



HAL
open science

Learning New Behaviors: Toward a Control Architecture Merging Spatial and Temporal Modalities.

Matthieu Lagarde, Pierre Andry, Christophe Giovannangeli, Philippe Gaussier

► To cite this version:

Matthieu Lagarde, Pierre Andry, Christophe Giovannangeli, Philippe Gaussier. Learning New Behaviors: Toward a Control Architecture Merging Spatial and Temporal Modalities.. 2008. hal-00356676

HAL Id: hal-00356676

<https://hal.science/hal-00356676>

Preprint submitted on 28 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Learning new behaviors : Toward a Control Architecture merging Spatial and Temporal modalities

Matthieu Lagarde*, Pierre Andry*, Giovannangeli Christophe* and Philippe Gaussier*[†]

Email: {lagarde,andry,giovannangeli,gaussier}@ensea.fr

*Neurocybernetic team, ETIS Laboratory, CNRS UMR 8051, Univ of Cergy-Pontoise, ENSEA, Pontoise 95000, France

[†]Institut Universitaire de France

Abstract—This paper discusses the role of two antagonist neural networks for the learning and control of complex behaviors composed as a sequence of elementary states. Learning a pathway with a mobile robot or a sequence of actions with a robot arm can be seen either as the result of the learning of a temporal sequence or as the result of the natural dynamics of a sensory-motor system using appearance based approaches for instance. As a result, we will discuss the performances and the complementary features of each system, and propose a unique control architecture embedding both systems for long life learning.

I. INTRODUCTION

Our long term goal is to design a control architecture allowing a robot to learn, as autonomously as possible, sequences of actions related either to spatial or temporal constraints (displacements between places or gestures for instances). Learning a behavior is often related to learning by reinforcement, by demonstration or learning by imitation. Learn by imitation has often been considered as a complex behavior, but in previous work we have showed that the imitation can emerge from elementary mechanisms. For example: a robot that learns a “behavior” consisting in moving at different places and performing some very simple but different manipulations of different objects at each places as shown in figure 1. This

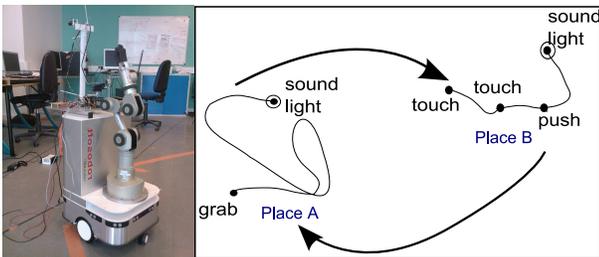


Fig. 1. Scenario illustrating our long term goal.

objective raises the issue of learning a behavior composed of actions : the nature of the relevant information is different between “moving from a place to another”, and “using an arm to push an object”. Indeed, the working spaces, the type of inputs, the motor commands, are different.

In the field of navigation systems, learning a sequence of

displacements between places (also known as navigation and planning) is strongly related to the localization and mapping (see for example the SLAM literature). The sequence is generally the result of a plan composed of motor actions or the result of an imitation (wheel orientation and speed) associated to the recognition of places (localization) anchoring the behavior in the robot’s cognitive map.

In the field of “manipulating systems”, i.e. non-mobile systems performing gestures and/or object manipulation. Different models propose to learn and adapt motor trajectories of the mechanical system in order to fit with the desired one of the model. The sequence is strongly related to the dynamical parameters allowing shaping the trajectory of the arm’s joints in order to obtain the right reproduction of the behavior.

This very short presentation of two important fields of autonomous robotics illustrates how complex the issue of building a global system that deals with navigation and arm movements as a single problem is. Our approach implicitly raises two crucial questions : how to build control architecture for articulated and mobile robots (to consider manipulation and navigation as a single problem)? How to build a neural architecture for spatial and temporal sensory-motor learning in which each modality could complement, confirm, infirm and/or enrich the other? What are the minimal requirements for such a merging? Which level for fusion making? Which coding to employ? In order to start to answer to these issues, we compare two models in the purpose of a unified model. Both solutions are based on artificial Neural Networks (NN) inspired from different properties of the cerebellum and the hippocampus loop.

II. MODELS

Complex Temporal Sequences. The model allows a robot to learn a sequence as a succession of transitions between the different sensory-motor situations. An associative learning rule allows learning and predicting the timing of the transitions. Moreover, neural oscillators composed of coupled CTRNN [Beer, 1994], play the role of an internal context and provide additional information in order to remove ambiguities in complex sequences [Lagarde et al., 2007]. Applied to the navigation, the sequence is based on the succession of

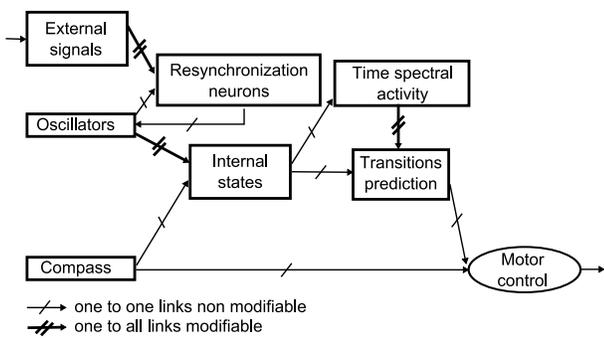


Fig. 2. Model of complex temporal sequences or orientations.

orientations (orientation is obtained from the compass). One of the main problems was the time spent by the robot to turn delaying the perception of the orientation during the reproduction. During this time lag, the internal context (i.e. the activities of the oscillators) changes. Consequently, the system loses the internal state and fails to reproduce the sequence. In order to avoid this problem, we propose to resynchronize the oscillators, according to external signals, when a new internal state is learnt. The context can be associated according to a Least Mean Square (LMS) learning rule. The property of resynchronization is crucial so that the system is able to correctly reproduce a sequence. This property is close to the one used in others models like the Echo States Network (ESN) [Jaeger, 2001].

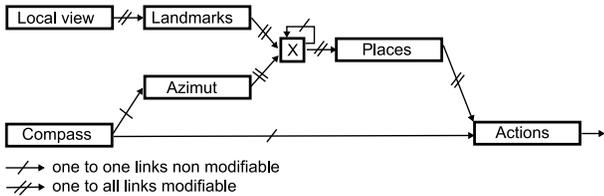


Fig. 3. Model of associations between places and actions.

Association between places and actions. This model [Giovannangeli and Gaussier, 2007] associates places with actions (figure 3). A place is a constellation of visual features (landmark, azimuth). The constellation results from the merging of “what” information provided by the visual system that extracts local-view centred on points of interest. The “where” information provided by the compass. A simple associative learning between places and actions enables to generate a sensory-motor attraction basin for homing or path following behaviors.

III. ROBOTIC APPLICATION

The robot used in our experiments is a Robulab10 (Robosoft) with a pan-tilt video camera and a compass.

Association between places and actions. During the learning, the robot moves in the environment (figure 4). When the robot escapes too far from the desired trajectory, we correct it with a joystick as a dog guided with a leash. At this moment, the NN learns online a new association between the correct

of motor command and the current place. When the robot recognizes a place, it triggers the associated action. After 3 or 4 iterations of learning, the robot navigates autonomously without correction from the teacher. The system does not try to recognize a place, but use a competitive mechanism between the learnt associations to build an attraction basin. The system adapts to the dynamic of the environment (obstacles, others agents).

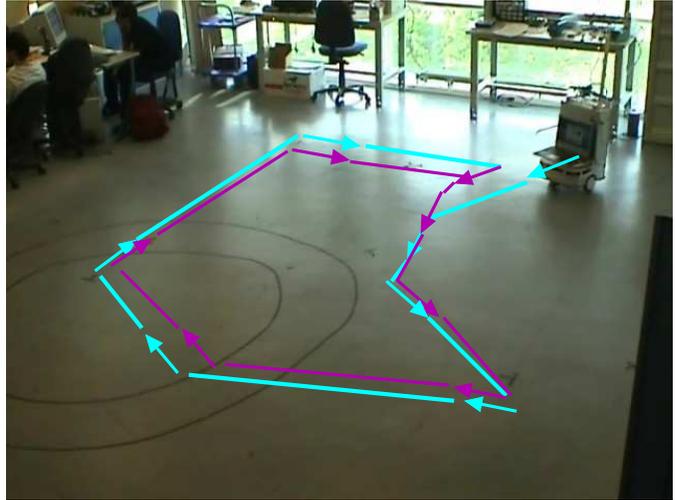


Fig. 4. Spatial navigation: picture of learnt (light arrows) and reproduced (dark arrows) trajectories by the robot with places-actions associations.

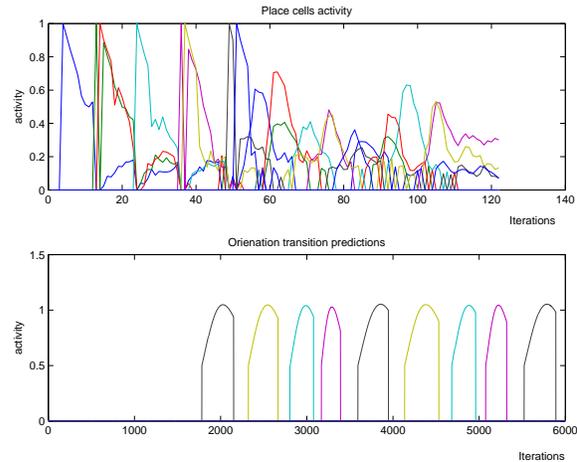
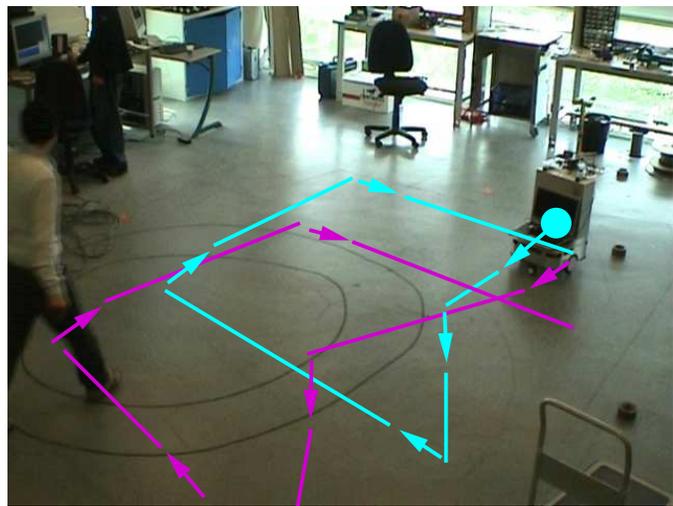


Fig. 5. Each curves represent the activity of each neuron coding place (Up) or transition prediction (Down). Up : responses of the place cells during learning (from iteration 0 to 55) and reproduction (from iteration 55) allowing the triggering of each action of the trajectory. Bottom: Activity of the transition prediction group, allowing the reproduction of the sequence of orientations with the correct timing

Complex Temporal Sequences. During the learning, the robot moves in the environment as shown in figure 6. When the robot makes mistakes, we use a joystick to correct its trajectory. At this moment, the NN learns online (one shot learning) a new transition between the previous and the

new orientation. To initiate the sequence of displacement, we set the robot heading to the first learnt orientation (the robot moves at a constant speed). The orientation information triggers on time the prediction of the next orientation that will drive the robot's new rotation, and begin the step by step reproduction of the sequence : each new orientation is recognized and resynchronizes the oscillators, inducing the next prediction and the realization of the associated action.



— Learning of the sequence of orientation
 — Reproduction of the sequence of orientation

Fig. 6. Learning and reproduction of temporal sequences: picture of another learnt trajectory (light arrows) and reproduced trajectory (dark arrows) by the robot.

IV. DISCUSSION

The architecture that learns places-actions associations has shown to be robust and reliable. It allows the robot to successfully navigate indoor as well as outdoor. In parallel, learning sequences of orientations has been successfully used in previous works in the frame of imitation with mobiles or articulated robots. Of course, the robustness of the navigation is strongly dependent on the quality of the visual environment. If the visual mechanism has shown to be robust to partial changes of the environment, a failure of the camera or very bad lighting conditions will prevent the system from working. Considering this, learning the sequence of orientations for a simple navigation task becomes interesting. Indeed, the robot uses little information from the environment : only a detection of the orientation variations. During the reproduction of the sequence, the robot acts as a “blind” automata. It can work correctly during little iteration without visual information. Resynchronization of internal dynamics with the current state is necessary after a while. It can not adapt to sudden changes of the environment (e.g. a new obstacle). Nevertheless, we think that the models of places-actions associations and sequence learning should work in parallel. Each architecture seems to

complete the other one in order to learn the spatial and temporal properties of complex behaviors. Moreover, the learning of the timing of the orientation changes should (1) contribute to confirm or infirm the visual place recognition (being the right orientation at the right time), (2) punctually replace the place cells if their activity is not strong enough (bad visual condition, conflict between different places) and (3) contributes to build long sequences, allowing to concatenate behaviors composed of displacements with those made of sequences of manipulations. Neurobiological and psychological studies suggest that both types of learning cohabit in the brain of mammals. For example, The results in [Packard and McGaugh, 1996] show the different roles of the hippocampus and basal ganglia in the task learning with different learning scales and learning rates, and the implication of different modalities (visual vs. proprioceptives). Hence, we propose a new architecture in Figure 7. This new architecture shows the model connecting

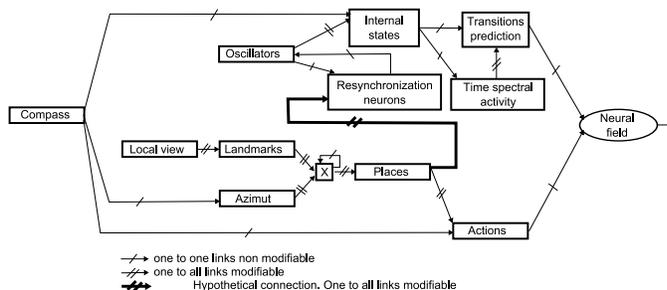


Fig. 7. Proposition of unified model with a hypothetical connection (bold arrow) could allow re-synchronizing the internal context (i.e. the oscillators) on place recognition signals.

both predictors. In order to merge the predictions, both systems will have their outputs merged in one neural field [Amari, 1977], [Schöner et al., 1995] allowing the cooperation of the predictions in case of similar responses, but also their competition in the case of too different responses (capacity of bifurcation of the neural field). Moreover, the neural field will allow coping easily with two systems working at different time scales. The emerging behavior will be the result of two subsystems having different dynamics and categorizing, predicting complementary modalities.

In a future robotic experiment, this new model will also help to enhance human/robot interaction allowing a mobile robot to learn the navigation path directly from following a naive user. During the displacement, the robot will focus on the demonstrator and will learn online the temporal succession of its orientations (short time learning). To anchor the displacement in the environment (which is not possible when focusing on the naive user), the robot will reproduce by oneself the displacement (i.e. the sequence of successive orientations) and will learn during this reproduction the associations between places and actions (long time learning). This experiment would help to study how experiences are stored in the brain. Moreover, it would help to study how and why the brain needs to use different kinds of memories according to learn and store behaviors between the episodic memory (hippocampus) and

the long term memory of the know-how (basal ganglia and/or cortical structures).

In the purpose of using an arm mounted on the mobile robot, it is interesting to anticipate that a similar visual mechanism as the “places cells” could guide the arm (for example, the location of interesting visual objects). This mechanism could allow anchoring in the visual working space of the arm temporal sequences of gestures, as well as the navigation model anchors actions in the wide visual environment. Previous works on robot arms have show the importance of the visuo-motor learning for gesture imitation. This solution consists in learning on a multi-modal map the associations between the motor and the visual information of the end-effector.

ACKNOWLEDGMENT

This work is supported by the French Region Ile de France, the FEELIX GROWING European project (FP6 IST-045169), the French Direction Générale des Armées (DGA) and CNRS Neuroinformatique project.

REFERENCES

- [Amari, 1977] Amari, S. (1977). Dynamic of pattern formation in lateral-inhibition type by neural fields. *Biological Cybernetics*, 27:77–87.
- [Beer, 1994] Beer, R. (1994). On the dynamics of small continuous-time recurrent networks. Technical Report CES-94-18, Cleveland, OH.
- [Giovannangeli and Gaussier, 2007] Giovannangeli, C. and Gaussier, P. (2007). Human-robot interactions as a cognitive catalyst for the learning of behavioral attractors. In *16th IEEE on Robot and Human Interactive Communication 2007*, pages 1028–1033.
- [Jaeger, 2001] Jaeger, H. (2001). The “echo state” approach to analysing and training recurrent neural networks. GMD Report 148, GMD - German National Research Institute for Computer Science.
- [Lagarde et al., 2007] Lagarde, M., Andry, P., and Gaussier, P. (2007). The role of internal oscillators for the one-shot learning of complex temporal sequences. In *ICANN 2007*, volume 4668 of *LNCS*, pages 934–943.
- [Packard and McGaugh, 1996] Packard, M. and McGaugh, J. (January 1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of Learning and Memory*, 65:65–72(8).
- [Schöner et al., 1995] Schöner, G., Dose, M., and Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16(4):213–245.