robot state. Later, this learning allows the recognition of the associated emotional state. Yet, we have to face several problems. First, the delay for the human (and robot) to recognize the change in the facial expression of the other and the motor delay to produce a facial expression introduces complex transitory states that the neural network has to filter. Second, to avoid long learning, it is important that the robot modulates its learning according to the fact there is something interacting with its own activity or not (for instance a human mimicking the robot facial expressions).

After two minutes of real time interaction, the robot is able to recognize the human facial expression as well as to mimic its facial expressions[13]. Fig. 11 shows the success rate for each facial expression (sadness, happiness, anger, surprise) and a neutral face. These results are obtained during the natural interaction with the robot head.
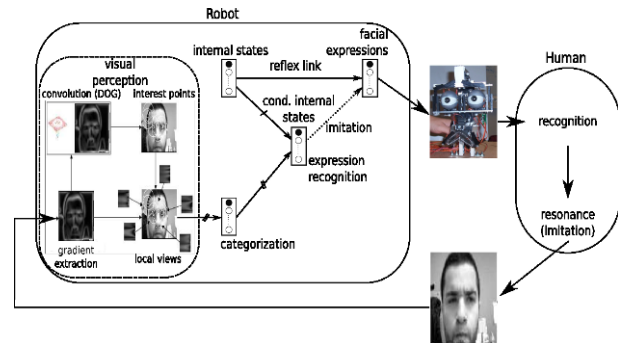

Figure 10. Architecture used to associate a collection of local views around feature points extracted from the visual flow with the expressed emotion by the robot. If a human comes in front of the robot and imitates the robot's expressions, (s)he will close the loop between vision and proprioception and allows the system to learn to recognize the facial expression of the human partner.

10 persons interacted with the robot head (32 images by facial expression by person). During the learning phase, each subject interacted during a 2 minutes period.
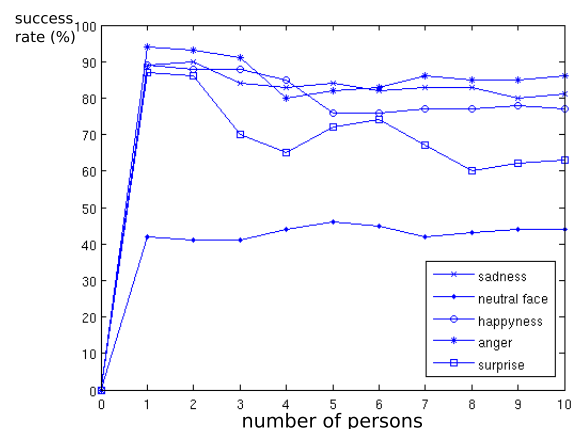

Figure 11. Success rate for each facial expression (sadness, neutral face, happiness, anger, surprise). These results are obtained during the natural interaction with the robot head. 10 persons interacted with the robot head (32 images by facial expression by person). During the learning phase (only a 2 minutes period), these humans imitate the robot, and then the robot imitates them. In order to build statistics, each image was annotated with the response of the robot head. The annotated images were analyzed and the correct correspondence was checked.

One interesting result is that the face detection system usually used for preprocessing facial expressions recognition is not really necessary and can even be the result of the facial expression recognition. This inversion in the classical way of learning to recognize facial expressions with a robot head [68], would allow to show the power of the emotional system to shape the individual development through the interaction with another agent. Practically, it is also very interesting to suppress the need to first detect the head since in our previous systems we

were unable to use an autonomous learning for this step (the learning was supervised since we had no way to find an autonomous criteria to decide what is a face or not). In our architecture, the autonomous recognition of face / non-face discrimination results from the facial expression recognition. A human face is recognized as such because his/her local views are associated to the emotion recognition and not the opposite.

4.2. Social referencing for places and objects

In this section, we try to verify the postulate that in a social environment, the emotion communication participates in the shaping and the triggering of more and more complex behaviors. Usually, robots which are able to learn navigation tasks, are taught under supervision of an experimenter [19, 78]. These techniques have the advantage of being fast in terms of learning time but the experimenter has to know exactly how the robot works and to be expert in order to use it. In other words, the experimenter has to strongly adapt itself to the robot's underlying architecture to achieve satisfactory learning performances. The autonomy of a mobile robot can be more easily reached if the robot has the ability to learn through emotional interactions. The social referencing is a concept issued from developmental psychology describing the ability to recognize, understand, respond to and alter a behavior in response to the emotional expressions of a social partner [51, 76, 86]. Besides, being non verbal and thus not needing high-level cognitive abilities, gathering information through emotional interactions seems to be a fast and efficient way to trigger learning at the early stages of human cognitive development (compared to stand alone learning). Even not at their full extent, these abilities might provide the robot valuable information concerning its environment and the outcome of its behaviors (e.g. signaling good actions). In that case, the simple sensory-motor associations controlling the robot's learning are defined throughout their interactions with the experimenter. This interactive learning does not rely on the experimentor's technical expertise, but on his/her ability to react emotionally to the robot's behavior in its environment (both human and robot have to adapt reciprocally with each other).

Social referencing can refer to an object, a person, an action, a place in the environment and probably different other things. This means that there are many ways for the recognition of an emotional facial expression to be interpreted and used by the navigation system. In our case, when the experimenter displays an expression of happiness, the robot can use this expression as a signal qualifying its behavior. In that case, its action in a specific place must be learned as having a positive value. But the robot could also use this signal to qualify its surrounding environment indicating a useful place that the robot should eventually seek. We studied these two different possible couplings between the navigation and the emotional interaction inside our architecture. We think this approach can be useful for the design of interacting robots and more generally, for the design of natural and efficient human-machine interfaces. Moreover, this approach provides new interesting insights about how, in their early age, humans can develop social referencing capabilities from simple sensory-motor dynamics.

The behavioral coupling refers to the situation where the recognition of an emotional facial expression is used to qualify the behavior of the robot. For instance, when the human displays a happy face, it means the robot must reinforce its current behavior positively while an angry face means the robot must reinforce its current behavior negatively. In order to do this, the PerAc architecture [33, 36] learns positive and negative action conditionings. To ensure this classical conditioning, the least mean square learning rule [88] is used. The difference between the neural network output and the desired output is used to compute the amount by which the connections weights have to be changed

(weight adaptation due to learning):

$$\Delta w_{ij} = \epsilon . I_i . (O_j^d - O_j) \qquad (3)$$

Δw is the difference between the old and the new weight, ϵ is the learning rate (neuromodulation of the network), I is the input, O is the output (of the conditioning network) and O^d the desired output. A positive conditioning refers to a direction to head for (to reach the goal), while a negative conditioning refers to a direction to inhibit (to avoid a dangerous place). Instead of one sensory-motor neural network that can only learn positive conditionings, we used one associative neural network for positive and one for negative conditionings. A third group of neurons is used to compute the sum of their two outputs (see figure 12).
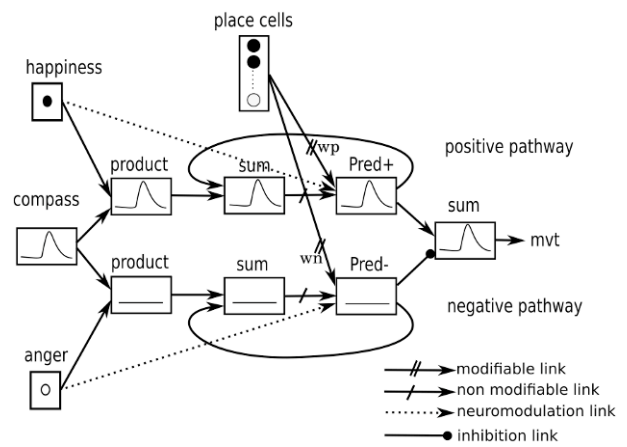

Figure 12. *Behavioral coupling model. When one of the "conditioning" groups of neurons using equation 3 receives neuromodulation from the recognition of the corresponding facial expression (happiness in this example), it learns the association between the current robot location (perceived as a specific winning place cell) and its direction (summed with what has already been learned by this group of neurons). Happiness and Anger are neurons associated to the recognition of an happy or an angry human face.*

While the positive conditioning group of neurons has a positive connection with this integration group of neurons (activations), the negative conditioning group of neurons has a negative connection (inhibition). This solution allows much more information to be stored about what is learned by the robot than outputs with positive or negative values (and is also more biologically plausible). For instance, having learned that one particular behavior is good and later that the same behavior is wrong could mean that something has changed in the nature of the environment or in the experimenter's objectives. If both reinforcements had been learned on the same group of neurons, they would have been averaged and the conflictual nature of the learning would be invisible. The model is described in figure 12. When the robot receives a social interaction signal (the display of an emotional facial expression of anger or happiness), it triggers the learning of a new visual place cell as well as the learning of the conditioning between this visual place cells and the current action. Nonetheless, if an existing place cell is too close to the robot current position (defined by a threshold on the place cells recognition level) the learning of a new place cell is inhibited and the sensory-motor conditioning is learned according to the nearest place, completing an eventually previously learned sensory-motor conditioning.
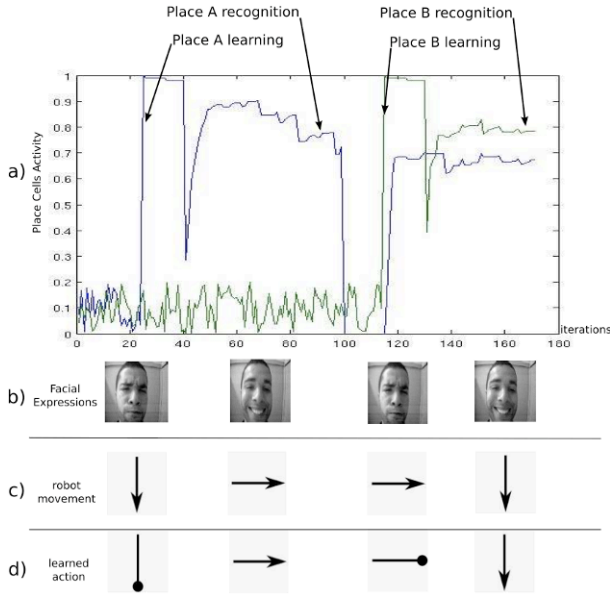
Figure 13. *a) Place cell signal. b) The experimenter's facial expressions recognized by the robot. c) Current robot direction of movement. d) Action learned by the robot (an arrow means a direction to activate and a dot a direction to inhibit). The experimenter facial expressions give the robot the information needed about its behavior to learn the necessary sensory-motor associations between the visual signal (recognition of the current place) and the learning of the activation or inhibition the current movement direction.*

The robot is thus able to learn progressively which direction to avoid and which direction to head at a given "place" according to the goal of the person interacting with it. We tested this architecture in the following situation: the robot's environment contains one place of interest and the experimenter wants to teach the robot how to reach it. Each time the experimenter thinks the robot's behavior is wrong, he expresses anger toward the robotic head and, conversely, he smiles for good behaviors (happiness). Figure 13 is an illustration of the learning chronology (as explained above). Figure 14 shows the robot's trajectories after learning. The robot is dropped from different positions of the environment. It is always able to reach the interesting place. Nevertheless, it is important to take into account the fact that the robot learns much more information about the task when its behavior is qualified as "good" by the experimenter than when it is qualified as bad (although both are needed). Knowing what is "good" is a faster way to converge to a solution than knowing what is "bad". The learning of the attraction basin around the goal place (i.e. set of place-actions that ensure a converging navigation dynamics) takes between three to five minutes.

If no reflex pathway is available, an instrumental conditioning can always work (and be superposed to the previous classical conditioning mechanism). When the robot receives the social interaction signal, it has to learn a new place cell characterizing its location and to learn to predict the interaction signal (happiness or anger) which is considered as a reward associated with this place (figure 15).

As the robot gets closer to the learned place, the place cell response will increase, such as the associated predicted reward. The opposite happens as the robot gets farther from the learned place. Instead of using a conditioning learning between a perception (a place) and an action (a direction), the derivative of the predicted reward (fig. 15) is
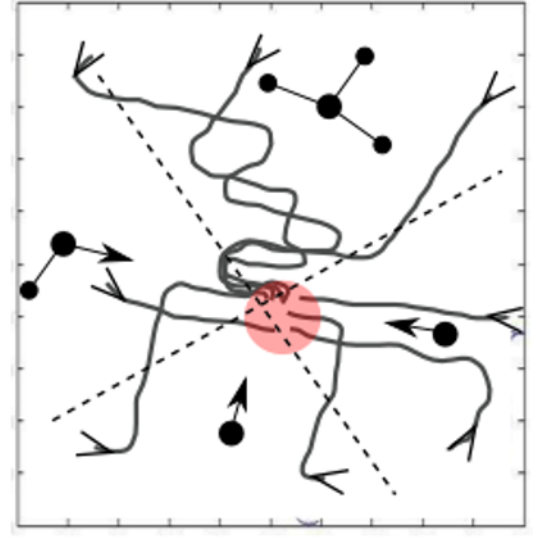

Figure 14. *Robot's trajectories from different starting points: the robot is able to reach the place associated with the happiness facial expression. The grey zone represents the goal place. These trajectories are obtained by video tracking. The size of the experimental area is 3m x 3m.*
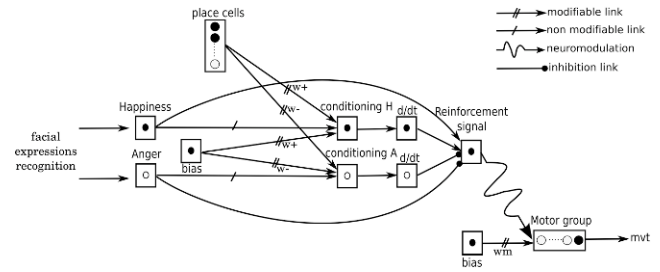

Figure 15. *Environmental coupling model. Using the least mean square learning rule, the conditioning neurons allow the association between a place cell (a zone of the environment) and an experimenter's facial expression (modifications of weights w+ and w- follow equation 1). The predicted expression signal temporal derivate is used as a reinforcement signal (Sutton and Barto learning rule) to maintain or change the direction on the motor group (modifications of weights w follow equation 3). The bias on the conditioning groups allows the learning of the frontier between the zone associated with a facial expression and the rest of the environment.*

used as a reinforcement signal [11] :

$$\Delta R \;=\; (\frac{dPredH}{dt} - \frac{dPredA}{dt}) + (H - A) \qquad (4)$$

$$\Delta w^{+/-} \;=\; \epsilon.\frac{dR}{dt}.\frac{dO_j}{dt}.I_j \qquad (5)$$

ΔR is the reinforcement signal, $\frac{dPredH}{dt}$ is the predicted happiness signal derivate, $\frac{dPredA}{dt}$ is the predicted anger signal derivate, H is the

happiness facial expression recognition signal and A is the anger facial recognition signal. $\Delta w^{+/-}$ is the difference between the old and the new weight, ϵ is the learning rate (neuromodulation of the network), $\frac{dR}{dt}$ is the temporal variation of the reward R, O is the output and I is the input. A motor group only connected to a constant input is used to control the robot movements. Without any reinforcement, this motor group basically produces random outputs (a small noise is added to the output) allowing the robot to "try" another action. A positive reinforcement will make it reinforce its current output while a negative reinforcement will make it inhibit its current output. We used the outputs to control the robot actions. After the robot has learned by interaction that the place at the center of its environment is dangerous (i.e. associated with the anger expression), we assigned various fixed directions to the robot in order to test the robot robustness of the robot learning. Figure 16 shows how directions that produce positive predicted reward derivative (going away from the dangerous place) are reinforced positively while directions that produce negative predicted reward derivative (going toward the dangerous place) are reinforced negatively.
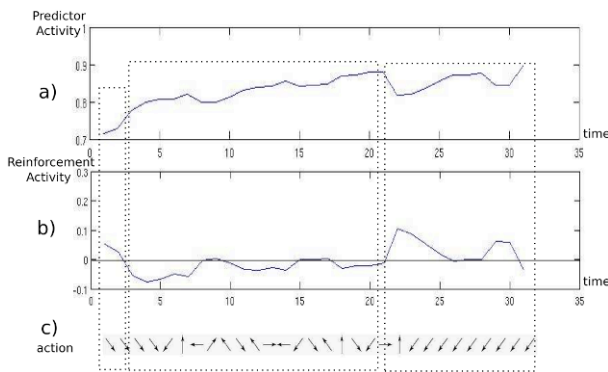

Figure 16. *a) The reward prediction (positive with happiness and negative with anger) informs the robot about its behavior outcome in the environment. b) Derivatives of this value are used as a reinforcement signal(see equation 5). c) when the derivative is negative, the robot direction changes and when it is positive it is maintained and reinforced*

Figure 17 shows the robot's trajectories from different starting points with different fixed directions while, at the same time, it has to avoid the dangerous place of its environment. The referencing of that place through interactions with the experimenter allows the robot to quickly learn to avoid it (the first interaction already allows the robot to avoid the "dangerous" place). Nevertheless, the task would be much more difficult if we wanted to teach the robot to reach one place instead of avoiding it. Indeed, avoiding a place needs to be efficient at the vicinity of the place in question. This is the role of the bias on the conditioning groups shown in figure 15. Reaching a place means being able to use variation of the corresponding place cell but far from the learning place. Yet, the place cells dynamics are not meaningful when the robot is too far away from the learning location.

One remaining difficulty is related to the intrinsic ambiguity of the emotional interaction signal. In our case, the same signal can be used to learn two different information: "this place is good" as well as "this place/action is good". A solution to this problem could be the way the system treats the interaction inputs. For a behavioral coupling (associating emotions to the robot's actions) a phasic signal (the moment the signal appears) should be used while for the environmental coupling (associating emotions to the robot's environment) a tonic signal (the whole time the signal is present) is sufficient. This way, both couplings
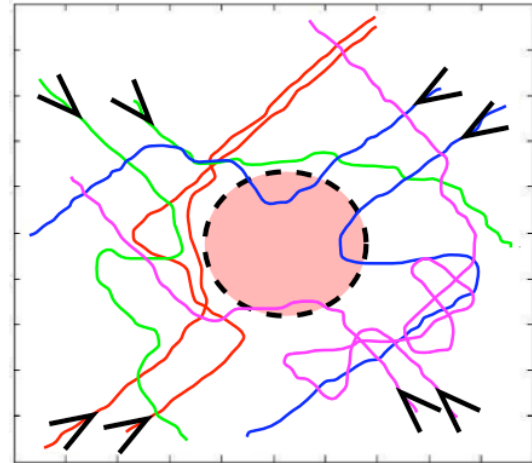

Figure 17. *Robot's trajectories from different starting points (with a fixed direction) after interactive learning of the association of the grey zone to the anger facial expression. The robot is able to avoid the place associated with the anger facial expression. The prediction of the negative reinforcement is sufficient to inhibit a movement in direction of the dangerous zone (when it is near it).*

could function with the same inputs but used differently. Of course, the question of the coherence of what is learned is asked : if the robot is doing something wrong (e.g. going away from a resource it needs) the experimenter will display an angry face and the robot will learn at the same time that its behavior was wrong but also that the place it is in has to be avoid. The problem is that, usually, the experimenter intended only one of the two learning. Nonetheless, because of the continuous nature of neural networks learning algorithms, the coherence of the learning should not be reached at the early stages of the interaction but rather for the more consistent ones. A place will have a well-defined emotional value (given by the social referencing) only if the reinforcement signal it receives is coherent over time.

5. Discussion and conclusion

In this paper, we have addressed three different aspects of the emotional mechanisms. First, we have shown, a simple architecture allowing a robot to self-monitor its success in a task completion, can be used to modify the robot drives, goals or strategies in order to avoid deadlocks. Next, we have summarized recent works performed with a robotic expressive head showing how a robot can learn to recognize the human facial expression of a person when he reproduces the robot own facial expressions. Finally, we have shown the coupling of both systems can be a simple way to teach a robot some arbitrary tasks like going or avoiding a given place. In other works not presented here [14], we have shown this strategy can be generalized to object reaching or to obstacle avoidance using a robot arm that can bootstrap some simple social referencing.

Following this, it is clear that the different parts of the proposed architecture can be easily improved by taking into account other emotional mechanisms [11, 61, 69, 81] or the robot expressiveness [16, 50, 90]. An important practical issue for a routine use of our system is that the facial expression recognition from a distance more than one meter needs the use of either a high resolution camera or a system with twin cam-
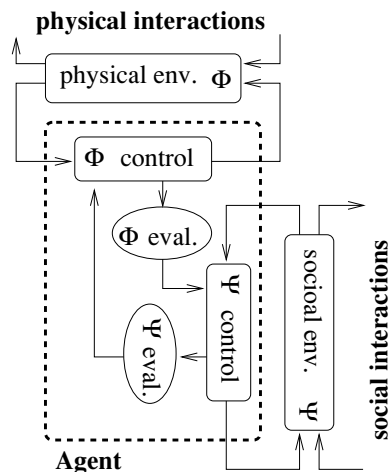
Figure 18. *Bidirectional interactions between abstract controllers devoted to the physical interactions and the social interactions. In the real systems they can be merged together but the interplay between both kinds of interactions is very important for the system autonomy and its capability to interact with other agents.*

Figure 19. *Illustration of the path integration computation. Left figure shows a simulated trajectory composed of two segments of different lengths (the first is three times the length of the second) and orientations (25 and 90 from an arbitrary absolute direction). Right figure shows the two inputs (dotted curves) as bell shape centered on the absolute direction of the movements (α and $\hat{\beta}$) and their sum (the bold curve).*

eras (one with a small field of view to focus on the face and one with a large field of view to find the partner in the room when he/she has moved). As a result it appears that a powerful attentional mechanism is necessary in such architectures to switch back and forth the attention between the navigation task and the human partner. Future work will focus on the need of a more realistic interaction where a bidirectional communication must exist between the human and the robot. The robot head can express the robot internal state and it can mirror the human facial expression. The problem is that currently, the robot head always mirrors the human facial expression to allow the experimenter to see that his/her mood has been well understood by the robot. Allowing a real interaction could provide a solution for expressing something related both to the expressive feedback of the experimenter and the robot's internal state. Control of the expression intensity and its duration is a lead we will explore. Moreover, some experiment have shown the difficulty in deciding which expression has to be displayed by the robot when globally it fails to solve its initial task but, because the meta-control succeeds, at least one goal can be satisfied. In this case, we choose it is better for the robot to express the more recent change in its emotional state (here for instance happiness while on the long term the robot has no solution to satisfy its primary need). A mechanism selecting the expression to be displayed according to some long-term reinforcement or planning would allow displaying an expression different from the robot internal state (faking an emotion or cheating) and would be necessary to transform the network producing the facial expression in a real communication device able to build and categorize complex emotional states.

Yet, a major question is: does the robot really feel the emotion or it is just an engineering trick? One easy answer could be to take a purely behavioral point of view and to consider the emotions even in the human perspective are nothing more than that. In the present state of the architecture, this answer is not satisfactory since we miss the capability to categorize new emotional situations. The emotional states are directly related to the emotional signals: pain would induce sadness, pleasure/happiness, surprise/surprise, and frustration/hanger. But what should be categorized as an emotional state? We believe the diffi-
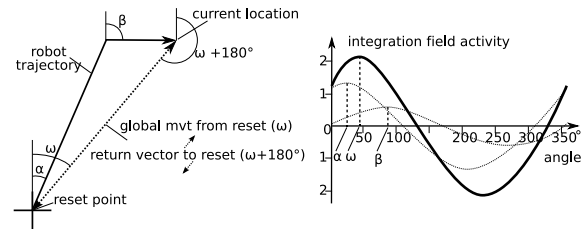
culty to define the emotions is certainly related to the feeling emotions could be defined as a static configuration of the different kind of internal variable. Fig. 18 proposes a simplified representation of the brain where we enlighten the interactions between two kinds of controllers: the controllers devoted to physical interactions and the controllers devoted to social interactions. Then, an emotion appears as something more than the result to a particular stimulation. The emotion can result from the interactions with another agent and from the interactions between different sub-controllers inside one agent. We can question the existence of a locus for the emotion even in the sense of a distributed network. The emotion may rely more on the network dynamics[4] than on the activation of particular neurons. Adding the capability to categorize such internal dynamical states could be a way to provide the robot a real perception or re-enaction of an emotional state.

Acknowledgments

Appendix

Navigation strategies

Two different navigation strategies were used in order to study inhibitory effects of the frustration regulation on the action selection process.

Proprioceptive navigation : path integration is the ability to determine the return vector (angle and distance) to an arbitrary reference

---

[4] By extension, the network can or perhaps should be extended to the network composed by the different agents in interaction.

point using odometric information. We designed a motor working memory [44] to use path integration implementation presented in [32] (figures 19. The global movement vector orientation (ω) is coded as the position of the maximum activity in a neural field and its norm is coded as the value of the maximum activity. Here, a neural field only means a group of neurons (no connections between each other) but the topology is important since position in the field has a meaning (an angle in our case). Fig. 20) shows the NN. used to compute the path integration and to propose a homing vector (inverse vector of the path integration). In the case of a multiple goal task, detection of a new goal allows the
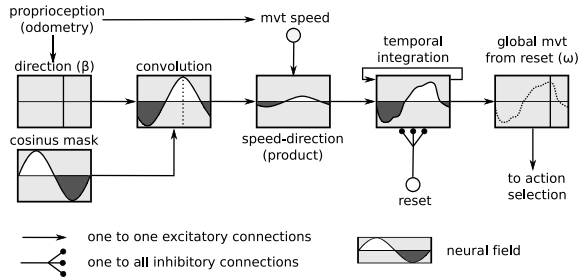

Figure 20. *Neural network for path integration : speed is coded as the activity of one neuron and orientation as the most active neuron of a field (i.e. a simple linear collection of neurons). At every time step, the integrator takes as input the activity of the orientation field (convoluted by a bell shape curve e.g. a Gaussian or a cosine) multiplied by the activity of the speed neuron. This input represents the orientation and distance traveled since the last time step. Summing this input with its own activity, the integration field computes the return vector.*

recruitment of a dedicated integration field. Every integration field computes dynamically the return vector to its associated goal (figure 21). A short term memory is used to store the relevant return vector. This model is fully described in [44]. Recruitment reset is the recruitment of a new integration field when a new goal is found. Recognition reset is the reset of the field corresponding to a detected known goal. Field selection is the selection of the integration field corresponding to the closest goal satisfying the active drive.
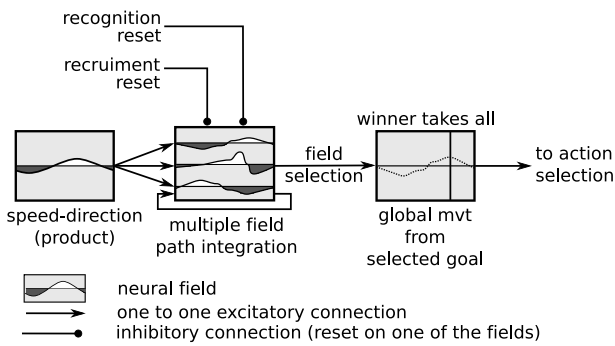

Figure 21. *Multi goals path integration navigation : return vectors to several places (goals) are computed dynamically.*

Visual navigation: The visual system learns place cells i.e. neurons that code information about a constellation of local views (visual cues) and their azimuths from of a specific place in that environment [38, 40] (see figure 22).
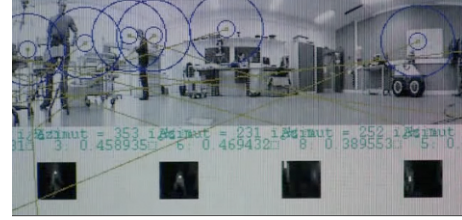

Figure 22. *Landmarks and their azimuths extracted from the raw visual flow and learned by the visual system.*
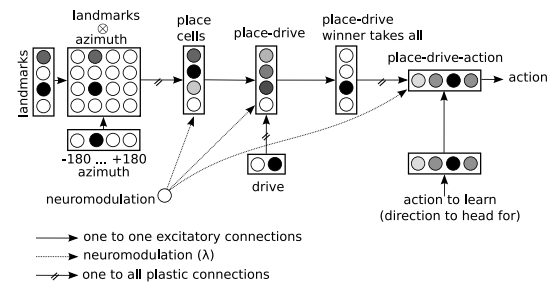

Figure 23. *Sensory-motor visual navigation: a visual place cell is constructed from recognition of a specific landmarks-azimuths pattern (tensorial product) and an action (the return vector) is associated with this place cell.*

Next, the network system is able to learn to characterize (and thus recognize) different places of the environment (see fig. 23). Activities of the different place cells depend on the recognition levels of these visual cues and of their locations. As shown in figure 24, a place cell will then be more and more active as the robot gets closer to its learning location.

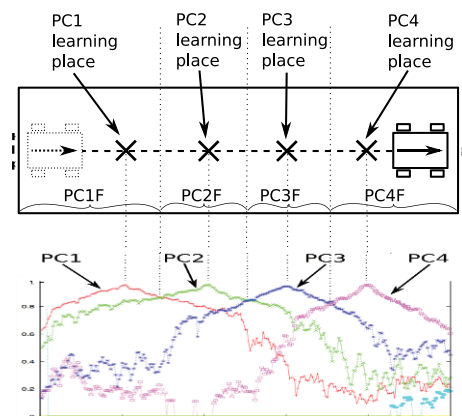The area where a given place cell is the more active is called its place

Figure 24. *As the robot gets closer to each place cell learning location, the corresponding place cell (PCn) gets more active. the maximum activity of a place cell corresponds to its learning location. And the area where a place cell activity is the highest is its place field (PCnF).*

field. An associative learning group of neurons allows sensory-motor learning (place-drive-action group on figure 23). Place-drive neurons learn the conditioning between place cells and drives (Hebbian learning). They are associated with the return vector of the corresponding goals to build a visual attraction basin around each goal.

References

[1] R. Adolphs, D. Tranel, and A. R. Damasio. The human amygdala in social judgment. *Nature*, 393(6684):470–474, Jun 1998.

[2] R. Adolphs, D. Tranel, H. Damasio, and A. R. Damasio. Fear and the human amygdala. *J Neurosci*, 15(9):5879–5891, Sep 1995.

[3] R. Adolphs, D. Tranel, S. Hamann, A. W. Young, A. J. Calder, E. A. Phelps, A. Anderson, G. P. Lee, and A. R. Damasio. Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia*, 37(10):1111–1117, Sep 1999.

[4] P. Andry, A. Blanchard, and P. Gaussier. Using the rhythm of nonverbal human–robot interaction as a signal for learning. *IEEE transactions on Autonomous Mental Development*, 3(1):30–42, 2011.

[5] P. Andry, P. Gaussier, S. Moga, J.P. Banquet, and J. Nadel. Learning and communication in imitation: An autonomous robot perspective. *IEEE transactions on Systems, Man and Cybernetics, Part A*, 31(5):431–444, 2001.

[6] R.C Arkin. *Behavior based robotics*. MIT press, 1998.

[7] M.B. Arnold. *Feelings and emotions : The mooseheart sympo- sium*, chapter An excitatory theory of emotion. Number 25. New York : Mc-Graw- Hill, 1950.

[8] B. Bakker and J. Schmidhuber. Hierarchical reinforcement learning based on subgoal discovery and subpolicy specialization. In *Proc. 8th Conference on Intelligent Autonomous Systems IAS-8*, 2004.

[9] J.P. Banquet, P. Gaussier, J.L. Contreras-Vidal, and Y. Burnod. A neural network model of memory, amnesia and cortico hippocampal interactions. In R. Park and D. Levin, editors, *Fundamentals of neural network modelling for neuropsychologists*, pages 77–121, Boston, 1998. MIT Press.

[10] P.A. Bard. A diencephalic mechanism for the expression of rage with special reference to the central nervous system. *American Journal of Physiology*, 84(24, 32):490–513, 1928.

[11] A.G. Barto, R.S. Sutton, and D.S. Brouwer. Associative search network : A reinforcement learning associative memory. *Biological cybernetics*, 40:201–211, 1981.

[12] T. Bosse, C. M Jonker, and J. Treur. Formalisation of damasio's theory of emotion, feeling and core consciousness. *Conscious Cogn*, Aug 2007.

[13] S. Boucenna, P. Gaussier, P. Andry, and L. Hafemeister. Learning to recognize facial expressions through an online mimicking game. *Pattern Analysis and MAchine Intelligence*, submitted, 2012.

[14] S. Boucenna, P. Gaussier, and L. Hafemeister. Development of joint attention and social referencing. In *Proc. IEEE Int Development and Learning (ICDL) Conf*, volume 2, pages 1–6, 2011.

[15] S. Boucenna, P. Gaussier, and J. Nadel. What should be taught first: the emotional expression or the face? In *8th International conference on Epigenetic Robotics*, 2008.

[16] C. Breazeal. *Designing Sociable Robots*. Cambridge MA: The MIT Press, 2002.

[17] R.A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, R.A. 2(1):14–23, March 1986.

[18] J. Burgdorf and J. Panksepp. The neurobiology of positive emotions. *Neurosci Biobehav Rev*, 30(2):173–187, 2006.

[19] S. Calinon, F. D'halluin, E. Sauser, D. Caldwell, and A. Billard. Learning and reproduction of gestures by imitation: An approach based on Hidden Markov Model and Gaussian Mixture Regression. *IEEE Robotics and Automation Magazine*, 17(2):44–54, 2010.

[20] L. Canamero. Emotions and adaptation in autonomous agents: A design perspective. *Cybernetics and Systems*, 32(5):507–529, 2001.

[21] L. Canamero and P. Gaussier. Emotion understanding: robots as tools and models. In J. Nadel and D. Muir, editors, *Emotional Development*, pages 235–258. New York: Oxfort University Press., 2005.

[22] L. Canamero, P. Gaussier, C. Hasson, and A. Hiolle. *Ouvrage transdisciplinaire sur les emotions*, chapter Emotion et cognition : les robots comme outils et modeles. Hermes, in press.

[23] Richard Canham, Alexander H. Jackson, and Andy Tyrrell. Robot error detection using an artificial immune system. *Evolvable Hardware, NASA/DoD Conference on*, 2003.

[24] W.B. Cannon. *Bodily changes in pain hunger fear and rage*. D.Appleton And Company, 1920.

[25] G.A. Carpenter and S. Grossberg. Invariant pattern recognition and recall by an attentive self-organizing art architecture in a non-stationary world. *Proceeding of Neural Network*, 2:737–745, 1987.

[26] A. Damasio. *Descartes' Error : Emotion Reason and the Human Brain*. Putnam Publishing, 1994.

[27] A. R. Damasio. Emotion in the perspective of an integrated nervous system. *Brain Res Brain Res Rev*, 26(2-3):83–86, May 1998.

[28] C. Darwin. *The Expression of the Emotions in Man and Animals*. John Murray (publisher), 1872.

[29] P. Dayan. Motivated reinforcement learning. In *Advances in Neural Information Processing Systems (MIT Press)*, 2001.

[30] U. Dimberg and M. Thunberg K. E. Unconscious facial reactions to emotional facial expressions. *Psychol Sci*, 11(1):86–89, 2000.

[31] P. Ekman, W. V. Friesen, and P. Ellsworth. *Emotion in the human face: Guide-lines for research and an integration of findings*. Putnam Publishing, 1972.

[32] P. Gaussier, J.-P. Banquet, F. Sargolini, C. Giovannangeli, E. Save, and B. Pousset. A model of grid cells involving extra hippocampal path integration, and the hippocampal loop. *Journal of Integrative Neuroscience*, 2007.

[33] P. Gaussier, C. Joulain, J.P. Banquet, S. Lepretre, and A. Revel. The visual homing problem: an example of robotics/biology cross fertilization. *Robotics and autonomous system*, 30:155–180, 2000.

[34] P. Gaussier, S. Leprêtre, M. Quoy, A. Revel, C. Joulain, and J.P. Banquet. Experiments and models about cognitive map learning for motivated navigation. In J. Demiris and A. Birk, editors, *Interdisciplinary approaches to robot learning*, volume 24, pages 53–94. Robotics and Intelligent Systems Series, World Scientific, ISBN 981-02-4320-0, 2000.

[35] P. Gaussier, K. Prepin, and J. Nadel. Toward a cognitive systems algebra: Application to facial expression learning and imitation. In *Embodied Artificial Intelligence, F. Iida, R. Pfeifer, L. Steels and Y. Kuniyoshi (Eds.) published by LNCS/LNAI series of Springer*, pages 243–258, 2004.

[36] P. Gaussier and S. Zrehen. Perac: A neural architecture to control artificial animals. *Robotics and Autonomous System*, 16(2-4):291–320, December 1995.

[37] G. Gergely and J.S. Watson. Early socio-emotional development: Contingency perception and the social-biofeedback model. In P. Rochat, editor, *Early Social Cognition: Understanding Others in the First Months of Life*, pages 101–136. Mahwah, NJ: Lawrence Erlbaum, 1999.

[38] C. Giovannangeli and P. Gaussier. Interactive teaching for vision-based mobile robots: A sensory-motor approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 40(1):13–28, 2010.

[39] C. Giovannangeli, P. Gaussier, and J.P. Banquet. Robustness of visual place cells in dynamic indoor and outdoor environment. *International Journal of Advanced Robotic Systems*, 3(2):115–124, June 2006.

[40] C. Giovannangeli and Ph. Gaussier. Autonomous vision-based navigation: Goal-oriented action planning by transient states prediction, cognitive map building, and sensory-motor learning. In *Proc. of the 2008 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS 2008)*, 2008.

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

[42] J.M. Harlow. Passage of an iron rod through the head. *Boston Medical and Surgical Journal*, 39:389–393, 1848.

[43] M.E. Hasselmo, E. Schnell, J. Berke, and E. Barkai. A model of the hippocampus combining self-organization and associative memory function. In G. Tesauro, D. Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems*, volume 7, pages 77–84, Washington DC, July 1995. The MIT Press.

[44] C. Hasson and P. Gaussier. Path integration working memory for multi task dead reckoning and visual navigation. In *(submitted) From animals to animat 11 (SAB)*, 2010.

[45] S. Ikemoto and J. Panksepp. The role of nucleus accumbens dopamine in motivated behavior: a unifying interpretation with special reference to reward-seeking. *Brain Res Brain Res Rev*, 31(1):6–41, Dec 1999.

[46] C.E. Izard. *The Face of Emotion*. Appleton Century Crofts, 1971.

[47] W. James. What is an emotion ? *Mind*, 9(34):188–205, 1884.

[48] J.O'Keefe and N. Nadel. *The hippocampus as a cognitive map*. Clarendon Press, Oxford, 1978.

[49] K. Jungho, B. Yunsu, and K. In So. Robust vision-based autonomous navigation against environment changes. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008.

[50] Takayuki Kanda, Takayuki Hirano, and Daniel Eaton. Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, 19:61–84, 2004.

[51] M. Klinnert, J. Campos, J. Sorce, R. Emde, and M. Svejda. The development of the social referencing in infancy. *Emotion in early development*, (2), 1983.

[52] A.H. Klopf. *The Hedonistic Neuron : A Theory of Memory, Learning and Intelligence*. Hemisphere, Washington D.C., 1982.

[53] C. Lange. *The mechanisms of the emotions*. 1885.

[54] J.E. LeDoux. *Learning and computational neuroscience: foundations of adaptive networks*, chapter Information Flow from Sensation to Emotion: Plasticity in the Neural Computation of Stimulus Value, pages 3–52. M. Gabriel and J. Moore, 1990.

[55] J.E. Ledoux. Brain mechanisms of emotion and emotional learning. *Current Opinion in Neurobiology*, 2(2):191–197, april 1992.

[56] M.D. Lewis. Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Science*, 2005.

[57] P. Maes. Bottom-up mechanism for behavior selection in an artificial creature. In *Simulation of Adaptive Behavior*, 1991.

[58] P. Maes. *Designing Autonomous Agents*, chapter Situated agent can have goals. The MIT Press, 1991.

[59] H.R. Mataruna and F.J. Varela. *Autopoiesis and Cognition: the realization of the living*. Reidel, Dordrecht, 1980.

[60] J.A. Meyer and S.W. Wilson. From animals to animats. In MIT Press, editor, *SAB91: From animals to animats*. Bardford Books, 1991.

[61] M. Mirolli, F. Mannella, and G. Baldassarre. The roles of the amygdala in the affective regulation of body, brain, and behaviour. *Connect. Sci.*, 22(3):215–245, 2010.

[62] M.L. Mittelstadt and H. Mittelstadt. Homing by path integration in a mammal. *Naturwissenschaften*, 1980.

[63] Murray and Trevarthen. *Emotional regulation of interaction between two-month-olds and their mothers*, chapter Emotional regulation of interaction between two-month-olds and their mothers, pages 101–125. Norwood, NJ: Ablex, 1985.

[64] D. G. Myers. *Theories of Emotion*. NY: Worth Publishers, psychology: seventh edition, new york edition, 2004.

[65] J. Nadel, S. Croué, J.M. Mattlinger, P. Canet, C. Hudelot, C. C. Lécuyer, and M. Martini. Do children with autism have expectancies about the social behaviour of unfamiliar people? *Autism*, 4(2):133–145, 2000.

[66] J. Nadel, C. Guerini, A. Peze, and C. Rivet. The evolving nature of imitation as a format for communication. In G Nadel, J. Butterworth, editor, *Imitation in Infancy*, pages 209–234. Cambridge: Cambridge University Press, 1999.

[67] J. Nadel, M. Simon, P. Canet, R. Soussignan, P. Blancard, L. Canamero, and P. Gaussier. Human responses to an expressive robot. In *Epirob 06*, 2006.

[68] M. Ogino, A. Watanabe, and M. Asada. Mapping from facial expression to internal state based on intuitive parenting. In *Epirob 06*, 2006.

[69] P.-Y. Oudeyer and F. Kaplan. Intelligent adaptive curiosity: a source of self-development. In *Proceedings of the 4th International Workshop on Epigenetic Robotics*, 2004.

[70] J. Panksepp. A critical role for "affective neuroscience" in resolving what is basic about basic emotions. *Psychol Rev*, 99(3):554–560, Jul 1992.

[71] J. Panksepp. Neuroscience. feeling the pain of social loss. *Science*, 302(5643):237–239, Oct 2003.

[72] J. Panksepp. Psychology. beyond a joke: from animal laughter to human joy? *Science*, 308(5718):62–63, Apr 2005.

[73] J.W. Papez. A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 1937.

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

[75] R. Plutchik. A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience*, 1980.

[76] C. L. Russell, K. A. Bard, and L. B. Adamson. Social referencing by young chimpanzees (pan troglodytes). *Journal of Comparative Psychology*, 111:185–191, 1997.

[77] J. A. Russell and L. F. Barrett. Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, (76):805–819, 1999.

[78] S. Schaal, A.J. Ijspeert, and A. Billard. Computational approaches to motor learning by imitation. *Philosophical Transactions: Biological Sciences (The Royal Society)*, 358:1431:537–547, 2003.

[79] S. Schachter. *Handbook of psychobiology*, chapter Cognition

and peripheralist-centralist controversies in motivation and emotion, pages 529–564. New York : Academic Press, 1975.

[80] K. Scherer. Emotion as a multicomponent process. *Rev. Person. Soc. Psychol.*, 5:37–63, 1984.

[81] J.H. Schmidhuber. *Advances in Neural Information Processing Systems*, chapter Reinforcement Learning in Markovian and non-Markovian environments, pages 500–506. Morgan Kaufmann, San Mateo, CA, 3 edition, 1991.

[82] R. Sutton and A. Barto. *Reinforcement Learning*. MIT Press, 1998.

[83] S.S. Tomkins. *Affect Imagery Consciousness*. Springer Publishing Company, 1963.

[84] Edward Z. Tronick and Jeffery F. Cohn. Infant-mother face-to-face interaction: Age and gender differences in coordination and the occurrence of miscoordination. *Child Development*, 60(1):85–92, 1989.

[85] F. Varela, E. Thompson, and E. Rosch. *The Embodied Mind*. MIT Press, 1993.

[86] T. Walden and T. Ogan. The development of social referencing. *Child development*, 59(5), 1988.

[87] D. F Watt. Panksepp's common sense view of affective neuroscience is not the commonsense view in large areas of neuroscience. *Conscious Cogn*, 14(1):81–88, Mar 2005.

[88] B. Widrow and M. E. Hoff. Adaptive switching circuits. In *IRE WESCON*, pages 96–104, New York, 1960. Convention Record.

[89] S.I. Wiener, A. Berthoz, and M.B. Zugaro. Multisensory processing for the elaboration of place and head direction responses in the limbic system. *Cognitive Brain Research*, 2002.

[90] T. Wu, N. J. Butko, P. Ruvulo, M. S. Bartlett, and J. R. Movellan. Learning to make facial expressions. In *IEEE 8th Int. Conf. on development and Learning*, 2009.