

Bio-inspired Model of Robot Adaptive Learning and Mapping

Alejandra Barrera Ramírez and Alfredo Weitzenfeld Ridel

Computer Engineering Department - Robotics and CANNES Laboratories
Instituto Tecnológico Autónomo de México (ITAM)
Mexico City, Mexico

E-mail: abarrera@itam.mx, alfredo@itam.mx

Abstract - In this paper we present a model designed on the basis of the neurophysiology of the rat hippocampus to control the navigation of a real robot. The model allows the robot to learn reward locations dynamically moved in different environments, to build a topological map, and to return home autonomously. We describe robot experimentation results from our tests in a T-maze, an 8-arm radial maze and an extended maze.

Index Terms - *Affordances, mapping, path integration, rat hippocampus, reinforcement learning.*

I. INTRODUCTION

Robot localization and navigation requires an internal representation of the environment. Simultaneous localization and mapping (SLAM) is the problem of building a map of the environment based on the robot position and simultaneously localizing the robot within the map built.

The approaches that have been proposed to solve the SLAM problem relied on the construction of different kinds of maps: topological, metric, feature based and hybrid maps.

Topological approaches represent the robot environment by using graphs, which are compact and efficient representations for solving tasks as path planning. They are used mainly in indoor environments, where clear distinctive places can be found. A known problem of the topological approach is place recognition in complex environments due to the lack of metric information to discriminate between two places that look alike.

The most common metric representations of the environment are occupancy grids, which maintain stochastic estimates of the occupancy state of each cell and where object locations are defined in a Cartesian coordinate frame. This technique provides very rich representations of the environment used for obstacle avoidance and path planning, but do not provide consistent global maps estimates when working in large environments [1].

Feature based techniques represent the environment with parametric features such as lines, cylinders, corners, etc. The location of the features besides other information such as geometry or color can be used to represent and maintain the map. Feature maps are suitable in environments where it is possible to distinguish geometric features. The map obtained could be a sparse representation of the environment that does not provide information to performed detailed navigation.

In the past few years various studies have presented

implementations of maps combining different approaches; e.g., in [2] local perceptual maps are built as occupancy grids in local regions of the environment, and a topological map is built to describe the structure of the large scale space. In [3] higher level topological maps derived from images are combined with lower level metric maps. Ref. [4] proposes the construction of a graph of coordinate frames, where nodes represent local frames (metric maps) and edges represent the transformation between adjacent frames. Ref. [5] uses features to build a topological map with metric information associated with the edges. Ref. [1] presents a metric map structure that combines feature maps with occupancy grids.

Currently, the main challenges in SLAM research include map consistency through closing loops [1], [2], [4], data association [1], [5], perceptual ambiguity [2], and dynamic environments [6].

On the other hand, the navigation strategies of animals such as rats and primates also relies on internal representations of the environment often referred to as cognitive maps [7], which are located in the brain region called hippocampus [8]. Experimental work has shown the existence of neurons called place cells in the rat hippocampus, which codify information about physical locations of the animal. Studies on the rat brain have provided inspiration in implementing those biological findings to allow adaptive navigation behavior in mobile robots coping with unknown environments. In this way, several robotic navigation models have been developed such as [9], [10], [11], [12], [13], [14], [15].

In this paper we present a navigation model based on the neurophysiology of the rat hippocampus that allows an actual robot to learn reward locations in different mazes, to build a topological map of the environment and to return home autonomously. This model was tested in different learning and mapping experiments. We do not pretend to compare our results with those obtained by classical SLAM approaches because until now we have been concerned in understanding nature to allow the model's capability of adaptive learning in small and controlled environments, but our next goal is to work in proposing an alternative solution to some of the SLAM challenges previously mentioned: the map consistency and the data association. Section II of the paper presents the model description, Section III discusses the robotic architecture and experimentation results, and we conclude in Section IV.

II. MODEL DESCRIPTION

Our model is composed of layers of neurons that implement Hebbian [16] and reinforcement learning [17] in order to allow the expression of goal-oriented behavior. Fig. 1 shows the different layers in the model. We discuss the model components in the following subsections.

A. Sensory Inputs

The sensory inputs to the model are composed of **affordances**, **kinesthetic information** and the **internal drive** of the rat.

The notion of **affordances for movement** represents all possible motor actions that a rat can execute through the immediate sensing of its environment; e.g., visual sighting of a corridor – go straight ahead, sensed branches in a maze – turn. In our model, affordances for movement are coded in a linear array of cells called **perceptual schema** (PS) that represents possible turns from -180° to $+180^\circ$ in 45° intervals. In this way, when the rat is in the center of an 8-arm radial maze, it is able to visually sense eight different arms and consequently perceive eight different affordances. Determination of the affordances for movement is based on a local coordinate system that is relative to the rat's head (see Fig. 2 Left).

The activation level of the neurons in PS is computed through an exponential equation. For a specific affordance (aff), the activation level of neuron i is determined as follows:

$$aff_i = e^{-\frac{(i-a)^2}{2d^2}}, \quad (1)$$

where d is a constant value representing the width of the exponential, and a is a constant that depends on the relative direction of the affordance: from -180° to $+180^\circ$, $a = 4 + 9m$ with m from 0 to 8.

Every available affordance has the form shown in (1); e.g. the information picked up by the PS when the rat is in the center of an 8-arm radial maze is a sum of all available affordances and the value of each of its neurons is computed using (2).

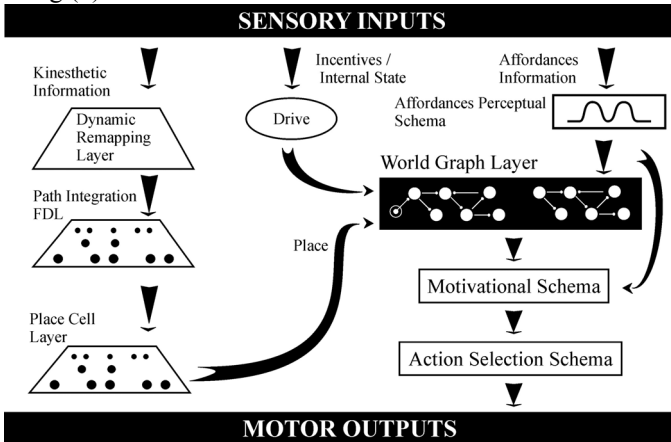


Fig. 1. The layers of the hippocampus-based navigation model. FDL stands for Feature Detector Layer.

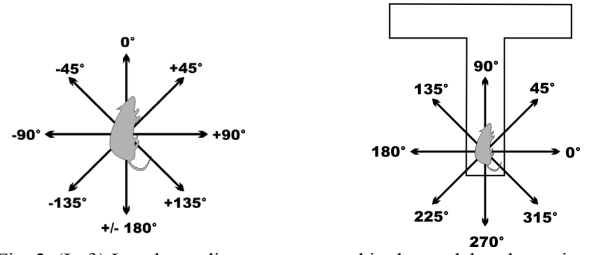


Fig. 2. (Left) Local coordinate system used in the model to determine the relative affordances for movement. (Right) Global coordinate system used in WGL to build the map-based representation of the external environment; e.g., a T-maze. This global system is relative to the departure location (the base of the T in the figure).

$$aff_i = \sum_{m=0}^8 e^{-\frac{(i-(4+9m))^2}{2d^2}} \quad (2)$$

The second kind of sensory input to the model is determined by kinesthetic information; i.e. rat internal body signals generated during rat's locomotion (the magnitude and direction of movement). These signals are used by rats to carry out the **path integration** process, by which kinesthetic information allows them to update the position of their point of departure each time they move in relation to their current position. In this way, path integration allows the animal to return home. As can be seen in Fig. 1, the model includes a path integration module composed of a **dynamic remapping layer** (DRL), and a **path integration feature detector layer** (PIFDL).

DRL generates a dynamic remapping perceptual schema (DRPS) defined as a two-dimensional array of neurons, whose activation level is computed similarly to the activation level of the neurons in PS. Initially, DRPS codifies the position of the departure location (the environmental anchor). Then, DRPS is updated each time the rat moves by displacing the anchor position in the same magnitude but in opposite direction to the rat's movement. Every neuron in DRPS is randomly connected to 50% of the neurons in PIFDL. Connection weights between layers are randomly initialized and normalized between 0 and 1. The activation level A_j of neuron j in PIFDL is computed by adding the products between each input value I_i coming from neuron i in DRPS and the corresponding connection weight w_{ij} as follows:

$$A_j = \sum_{i=1}^n I_i w_{ij} \quad (3)$$

We used Hebbian learning to update the connection weights between layers, modeling the association between the activation of a group of neurons in DRPS, the activation of a specific set of neurons in PIFDL and the inhibition of others as follows:

$$\Delta w_{ij} = \alpha I_i w_{ij} G_j, \quad (4)$$

where α is the learning rate, and G_j is a new activation value of neuron j that depends on the number of the place occupied by the original activation level A_j within the hierarchy of activation levels in PIFDL. Weights are normalized between 0

and 1.

The third kind of sensory input to the model is the internal drive of the rat, which is related to its need to eat: the **hunger drive**. In general, drives d can be appetitive or aversive. The idea is that each appetitive drive spontaneously increases with every time step towards d_{\max} (input value), while aversive drives are reduced towards 0, both according to a factor α_d (input value) intrinsic to the animal. An additional increase occurs if an incentive $I(d,x,t)$ is present such as the sight of food. Drive reduction $a(d,x,t)$ takes place after food ingestion. If the animal is at place x at time step t , and the value of drive d at that time step t is $d(t)$, then the value of d for the animal at time step $t+1$ will be

$$d(t+1) = d(t) + \alpha_d |d_{\max} - d(t)| - a(d,x,t) |d(t)| + I(d,x,t) |d_{\max} - d(t)| \quad (5)$$

The amount of reward the animal gets by the presence of food is dependent on its current motivational state. If the rat is extremely hungry, the presence of food might be very rewarding, but if not, it will be less rewarding. In this way, the reward value depends on the current value of the animal's drive $d(t)$ and its corresponding maximum value, according to:

$$r(t) = d(t) / d_{\max} \quad (6)$$

B. The Place Cell Layer and the World Graph Layer

The pattern of kinesthetic information generated by PIFDL is the input to the **place cell layer** (PCL). The connection pattern between these two layers is the same as the one carried out between DRPS and PIFDL. The pattern of activity generated in PCL represents a single place or location in the environment.

The topological map is implemented by the **world graph layer** (WGL). Nodes represent distinctive places in the environment and store patterns of activity generated in PCL. If the current pattern in PCL is not recognized as stored in the map and the affordances sensed by the rat changed, then a new node is created and linked to the previous active node that represents the place currently visited by the rat. Each node can store eight different activity patterns, one for each direction, assuming that the animal can orient itself in eight directions and can experiment different views from the same place. The arcs in the graph are associated with the orientation of the rat's head when the animal went from one node to the next one. The determination of the direction of the rat's head is based on a global coordinate system, which is relative to the departure location in the exploration process (see Fig. 2 Right).

As we will describe in Section III, in the experiments tested the rat learns to find food locations in different mazes. To allow this goal oriented behavior we used **reinforcement learning** through an **actor-critic architecture** [17]. While the core ideas of reinforcement learning come from theories of animal classical conditioning, the influence of concepts from artificial intelligence and control theory has produced a

collection of computationally powerful learning architectures. Among Temporal-Difference Learning methods (such as Q-Learning), we decided to use the actor-critic architecture in our bio-inspired model since it has been related to the nervous system, specifically to the circuits of the basal ganglia and associated brain structures [18]. An actor-critic architecture processes expected values of future reinforcements through its components: an **adaptive critic unit** and **actor units**. The output of the critic unit is a prediction of the value of future reinforcement. In our model this predicted value is used to generate a reinforcement signal that is transmitted to WGL in order to reinforce eight actors (one per each rat's orientation) associated to every node in the map. The critic unit is implemented as a module of the model (not shown in Fig. 1) that supports the operation in WGL. This module receives the activity pattern registered in PCL as input and computes $p(t)$, the prediction value of future reinforcement at time t , using

$$p(t) = \sum_{i=1}^n v_i, \quad (7)$$

where v_i is the connection weight between neuron i in PCL and the critic unit. The weights added correspond to the most active neurons in PCL (we considered $n=5$ neurons). The connection weights to the critic unit are initialized to 0 and updated according to

$$v_i(t+1) = v_i(t) + \beta \hat{r}(t) \bar{x}_i, \quad (8)$$

where β is the learning rate; \bar{x}_i is the eligibility trace of the activity level of each neuron in PCL; and $\hat{r}(t)$ is the temporal difference error between any two adjacent predictions and is computed using

$$\hat{r}(t) = r(t) + \gamma p(t) - p(t-1). \quad (9)$$

The reinforcement learning process starts by updating \bar{x}_i values associated to the most active neurons in PCL. If the rat perceives food from its current location, the eligibility trace is increased, otherwise is decreased.

As previously mentioned, the actor modules are associated to the map nodes and represent the expectations of finding reinforcement in orienting to a certain direction at the current location. Every actor is implemented in a map node as a pair of "weight – eligibility trace". The reinforcement process is carried out in the actors when the reinforcement in the critic unit has finished, and consists on increasing the eligibility trace associated to the current rat's head direction in the active node in the map. Then, the actor weights are updated for all map nodes using:

$$w_k^d(t+1) = w_k^d(t) + \alpha \hat{r}(t) e_k^d(t), \quad (10)$$

where $w_k^d(t+1)$ is the actor weight associated to node k and direction d at time $t+1$, α is the learning rate, and $e_k^d(t)$ is the eligibility trace of actor d in node k at time t .

At a given location, the choice of the rat to turn to a specific direction is influenced by a general rewarding signal

generated through the expectations of future reward w_k^d of the actors. This rewarding signal is computed by the motivational schema in the model considering information that proceeds from the WGL. In this layer three nodes in the map are considered: the active one and two more in sequence. For each node the expectation of reinforcement values associated to the directions of the arcs pointing to other nodes are reviewed and the direction with the highest expectation value is selected. The different directions selected over the sequence of nodes as well as the corresponding expectation of reinforcement values are stored and sent to the motivational schema.

C. Motor Outputs

The mission of the **motivational schema** (MS) is to compute the input (*iss*) to the **action selection schema** (SS) by adding four perceptual schemas representing the affordances (*aff*), the food (*tf*), the global expectation of future reinforcement (*ger*), and the curiosity level (*cl*).

The food perceptual schema (*tf*) codifies the rotation the animal has to carry out to orient to food when it is visible from current location. When food is not visible, the rotation magnitude is determined randomly between available affordances. In this way, a random component is induced to the process of choosing a direction to orient to.

Through *cl* the model considers the fact that the animal may go to places that are not yet represented in the map when it is not motivated to go towards a previously experienced place. In this perceptual schema an exponential term like the one shown in (1) is added, corresponding to each available affordance associated to the direction of an arc not represented in the active map node.

MS uses the expectation of reinforcement values and the corresponding directions selected by WGL over the sequence of nodes to generate an expectation of future reward perceptual schema *efr*. The rotation the rat has to execute to orient to each direction is computed and added as an exponential term in this perceptual schema with strength depending on the expectation of reinforcement value associated with the direction, as follows:

$$efr_i = f_j e^{-\frac{(i-a)^2}{2d^2}}, \quad (11)$$

where efr_i is the activity level of neuron i in *efr*, f_j is the expectation of reinforcement value corresponding to direction j (the height of the exponential), d and a were defined in (1).

There can be at most three exponential terms associated to the three nodes in the sequence. In order to generate a global expectation of reinforcement signal (*ger*) that will influence the next behavior of the rat, the center of mass is computed. If there is no available affordance coinciding with the center of mass, it is moved to the neuron that corresponds to the selected direction from the active map node.

SS determines the next direction of the rat's head, from 0° to 315° , by considering the highest activation value in the

perceptual schema that receives as input (*iss*), the required rotation to point to the next direction, and the displacement. If the next direction is different from the current one, the displacement is set to 0 giving the rat the opportunity to perceive a different view from the same place. When the rat is returning to the departure point after having finished a trial in an experiment, SS computes the next rat's direction from the built map. This return process was documented in [19].

III. ROBOTIC ARCHITECTURE AND EXPERIMENTATION RESULTS

We tested the model using a Sony AIBO ERS-210 4-legged robot having a local camera. The model was designed and implemented using the **NSL system** [20] and can interact with a virtual or real environment through a **visual processing** module that takes as input the image perceived by either a simulated rat or a real robot, and a **motor control** module that executes rotations and/or translations on the rat or robot. We have used three different virtual and physically built experimental environments: a T-maze, an 8-arm radial maze and an extended maze. Different color papers were pasted over the walls of the mazes to simplify the recognition of the junctions, the food, the walls and the end of corridors, which is necessary to compute affordances and estimate distances.

The experiment carried out in the T-maze and in the 8-arm radial maze is inspired on the reversal task implemented by O'Keefe [21]. Our goal was not the replication of this experiment but to use it to validate our model by comparing our results with those obtained by O'Keefe qualitatively, and to extend the experiment into more complex robotic mazes.

O'Keefe divided the experiment in two phases: training and testing. The training phase was carried out in the T-maze and the testing phase was carried out combining trials in the T-maze with trials in the radial maze. We decided for simplicity to implement the reversal task in both the T-maze and the radial maze separately. We have also implemented it within an extended maze. The following subsections describe the robot experimentation results obtained in the three mazes.

A. Experiment I: T-Maze

In every experiment's trial in the T-maze, the rat navigates from the base of the T to either one of the two arm extremes, and then it returns to the departure location autonomously. At each step, the rat takes three pictures of the environment in different angles: 0° , $+90^\circ$ and -90° . During the experiment, the rat builds the map shown in Fig. 3(a). It is composed of seven nodes, each one created when the rat sensed a change in the available affordances and did not recognize the information pattern generated by PCL. Some of the nodes represent more than one place in the maze as can be seen in Fig. 3(b).

Table I summarizes the rat performance. During the training phase, the food is placed in the left arm of the maze and the rat turns left in every trial increasing the expectation of reward for the left arm. When the testing phase begins, the

food is moved to the right arm. The curiosity of the rat for that arm and the sight of food make the rat explore it. In the following trials, the rat goes through an unlearning process, where the expectation of reward for the left arm is the dominant influence in its behavior although it is decreasing continuously, while the expectation of reward for the right arm increases. When both levels of expectation are similar, the rat turns left or right randomly due to the noise level, starting a relearning process. From trail 32 the expectation of reward for the right arm is bigger than the one for the left arm, making the rat turn right consistently. After finishing a trial, the rat returns home [19]. Fig. 4 shows pictures of the robot's behavior and a "shortened" video is in [22].

B. Experiment II: 8-Arm Radial Maze

In the 8-arm maze shown in Fig. 5 the rat navigates from the 270° arm to any other arm extreme in every experiment's trial. Fig. 6 (a) presents the map built by a rat/robot during the experiment. We divided the experiment in three phases: training, pre-testing and testing. Table II summarizes the robot performance. Fig. 5 shows pictures of the robot's behavior and a video can be found in [22].

In terms of learning, our results match qualitatively with those obtained by O'Keefe experimenting with normal rats despite some variations done to the original experiment. As in O'Keefe's rats, the orientation of our rats after the training phase did not shift in a smooth manner but randomly as shown in Fig. 6 (b, c).

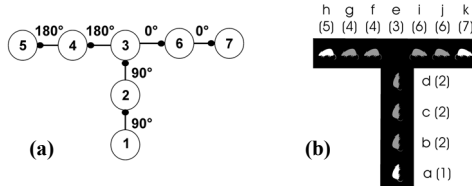


Fig. 3. (a) The map built by the rat during the reversal task in the T-maze. Nodes are numbered in order of creation. Arcs between nodes show the orientation of the rat when it moved from one node to the next one. (b) The T-maze locations are labeled with letters and associated to the nodes of the map.

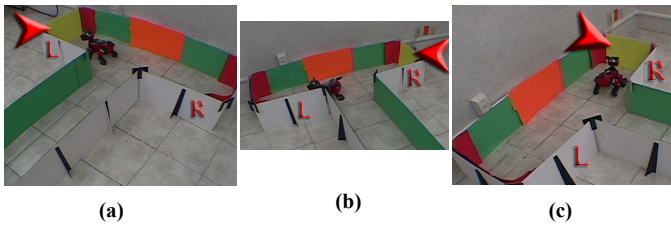


Fig. 4. (a) A typical trial in the training phase: the food is at left (L) and the robot is approaching it. (b) A trial during the unlearning process: the food is at right (R) and the robot is approaching the end of L. (c) A trial during the relearning process: the food is at right and the robot is approaching it.

TABLE I. THE PERFORMANCE OF THE RAT DURING THE T-MAZE EXPERIMENT.

Trial #	Chosen arm	Phase	Process
1 – 10	Left	Training	Learning
11	Right	Testing	Curiosity drive
12 – 19	Left	Testing	Unlearning
20 – 31	Left or right randomly	Testing	Relearning
32 – ...	Right	Testing	Relearning

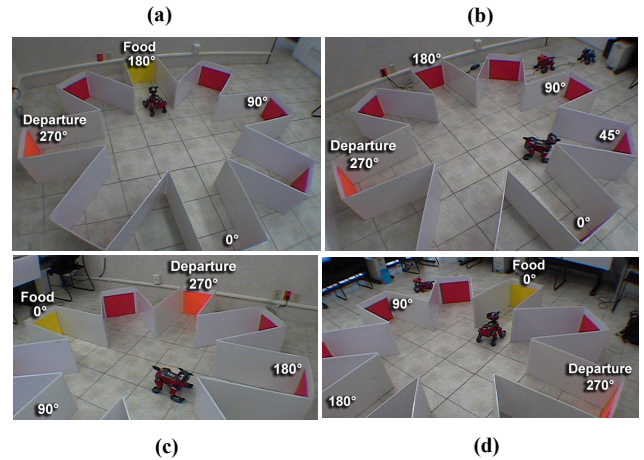


Fig. 5. (a) A typical trial in the training phase: the food is at 180° arm and the robot is approaching it. (b) A trial in the pre-testing phase: the food has been removed from the maze and the robot chooses any arm not yet visited. (c) A trial in the testing phase: the food is at 0° arm and the robot is approaching the 180° arm. (d) The robot is approaching the new reward arm (0°).

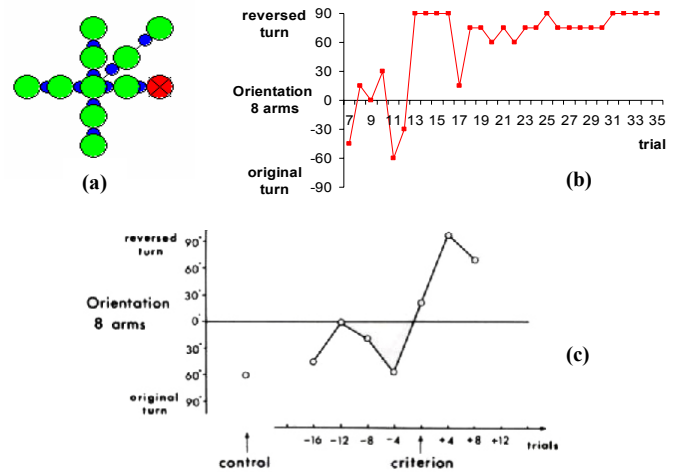


Fig. 6. (a) The map built by a rat/robot during the reversal task in an 8-arm radial maze. (b) Average performance of four simulated rats during the pre-testing and testing phases reported by our model. (c) Average performance of four real and normal rats during the testing phase reported by O'Keefe in [21]. In (b) and (c) rats turn randomly to different arms between the original turn (-90°) and the reversed turn (+90°).

TABLE II. THE PERFORMANCE OF THE RAT IN AN 8-ARM RADIAL MAZE.

Trial #	Chosen arm	Phase	Process
1 – 5	180°	Training	Learning
6 – 10	315°, 45°, 90°, 225°, 135° or 0° randomly	Pre-testing	Randomness and curiosity drives
11 – 12	180°	Pre-testing	Unlearning
13 – 17	180° or 0° randomly	Testing	Relearning
18 – 22	0°	Testing	Relearning

C. Experiment III: Extended Maze

We decided to extend the reversal task by considering a maze where the food was not visible by the rat at the first choice point but at the second one. To try this, we designed and built the maze shown in Fig. 7.

Since the rat does not see food at the first choice point, we had to extend the model to propagate the reinforcement from the goal location back towards the departure place in the map built while the rat returns from a trial. If the goal is reached by

the end of a trial, the eligibility trace of the actor corresponding to the arc direction of each map node in the path is increased, otherwise is decreased. Considering the map built by the rat in Fig. 7 (d, e), and supposing that it has reached the goal place in the training phase, the reinforcement process consists on increasing the following eligibility traces in sequence: actor 90° in node 6, actor 90° in node 5, actor 180° in node 4, actor 180° in node 3, actor 90° in node 2, and actor 90° in node 1. Table III summarizes the rat/robot behavior, related pictures are shown in Fig. 7 (a, b, c), and a video can be found in [22].

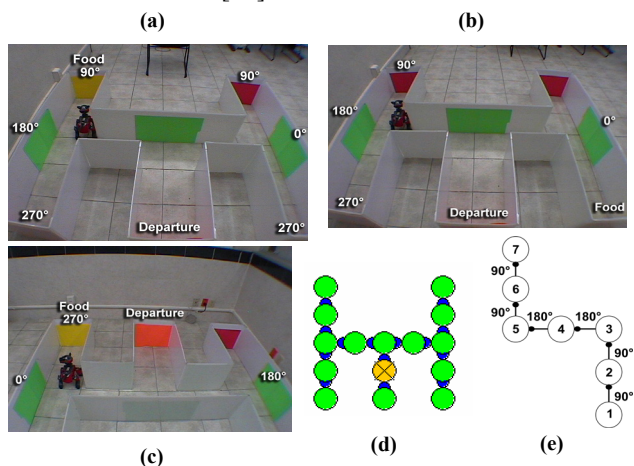


Fig. 7. (a) A trial in the training phase: the robot is approaching the food in the 180° arm. (b) A trial during the testing phase (unlearning process): the food is at 0° arm, but the robot still searches for it at the previous reward arm. (c) A trial during the relearning process: the robot is approaching the new reward location. (d) The map built by the robot. (e) The actors whose eligibility traces were reinforced positively during the training phase.

TABLE III. THE PERFORMANCE OF THE RAT IN THE EXTENDED MAZE.

Trial #	Chosen arm (1 st choice point)	Chosen arm (2 nd choice point)	Phase	Process
1 – 15	180° or 0° randomly	If the 1 st choice was 0°, the 2 nd one is 90° or 270° randomly. If the 1 st choice was 180°, the 2 nd one is 90°.	Training	Positive path reinforcement during at least five trials 180° – 90° (the path to food)
16	180°	270°	Testing	Curiosity
17 – 18	180°	90°	Testing	Unlearning
19 – 23	180° – 90° or 0° – 270°		Testing	Relearning
24 – 32	0°	270°	Testing	Relearning

IV. CONCLUSIONS

In this paper we have presented the adaptive learning and mapping capabilities of a robotic navigation model based on the physiology of the rat’s brain. We have shown that the rat/robot is able to learn the locations of food dynamically moved in different mazes, and to build a topological map of the environment. We plan to work in validating the consistency of the maps built by the model through solving cyclic mazes and recognizing places already visited. To do

this, we will merge information about spatial landmarks with kinesthetic information to identify the places of the environment. In the medium term we hope to propose a bio-inspired alternative solution to some of the SLAM challenges.

REFERENCES

- [1] J. Guivant, E. Nebot, J. Nieto, F. Masson. “Navigation and Mapping in Large Unstructured Environments.” *The International Journal of Robotics Research*, vol. 23, no. 4, pp. 449-472, 2004.
- [2] B. Kuipers, J. Modayil, P. Beeson, M. MacMahon, F. Savelli. “Local Metrical and Global Topological Maps in the Hybrid Spatial Semantic Hierarchy.” *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, New Orleans, USA, April, 2004.
- [3] Z. Zivkovic, B. Bakker, B. Kröse. “Hierarchical Map Building Using Visual Landmarks and Geometric Constraints.” *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Edmonton, Canada, August, 2005, pp. 7-12.
- [4] M. Bosse, P. Newman, J. Leonard, S. Teller. “SLAM in Large-scale Cyclic Environments using the Atlas Framework.” *International Journal on Robotics Research* 23(12):1113–1139, December, 2004.
- [5] J. Folkesson, H. Christensen, “Graphical SLAM - A Self-Correcting Map.” *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, New Orleans, USA, April 2004.
- [6] D. Hähnel, R. Triebel, W. Burgard, S. Thrun. “Map Building with Mobile Robots in Dynamic Environments.” *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1557-1563. Taipei, Taiwan, September 2003.
- [7] E. Tolman. “Cognitive maps in rats and men.” *Psychological Review* 55, pp. 189-208, 1948.
- [8] J. O’Keefe and L. Nadel. “The hippocampus as a cognitive map.” Oxford University Press, 1978.
- [9] N. Burgess, M. Recce and J. O’Keefe. “A model of hippocampal function.” *Neural Networks*, Vol. 7, Nos. 6/7, pp. 1065-1081, 1994.
- [10] D. Touretzky and A. Redish. “A theory of rodent navigation based on interacting representations of space.” *Hippocampus* 6, pp. 247-270, 1996.
- [11] K. Balakrishnan, R. Bhatt and V. Honavar. “A computational model of rodent spatial learning and some behavioral experiments.” In Gernsbacher & Derry (Eds.). *Proceedings of the 20th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ, Lawrence Erlbaum Assoc., 1998.
- [12] O. Trullier and J-A. Meyer. “Animat navigation using a cognitive graph.” *Biological Cybernetics* 83, pp. 271-285, 2000.
- [13] P. Gaussier, A. Revel, J. P. Banquet and V. Babeau. “From view cells and place cells to cognitive map learning: processing stages of the hippocampal system.” *Biological Cybernetics* 86, pp. 15-28, 2002.
- [14] A. Guazzelli, F. J. Corbacho, M. Bota, and M. A. Arbib. “Affordances, motivation, and the world graph theory.” *Adaptive Behavior*, vol. 6 (3/4), pp. 435-471, 1998.
- [15] M. Milford and G. Wyeth. “Hippocampal Models for Simultaneous Localization and Mapping on an Autonomous Robot.” *Proceedings of the Australasian Conference on Robotics and Automation*. Brisbane, 2003.
- [16] D. O. Hebb. “The Organization of Behavior: A Neuropsychological Theory.” Wiley-Interscience, New York, 1949.
- [17] R. S. Sutton and A. G. Barto. “Reinforcement Learning: An Introduction.” MIT Press, Cambridge, MA, 1998.
- [18] A. G. Barto. “Adaptive Critics and the Basal Ganglia.” In J. C. Houk, J. L. Davis and D. Beiser (Eds.), *Models of Information Processing in the Basal Ganglia*. Cambridge, MA: MIT Press, 1995, pp. 215-232.
- [19] A. Barrera and A. Weitzenfeld. “Return of the Rat: Biologically-Inspired Robotic Exploration and Navigation.” *Proceedings of the 1st IEEE / RAS-EMBS International Conference on Biomedical Robotics and Biomechanics (BioRob)*. Pisa, Italy. February 20 – 22, 2006, in press.
- [20] A. Weitzenfeld, M. Arbib and A. Alexander. “The Neural Simulation Language.” MIT Press, 2002.
- [21] J. O’Keefe. “Spatial memory within and without the hippocampal system.” In W. Seifert (Ed.), *Neurobiology of the Hippocampus*. New York, Academic Press, pp. 375 – 403, 1983.
- [22] A. Barrera and A. Weitzenfeld. “Robot experimentation results.” [ftp://ftp.itam.mx/pub/alfredo/ABarrera/Videos2005/](http://ftp.itam.mx/pub/alfredo/ABarrera/Videos2005/)