USING EMOTIONAL INTERACTIONS FOR VISUAL NAVIGATION TASK LEARNING

C. Hasson[1 *] **, S. Boucenna**[1] **, P. Gaussier**[1,2] **and L. Hafemeister**[1]

[1]*ETIS, CNRS UMR 8051, ENSEA, Univ Cergy-Pontoise, France*
[2]*IUF, France*

ABSTRACT

The aim of this study is to show how robots learning could be easier and accessible to non experts if it relies on emotional interactions, more precisely on social referencing abilities, rather than on specialized supervised learning technics. To test this idea, we coupled two systems : a robotic head able to learn to recognize and imitate emotional facial expressions and a mobile robot able to learn autonomous visual navigation tasks in a real environment. Two possible solutions for coupling these two systems are tested. First, the emotional interactions are used to qualify the robot's behavior. The robot shows its ability to learn how to reach a goal-place of its environment using emotional interaction signal from the experimentator. These signals are giving the robot information about the quality of its behavior and allow it to learn place-actions associations to construct an attraction basin around the goal-place. Second, the emotional interactions are used to qualify the robot's immediat environment. The robot shows its ability to learn how to avoid a place of its environment by associating it with the experimentator's anger facial expression. The first strategy allows the experimentator to teach the robot to reach a specific place from anywhere in its environment. However, this strategy takes more learning time than the second strategy that is very fast but seems to be inappropriate to learn to reach a place instead of avoiding it. While these two different strategies achieve satisfactory results, there is no reason why they should be mutually exclusive. In conclusion, we discuss the coupling of both type of learning. Our results also show that relying on the natural expertise of humans in recognizing and expressing emotions is a very promising approach to human-robot interactions. Furthermore, our approach can provide new interesting insights about how, in their early age, humans can develop high level social referencing capabilities from low level sensorimotors dynamics.

Keywords: Emotional interactions, interactive learning, autonomous robotics.

* **Corresponding author:** hasson@ensea.fr

1. INTRODUCTION

Robots able to learn navigation tasks are usually taught under supervision of an experimentator [1]. These techniques have the advantage of being fast in terms of learning time but the experimentator has to know exactly how the robot works and to be expert in its use. In others words, the experimentator 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. 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 [2, 3, 4]. Besides, being non verbal and thus not needing high level cognitive abilities, gathering information through emotional interactions seems to be a fast and effcient 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 sensorimotor associations controlling the robot's learning are defined throughout their interactions with the experimentator. This interactive learning does not rely on the experimentator technical expertise, but on his/her ability to react emotionally to the robot's behavior in its environment. In that case, the human and the robot have to adapt reciprocally to each other through means that are much more natural to humans. To test this idea, we developped a robotic head able to learn online to recognize emotional facial expressions [5]. The robot internal emotional state triggers one specific expression (e.g. happiness) and the human mimicks the robot in front of it. The robot then learns to associate its internal emotional state with the humans facial expression. After a certain amount of learning time, the robot is able to recognize the human facial expression as well as to mimick its facial expressions. Moreover, the robot is able to learn to navigate autonomously in its environment using visual and odometric information and is thus able to reach specific places of interest (figure 1). We study the merging of these two processes as a way to provide the robot the abil-

Figure 1: 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).

ity to learn navigation tasks via emotional interactions. Facial expression recognition affects the robot's behavior by providing information about its environment or its behavior. We studied these two different ways to connect the navigation and the emotional interaction systems. We think this approach can be usefull for the design of interacting robots and more generaly, 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 sensorimotors dynamics.

2. EMOTIONAL INTERACTIONS AND NAVIGATION SYSTEM

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

2.1. Sensorimotor based facial expression recognition system

This work was motivated by the question of how a robotic system (figure 2, right), able to exhibit a set of emotional expressions, can learn autonomously to associate theses expressions with those of others. Here, ”autonomously” refers 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. This question is close to the understanding of how babies learn to recognize the facial expressions of their caregivers without having any explicit teaching signal allowing them to associate, for instance, an ”happy face” with their own internal emotional state of happiness. Using the cognitive system algebra [6], we showed that a simple sensorimotor architecture (figure 2, left) 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 [7]. Moreover, psychological experiments [8, 9] have shown that humans reproduce involuntary a facial expression when observing 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

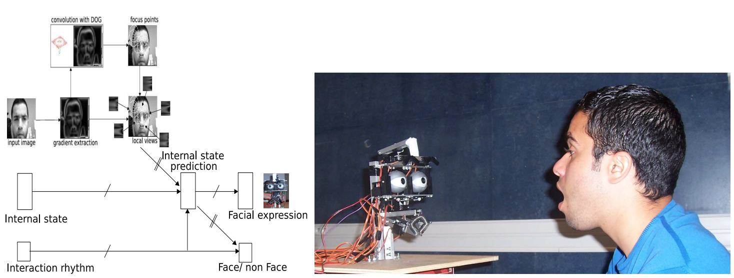

Figure 2: Architecture used to associate a collection of local views around feature points extracted from the visual flow with the expressed emotion by the robot. If a human comes in front of the robot and imitates the robot’s expressions, (s)he will close the loop between vision and proprioception and allows the system to learn to recognize the facial expression of the human partner.

also showed that, learning autonomously to recognize a face can be really more complex than to recognize a facial expression. We proposed an architecture (figure 2, left) using the rhythm of the interaction to allow first a robust learning of the facial expression without a face tracking [10], and second, to stop the learning when the visual stimuli (facial expression or absence of face) are not synchronized with the robot facial expression.

2.2. Sensorimotor based navigation system

The robot's navigation abilities are based on a bio-inspired sensorimotor associations learning system : the PerAc architecture [11]. This architecture allows the robot to learn the conditioning of an action by a sensory input. More precisely, the robot's visual system is inspired from visual navigation models issued from neurobiology [12]. It consists in 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 [13, 11]. 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. The area where a given place cell is the most active is called its place field. A conditioning neural network enables the learning of the association between a place field and an action (e.g., a direction to head for). Later recognition of this place field will trigger the linked action.

3. BEHAVIORAL AND ENVIRONMENTAL COUPLING BETWEEN THE NAVIGATION AND THE FACIAL EXPRESSION RECOGNITION SYSTEMS

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 experimentator 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 usefull place that the robot should eventually seek. We studied these two different possible couplings between the navigation and the emotional interaction part of our architecture.

3.1. Coupling behaviors and emotions

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, we have adapted the PerAc architecture [11] to be able to learn positive actions such as negative action conditionings. To ensure this classical conditioning, we used the least mean square learning rule [14] that uses the difference between the neural network output and the desired output to compute the amount by which the connexions weights have to be changed (weight adaptation due to learning) :

$$\Delta w_{ij} = \epsilon . I_i . (O_j^d - O_j) \tag{1}$$

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 sensorimotor association neural network that can only learn positive conditionings, we used one association neural network for positive and one for negative conditionings. A third group of neurons is used to compute the sum of these two outputs (see figure 3). While the positive conditioning group of neurons has a positive connection with this summative group of neurons (activations), the negative conditioning group of neurons has a negative connection (inhibition). This solution allows to store much more information about what is learned by the robot than ouputs 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 very nature of the environment or in the experimentator'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 3. 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 sensorimotor conditioning is learned according to the nearest place, completing an eventually previously learned sensorimotor conditioning. The robot is thus able to learn progressively which direction to avoid and which direction to head for in the different "places" of its environment and according to the goal of the person interacting with it. We tested this architecture in the following situation : the robot's environment contains one

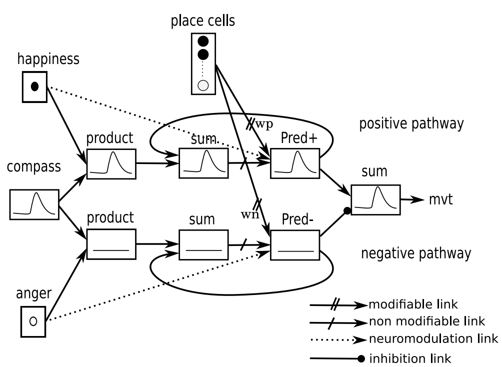

Figure 3: Behavioral coupling model. When one of the "conditioning" groups of neurons using equation 1 receives neuromodulation from the recognition of the corresponding facial expression (happiness in this example), it learns the association between the current robot location (perceived as a specific 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.

place of interest and the experimentator want to teach the robot how to reach it. Each time the experimentator thinks the robot's behavior is wrong, he expresses anger toward the robotic head and, conversely, he smiles for good behaviors (happiness). Figure 4 is an illustration of the learning chronology (as explained above). Figure 5 shows the robot's trajectories after learning. The robot

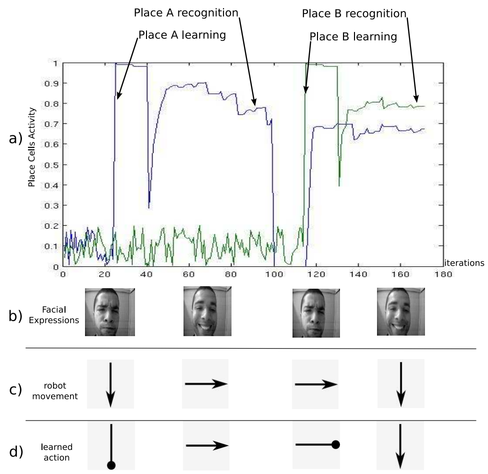

Figure 4: a) place cell signal. b) The experimentator's facial expressions recognized by the robot. c) current robot direction of movement. d) action learned by the robot (an arrow means a direction to activate and a dot a direction to inhibit). The experimentator facial expressions give the robot the information needed about its behavior to learn the necessary sensorimotor associations between the visu signal (recognition of the current place) and the learning of the activation or inhibition the current movement direction.

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 experimentator 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.

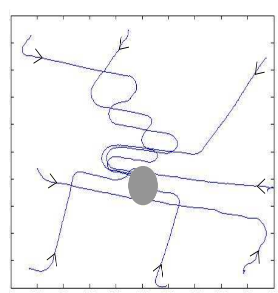

Figure 5: Robot's trajectories from different starting points : the robot is able to reach the place associated with the happiness facial expression. The grey zone represents the goal place. These trajectories are obtained by video tracking. The size of the experimental area is 3m x 3m.

3.2. Coupling environment and emotions

The environmental coupling refers to the situations where the recognition of an emotional facial expression is used to qualify the robot's immediat environment. When the experimentator displays a happy face, this means that the robot has to learn the place is "good" (e.g. containing a resource needed by the robot) and conversely for an angry face that the place is "bad". The robot has to seek good places and to avoid the dangerous ones. To do this, we used one more time a modified version of the PerAc architecture. When the robot receives the social interaction signal, it has to learn a new place cell characterising its location and to learn to predict the interaction signal (happyness or anger) which is treated as a reward associated with this place (figure 6). As the

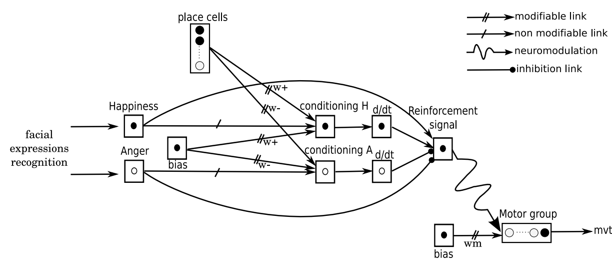

Figure 6: Environmental coupling model. Using the least mean square learning rule, the conditionning neurons allow the association between a place cell (a zone of the environment) and an experimentator's facial expression (modifications of weights w^+ and w^- follow equation 1). The predicted expression 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_m 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

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 is used as a reinforcement signal for neurons using the Sutton and Barto learning rule [15] (figure 6) :

$$\Delta R = (\frac{dPredH}{dt} - \frac{dPredA}{dt}) + (H - A) \quad (2)$$

$$\Delta w^{+/-} = \epsilon . \frac{dR}{dt} . \frac{dO_j}{dt} . I_j \quad (3)$$

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 ouput 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 ouptut. We used the ouputs to control the robot actions. To test this architecture, we assigned various fixed directions to the robot after it has learned through interactions with the experimentator that the place at the center of its environment is dangerous (i.e. associated with the anger expression). Figure 7 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 8 shows the robot's trajectories from different starting points

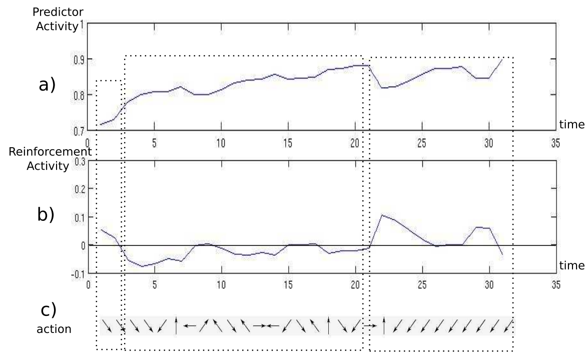

Figure 7: a) The reward prediction (positive with happiness and negative with anger) informs the robot about its behavior outcome in the environment. b) Derivatives of this value are used as a reinforcement signal(see equation 3). c) when the derivative is negative, the robot direction changes and when it is positive it is maintained and reinforced

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 experimentator allows the robot to quickly learn to avoid it (the first interaction is 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 6. 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 meaningfull when the robot is too far away from the learning location.

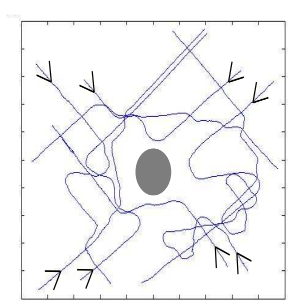

Figure 8: Robot's trajectories from different starting points (with a fixed direction) after interactive learning of the association of the grey zone to the anger facial expression. The robot is able to avoid the place associated with the anger facial expression.The prediction of the negative reinforcement is sufficient to inhibit a movement in direction of the dangerous zone (when it is near it).

4. DISCUSSION

While the behavioral (associating emotions to the robot's actions) and the environmental (associating emotions to the robot's environment) couplings have both shown their possibilities and limitations, one could easily argue that a more complete and convincing solution would be to give the robot the ability to perform both with the same architecture. Nonetheless, this complete coupling is far from trivial because of 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”. One way this problem could be solved is to give the robot the ability to recognize more than the two different facial expression we used in the earlier experiments. For instance, an anger facial expression could mean that the robot's behavior is ”bad” (according to the experimentator and/or to its environment) while a fear facial expression could mean that the robot is in a dangerous part of its environment. Nevertheless, this solution cannot be scaled to more complex problems because it does not provide a real coupling of these two possible types of learnings. This solution relies more on the experimentator expertise about the robot's architecture than on the natural ability to interact emotionaly. Another solution to this problem could be the way the system treats the interaction inputs. While the behavioral coupling uses a phasic signal (the moment the signal appears), the environmental coupling uses a tonic signal (the whole time the signal is present). This way, both couplings 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 experimentator 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, only one of the two learnings was intended by the experimentator. 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. In conclusion, we described a system where a human interacting with a robotic head is able to help a mobile robot to learn different navigation tasks. Yet, one can argue that in these experiments, there is no real interaction between the experimentator and the robot. The facial expressions are used more as commands than interaction signals. In a sense, the robot's behavior acts like a communication signal on the experimentator which reacts to it in order to improve the robot's learnings. But in a ”real” interaction, the robot should be able to express its internal states in order to give the experimentator more information about how it is dealing with its environment. Future works will focus on the need of a more realistic interaction where a bidirectionnal 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 mirror the human facial expression to allow the experimentator to see that his/her mood has been well understood by the robot. Allowing a real interaction means to find a solution for expressing something related both to the expressive feedback of the experimentator and the robot's internal state. Control of the expression intensity and its duration is a lead we will explore. Moreover, for the moment, the robotic head and the mobile robot are two distinct devices. Having a more sophisticated and realisitic set-up could have a major impact on the way the human and the robot interact.

5. ACKNOWLEGDMENTS

This study is part of the European project ”FEELIX Growing” IST-045169, the French Region Ile de France ”DIGITEO”and the Institut Universitaire de France.

REFERENCES

[1] Hersch, M, adn S. Calinon, F. G, and Billard, A. Dynamical system modulation for robot learning via kinesthetic demonstrations. *IEEE Transactions on Robotics*, 24:1463–1467, 2008.

[2] Klinnert, M, Campos, J, Sorce, J, Emde, R, and Svejda, M. The development of the social referencing in infancy. *Emotion in early development*, 2:57–86, 1983.

[3] Walden, T and Ogan, T. The development of social referencing. *Child development*, 59:1230–1240, 1988.

[4] Russell, C, Bard, K, and Adamson, L. Social referencing by young chimpanzees (pan troglodytes). *journal of comparative psychology*, 111(2):185–193, 1997.

[5] Gaussier, P, Boucenna, S, and Nadel, J. Emotional interactions as a way to structure learning. *epirob*, pages 193–194, 2007.

[6] Gaussier, P. Toward a cognitive system algebra: A perception/action perspective. *In European Workshop on Learning Robots (EWRL).*, pages 88–100, 2001.

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

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

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

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

[11] Gaussier, P and Zrehen, S. Perac: A neural architecture to control artificial animals. *Robotics and Autonomous System*, 1995.

[12] J. O'Keefe, N. N. The hippocampus as a cognitive map. *Clarenton Press, Oxford*, 1978.

[13] Gaussier, P, Lepretre, S, Quoy, M, Revel, A, Joulain, C, and Banquet, J. *Interdisciplinary Approaches to Robot Learning*, chapter Experimentsandmodels about cognitive map learning for motivated navigation. Robotics and Intelligent Systems Series World Scientific, 2000.

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

[15] Barto, A, sutton, R, and Brouwer, P. Associative search network: A reinforcement learning associative. *Biological Cybernetics*, 40(3):201–211, 1981.