Learning to navigate, progress and self-evaluation

C. Giovannangeli and P. Gaussier
ETIS Lab (CNRS 8051), Cergy Pontoise University

For the last decade, we have developed a neural network architecture inspired from the brain functioning, efficient for visual navigation and imitation tasks. This architecture involves a model of the visual system, the hippocampus, the prefrontal cortex, the basal ganglia and the motor cortex [2]. The use of a robot standing for a simulation of an animal or a human provides an efficient mean to validate our models and verify if the global dynamics of the robot/environment interactions corresponds to those observed by the neurobiologists and the psychologists. As navigation is concerned, our model enables a robot to learn visual landmarks, to associate them to their spatial properties (azimuth and elevation) in order to build a constellation of landmarks (a set of triplet *landmark-azimuth-elevation*). The activity of the neurons recognizing such a constellation is similar to the activity of the large place cells recorded in the entorhinal cortex of the rat [10]. A set of *place-action* associations, learned in one-shot creates an attraction basin enabling our robot to go back to a learned location or to follow an arbitrary visual path in indoor environments [5]. The robustness of our visual place cells has recently been optimized for outdoor environments [7], enabling the robot to achieve sensory-motor tasks in indoor as well as in large outdoor environments with a low computation load [8] (see fig. 1). The behavior is also robust to kidnapping, to object and landmark addition or removal, to the presence of mobile obstacles and to severe visual field occlusions[1].

FIG. 1 – Sensory-motor trajectory in outdoor environment. Arrows represents the learned sensory-motor associations.

[1] movies available on
http ://www.etis.ensea.fr/~neurocyber/Videos/homing/index.html
or on the home pages of the authors.

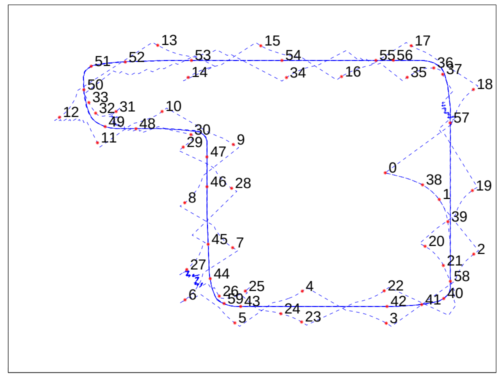

FIG. 2 – Simulation of the place-action learning during a path-learning session. The figure shows the trajectory during the learning (clockwise travel). Numbers correspond to the position and the order of the places autonomously learned by the robot. The professor first forbids the robot to escape from the path (two laps of proscriptive learning), and then guides the robot on the precise trajectory (several laps of prescriptive learning). The number of learned places stops increasing quickly.

In the context of long life learning, we are presently interested in addressing the problem of the semi-supervised building of the attraction basin and its refinement. Besides, we focus here on the capability of the robot to learn autonomously a sensory-motor task by imitating a human *teacher* [1]. Being guided by the human, the robot learns places (as illustrated by the fig. 2) and is able to compare the action associated with the current state (here places) to the action imposed by the teacher. A sensory-motor error between its *a priori* action and the proposed action in each places is available. For a given state, the learning progress is defined as the difference between the learned error and the current error. Previous works have shown how close is the maximisation of the learning progress to the manifestation of curiosity [9] in a completely unsupervised context. Here, the interaction with a teacher is necessary to learn the task since there is no environmental constraint that could guide the learning. Furthermore, [9] do not raise the problem of environmental changes that can lead learned associations to become erroneous. We propose a progress-based metacontroller that supervises a predictor by suspending

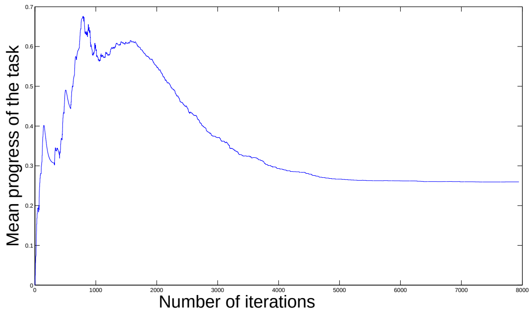

FIG. 3 – Mean progress of the task during the training phase. The curve shows the sum of the mean progress associated to all the places which can be viewed as the learning progress of the task. After a while, the progress does not change anymore : the learning can be suspended.

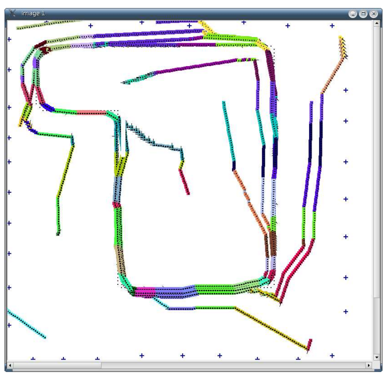

FIG. 4 – Simulated sensory-motor trajectories after the learning (*i.e.* when no more progress is expected by the robot). The crosses on the border of the environment represents the landmarks. Most of the trajectories converge to the precise attraction basin corresponding to the learned trajectory. However, if the starting point is far enough from the attraction area, the robot trajectory can diverge, as illustrated by the bottom left starting point.

or triggering the adaptation. The metacontroller evaluate when no more progress is expected and when the predictions are no longer carried out by means of a progress-based novelty-habituation detector, providing the robot to access with a self-evaluation measure (see fig. 3). We show the learning procedure is stable and accurate allowing the precise reproduction of an arbitrary trajectory according to very few global parameters that will be discussed in the poster (see fig. 4).

Future work will focus on the comparison of sensory-motor strategies versus planning strategies for the learning of an arbitrary path and the control of its reproduction. In our complete biological model, neurons in the hippocampus proper (CA1/CA3 regions) learn and predict transitions between successive multimodal states [6]. A cognitive map computes a latent learning of the spatial topology of the environment [11] and can be used to plan a sequence of actions to reach an arbitrary goal [4]. The influence of our progress-based metacontroller will be evaluated at all the level of this architecture. We will also study how an agent can autonomously detect it is not really doing what it aims at doing (as in the case of the bottom left starting point of fig. 4, when the robot get lost). We will wonder how an emotional system could be used as a second order controller [3] to adjust the shape of the attraction basins provided by the sensory-motor or the planning systems when the behavior becomes incorrect.

[1] P. Andry, Ph. Gaussier,J. Nadel and B. Hirsbrunner. Learning Invariant Sensorimotor Behaviors : A Developmental approach to Imitation Mechanisms. Adaptive behavior. 12(2) : 117-138, 2004.

[2] J.P. Banquet, P. Gaussier, M. Quoy, A. Revel, and Y. Burnod. A hierarchy of association in hippocampo-cortical systems : cognitive maps and navigation strategies. Neural Computation, 17 :1339–1384, 2005.

[3] L. Canamero and P. Gaussier. Emotion understanding : robots as tools and models. In Emotional Development. Ed by J. Nadel and D. Muir, Oxford university Press. pp235–258, 2005.

[4] N. Cuperlier, M. Quoy, P. Gaussier and C. Giovannangeli, Navigation and planning in an unknown environment using vision and a cognitive map, In IJCAI-05 Workshop : Reasoning with Uncertainty in Robotics (RUR-05), pp48-53, 2005,

[5] 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.

[6] 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, vol 86, pp15–28, 2002.

[7] C. Giovannangeli, P. Gaussier and J. Banquet. Robot as a tool to study the robustness of visual place cells. I3M'2005 : Int. Conf. on Conceptual Modelling and Simulation, pp 97–104, 2005.

[8] C. Giovannangeli, P. Gaussier and G. Désilles, Robust Mapless Outdoor Vision-Based Navigation, To appear in Proc. of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems.

[9] P-Y Oudeyer, F. Kaplan, V.V. Hafner and A. Whyte. The playground experiment : task-independent development of a curios robot. Proc. of the AAAI Spring Symp. Workshop on Developmental Robotics, pp 42-47, 2005.

[10] G.J. Quirk, R.U. Muller, J.L. Kubie and J.B. Ranck, Jr. The positional firing properties of medial entorhinal neurons : Description and comparison with hippocampal place cells. Journal of Neuroscience, 12(5) :1945–1963, 1992.

[11] E.C. Tolman. Cognitive maps in rats and men. The Psychological Review. 55(4) :189–208, 1948.