

Frustration as a generical regulatory mechanism for motivated navigation

Cyril Hasson¹ and Philippe Gaussier^{1,2}

¹ETIS Cergy Pontoise University, CNRS UMR8051 - ENSEA - F-95000 Cergy, ²IUF
{hasson,gaussier}@ensea.fr

Abstract—This paper explores the use of a mechanism to auto-regulate the robot behavior in situations of persistent failures. In order to give more autonomy to a mobile robot, a generical frustration mechanism based on the automonitoring of the progress (in terms of goal distance reduction) is studied in different situations and on different parts of the robot architecture. To escape failure situations and deadlocks, the frustration reaction can inhibit the robot navigation strategies, goals or drives.

I. INTRODUCTION

If we follow the animat approach [5], to ensure its survival, a robot must maintain a set of artificial physiological variables inside safe levels. Thus, the robot must look for different resources to fulfill its various needs i.e. the robot must have goals that depends on its motivations [12], [13]. However, sustaining a durably efficient behavior in a dynamic and complex environment remains a difficult task [15], [10], [1]. The very nature of the environment as well as the robot's own imperfections will most likely let arise situations where the learned behaviors are not sufficient and may lead to deadlocks [3]. 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.

In this paper, we describe a generical frustration mechanism based on the robot goals monitoring. Frustration is used to regulate the robot behavior in case of persistent failures. The robot navigates using biologically inspired mechanisms like path integration based on odometric information [6], [14] (return vector computing) and sensorimotor learning based on visual place recognition [8], [7], [18]. These navigation strategies are coupled to a low level motivational system (using the simulated physiology as input) in order to perform a survival task. Metacontrol of the robot success is an efficient way to design an emotional regulation [2] and enhance the robot autonomy. Figure 1 shows the robot and its environment.

Section 2 describes the simulated physiology and the motivational system. The proprioceptive and visual navigation architecture we use are briefly presented in section 3. Section 4 describes the robot goals multimodal automonitoring and the frustration mechanism. Section 5 shows experimental results with the robot. And section 6 contains the conclusions.



Fig. 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.

II. PHYSIOLOGY AND MOTIVATIONAL SYSTEM

A synthetic physiology simulates physiological variables like hydration or glucose levels. These variables levels constantly decrease as the robot consumes its internal resources. Collecting a simulated resource (i.e. detecting a needed resource) results in an increase of the corresponding resource level. Robot's survival is only possible if it collects periodically the resources it needs so these variables levels don't go 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 2 describes this system.

III. NAVIGATION STRATEGIES

Two different navigation strategies were used in order to study inhibitory effects of the frustration regulation on the action selection process. Furthermore, this provides different test situations for the frustration regulation.

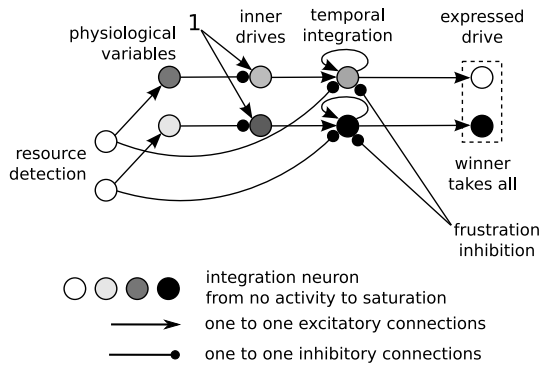


Fig. 2. Low level drive system : inner drives are computed from the physiological variables levels, integrated drives are signals that can be manipulated without affecting the inner states of the system and the expressed drive is the most active integrated drive. Frustration inhibition is described in section IV.

Proprioceptive navigation : path integration is the ability to determine the return vector (angle and distance) to an arbitrary reference point using odometric information. We designed a motor working memory [9] to use path integration implementation presented in [6] (figures 3 and 4) for multiple goal tasks.

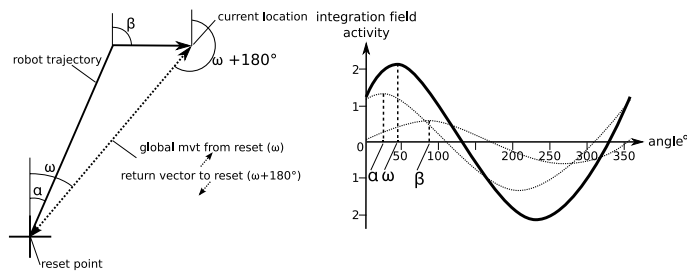


Fig. 3. Illustration of the path integration computation. Left figure shows a simulated trajectory composed of two segments of different lengths (the first is three times the length of the second) and orientations (25° and 90° from an arbitrary absolute direction). Right figure shows the two inputs (dotted curves) as bell shape centered on the movements absolute directions (α and β) and their sum (the bold curve). 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).

Detection of a new goal allows the recruitment of a dedicated integration field. Every integration field computes dynamically the return vector to its associated goal (figure 5).

Visual navigation : the visual system (figure 8) is able to learn to characterize (and thus recognize) different places of the environment. The visual system, a simulated neural network, 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 [8], [7] (see figure 6).

Activities of the different place cells depend on the recognition levels of these visual cues and of their locations. As shown in figure 7, a place cell will then be more and more

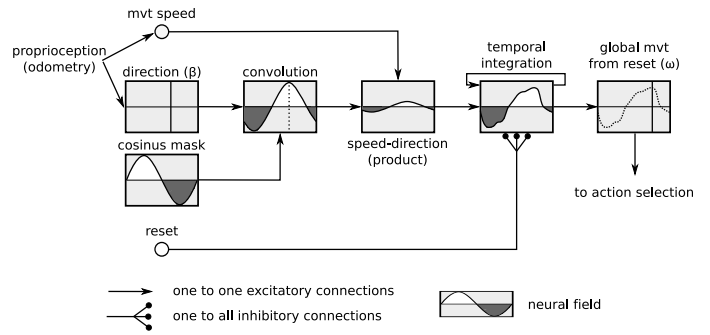


Fig. 4. Path integration : speed is coded as the activity of one neuron and orientation as the most active neuron of a field (i.e. a simple linear collection of neurons). At every time step, the integrator takes as input the activity of the orientation field (convoluted by a bell shape curve e.g. a gaussian 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.

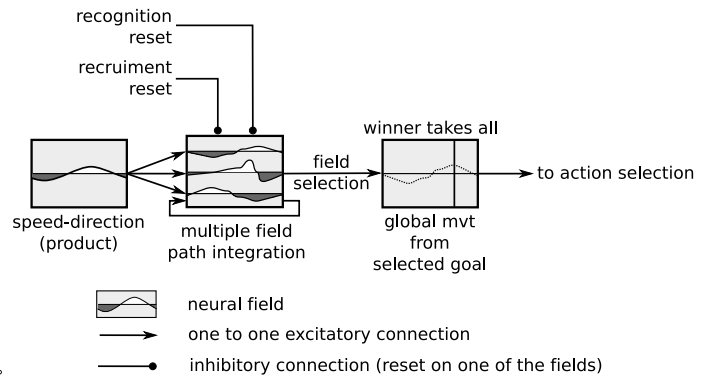


Fig. 5. multi goal path integration navigation : return vectors to several places (goals) are computed dynamically. This model is fully described in [9]. Recruitment reset is the recruitment of a new integration field when a new goal is found. Recognition reset is the reset of the field corresponding to a detected known goal. Field selection is the selection of the integration field corresponding to the closest goal satisfying the active drive.

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 field. An associative learning group of neurons allows sensorimotor learning (place-drive-action group on figure 8). 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.

IV. FRUSTRATION FROM GOAL DISTANCE MONITORING

In a dynamic environment, conditions will inevitably change in a very perturbing way for the robot. For instance, a resource might be unreachable, all lights in the environment might be turned off (no more visual navigation), the robot might be "kidnapped" and placed in another place (erroneous proprioceptive navigation), ... All these perturbations might get the robot trapped in a deadlock situation. They could each be detected by specific means at sensor level, but a progress (or failure) evaluation might offer a generical approach to this issue. The aim is not to combine navigation strategies in a way that avoids deadlocks but to construct

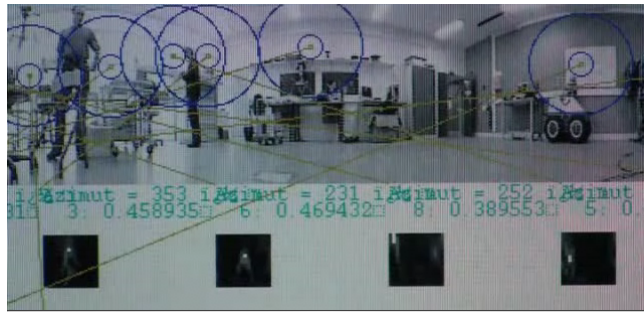


Fig. 6. Landmarks and their azimuths extracted from the raw visual flow and learned by the visual system.

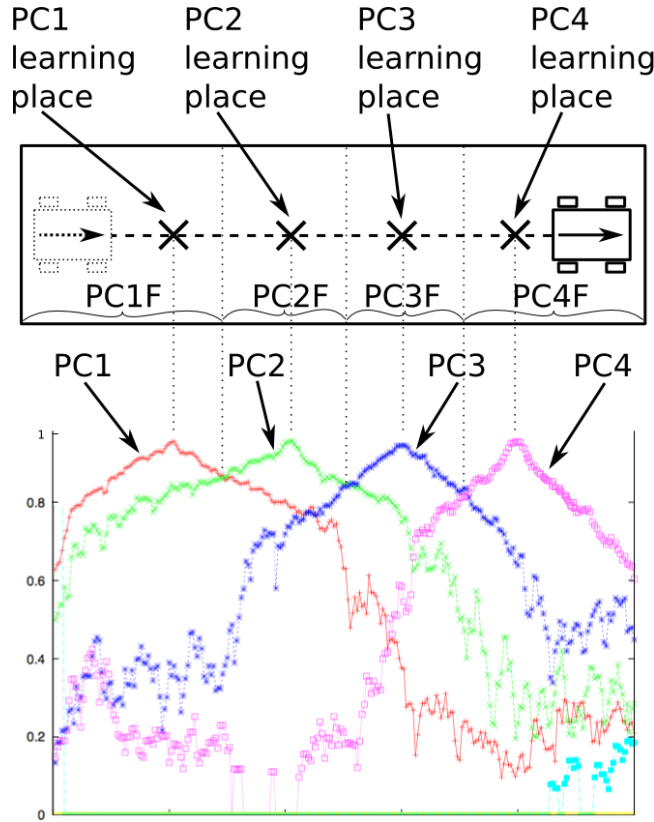


Fig. 7. As the robot gets closer to each place cell learning location, the corresponding place cell (PCn) gets more active, the maximum activity of a place cell corresponds to its learning location. And the area where a place cell activity is the highest is its place field (PCnF).

a simple mechanism to detect them. Monitoring the goal distance over time is an efficient and simple way for the robot to evaluate its progress. 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 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. From

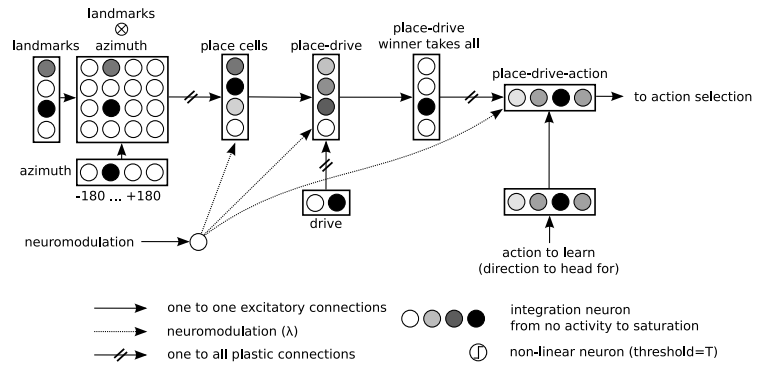


Fig. 8. Sensorimotor visual navigation : a visual place cell is constructed from recognition of a specific landmarks-azimuths pattern (tensorial product) and an action (the return vector) is associated with this place cell.

vision, the robot can monitor the activity of the place cell associated with the goal. The robot has to learn which place cell corresponds to the goal and then monitor its activity. As the robot gets closer to the goal, the corresponding place cell activity gets higher. When the robot has an active drive e.g. when it is hungry, until food is found, as long as the predicted distance between the food and the robot decreases, the robot might assume that everything is alright. But if goal distance does not decrease, the robot behavior is inefficient. And if this inefficiency is lasting this means the robot is caught in a deadlock. The threshold T (figure 9) defines the robot tolerance to frustration. This mechanism differs from a simple timeout because failure and not time is what increase frustration. According to this view, solving a long problem should not be frustrating as long as progress can be perceived. Furthermore, 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. The following equation describe the frustration mechanism :

$$F(t) =$$

$$\begin{cases} 1 & \text{if } [f(t - \Delta t) + [\frac{G(t) - G(t - \Delta t)}{\Delta t} + \epsilon]^+ - R]^+ > T \\ 0 & \text{otherwise} \end{cases}$$

with $[x]^+$ equals x if $x > 0$ and equals 0 otherwise.

$F(t)$ is the frustration value, $f(t)$ is the temporal integration of failure to get closer to the goal, $G(t)$ is the goal distance, Δt is the duration of each calculation time-step, 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 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 for 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 2 but the same kind of inhibition allows to switch from the active strategy or goal. Figure 9 shows the neural network used to detect failure situations and the way it can regulate the robot behavior.

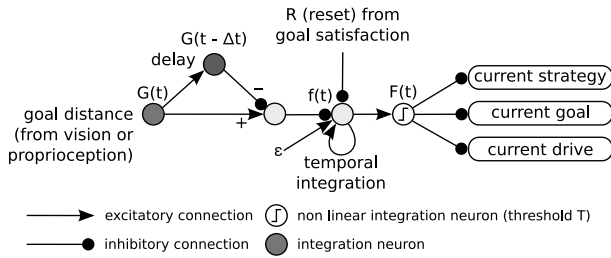


Fig. 9. Frustration mechanism : non linear integration of the goal distance derivative over time. When the goal distance comes from vision, its temporal difference is computed the opposite way as when goal distance comes from proprioception (goal place cell activity increases while integration field 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.

V. ROBOTIC EXPERIMENTS

In the first experiment¹, the effect of the frustration regulation is tested on 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 will get 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 10 shows the robot trajectories as well as its internal drive, failure detection and frustration signals.

In the second experiment, frustration regulation is tested on the goal 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 [9] (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 11 shows the robot trajectories as well as its internal goals, failure detection and frustration signals.

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

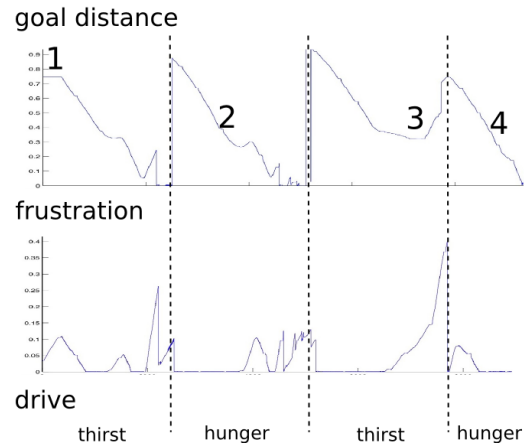
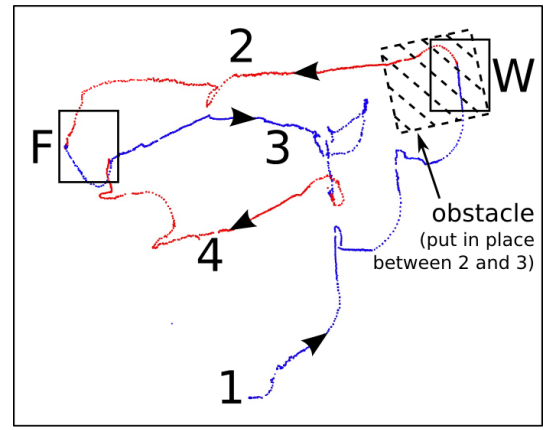


Fig. 10. Up : robot trajectories with frustration of the active drive (visual navigation). F stands for food and W for water. Down : goal distance, failure detection and drive signals. In 1, the robot starts the experiment with thirst as the most active drive. In 2, the robot satisfy its thirst and hunger becomes the active drive. In 3, the water resource is obstructed by an obstacle. When enough failure detection has been integrated, a frustration inhibition is sent to the active drive (thirst) and in 4, the robot go back the food location.

resource. After having learned to reach the two resources with each strategy, the robot uses the proprioceptive 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 converge toward the wrong locations. Inhibition of the active strategy allows the robot to switch from its proprioceptive to its visual navigation strategy which is robust to that kind of perturbation. Similarly, proprioceptive navigation is a good way to navigate in the dark thus offering a good alternative to visual navigation. Figure 12 shows the robot trajectories as well as its internal strategies, failure detection and frustration signals.

VI. CONCLUSION AND PERSPECTIVE

Monitoring progress is an efficient way to react to changing conditions of a dynamical environment. The robot can construct 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

¹Trajectories in all experiments are recorded from an onboard tracking device that is not used by the robot.

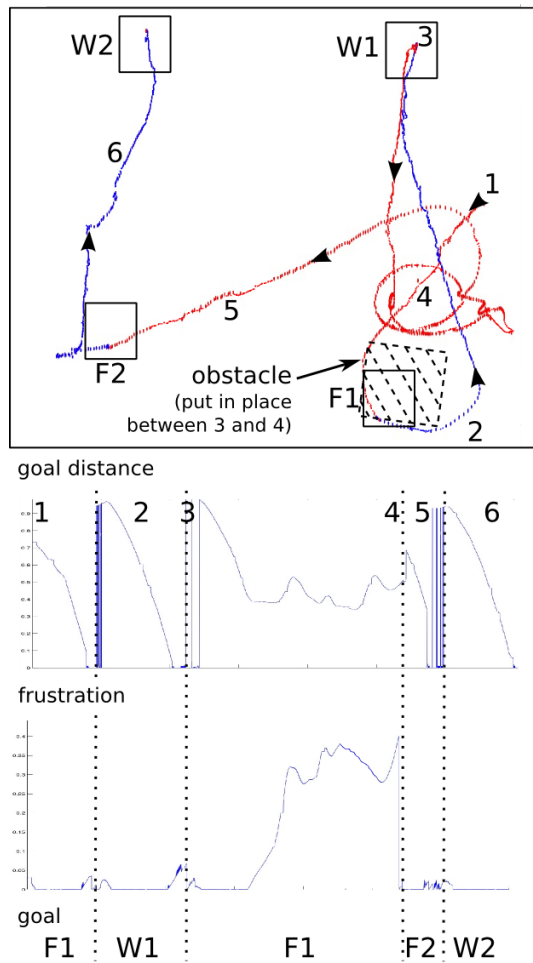


Fig. 11. Up : robot trajectories with frustration of the active goal (proprioceptive navigation). F1 and F2 are the two food resource and W1 and W2 are the two water resource. Down : goal distance, frustration and goals signals. In 1, the robot starts the experiment with hunger as the active drive and 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.

satisfaction) give 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 generical inhibition mechanism and the behavioral change it causes can be viewed as an emotional regulation : frustration. Using a metacontrol regulatory mechanism, the robot adapts its behavior to changing conditions rather than getting stuck in a deadlock situation where its learnings are not sufficient. In this paper, we described an empirical view of the frustration regulation which needs to be refined in later works. This emotional regulation was used for rapid behavioral changes but it could also be used for long term evaluation. 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

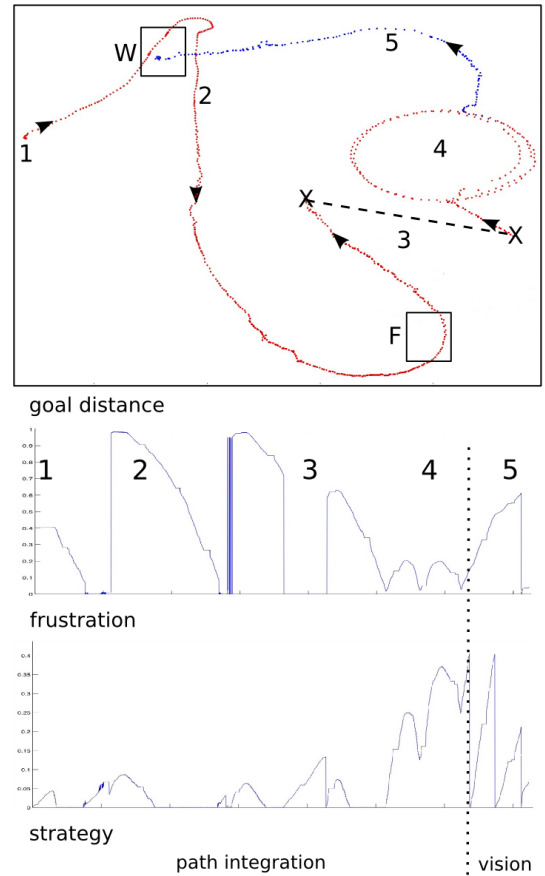


Fig. 12. Up : robot trajectories with frustration of the active strategy. Down : goal distance, frustration and strategy signals. In 1, the robot starts the experiment with path integration. The active drive is thirst. In 2, after having satisfied its thirst, hunger becomes the active drive and the robot 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.

or drive and can then be used to select them accordingly. Our model thus bears strong similarities with TD lambda [17], [4] and the possibilities of hedonist neurons [11]. This frustration regulation can also be compared to the novelty detection and curiosity mechanisms described by [16]. 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.

Future works will focus on the need to construct a multimodal perception of the goal distance able to use both or only a single modal goal distance according to the perceptual context and the ongoing frustration regulation. Furthermore, 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 this equation. Frustration inhibition has been applied to strategies, goals and drives, but no preference or selection mechanism has been used. Future works will thus also focus on the need to define

specific failure context and link them to effective frustration inhibition.

ACKNOWLEDGMENTS

This work has been supported by the european project Feelix Growing.

REFERENCES

- [1] 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.
- [2] L. Canamero, P. Gaussier, C. Hasson, and A. Hiole. *Ouvrage transdisciplinaire sur les emotions*, chapter Emotion et cognition : les robots comme outils et modeles. Hermes, in press.
- [3] Richard Canham, Alexander H. Jackson, and Andy Tyrrell. Robot error detection using an artificial immune system. *Evolvable Hardware, NASA/DoD Conference on*, 2003.
- [4] P. Dayan. Motivated reinforcement learning. In *Advances in Neural Information Processing Systems (MIT Press)*, 2001.
- [5] J.-Y. Donnat and J.-A Meyer. Learning reactive and planning rules in a motivationally autonomous animat. *Systems Man and Cybernetics*, 1996.
- [6] 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.
- [7] 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.
- [8] C. Giovannangeli and Ph. Gaussier. Interactive teaching for vision-based mobile robot: a sensory-motor approach. *IEEE Transactions on Man, Systems and Cybernetics, Part A: Systems and humans*, 2010.
- [9] 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.
- [10] 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.
- [11] A.H. Klopff. *The Hedonistic Neuron: A Theory of Memory, Learning and Intelligence*. Hemisphere, New York, 1982.
- [12] P. Maes. Bottom-up mechanism for behavior selection in an artificial creature. In *Simulation of Adaptive Behavior*, 1991.
- [13] P. Maes. *Designing Autonomous Agents*, chapter Situated agent can have goals. The MIT Press, 1991.
- [14] M.L. Mittelstadt and H. Mittelstadt. Homing by path integration in a mammal. *Naturwissenschaften*, 1980.
- [15] Viet Nguyen, Ahad Harati, Nicola Tomatis, Agostino Martinelli, and Roland Siegwart. OrthoSLAM: A Step toward Lightweight Indoor Autonomous Navigation. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, Beijing, China*, 2006.
- [16] 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.
- [17] R. Sutton and A. Barto. *Reinforcement Learning*. MIT Press, 1998.
- [18] 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.