



HAL
open science

On-Line Learning and Planning in a Pick-and-Place Task Demonstrated Through Body Manipulation

Antoine de Rengervé, Julien Hirel, Mathias Quoy, Pierre Andry, Philippe Gaussier

► **To cite this version:**

Antoine de Rengervé, Julien Hirel, Mathias Quoy, Pierre Andry, Philippe Gaussier. On-Line Learning and Planning in a Pick-and-Place Task Demonstrated Through Body Manipulation. IEEE International Conference on Development and Learning (ICDL) and Epigenetic Robotics (Epirob), Aug 2011, Frankfurt am Main, Germany. pp.1-6, 10.1109/DEVLRN.2011.6037336 . hal-00631114

HAL Id: hal-00631114

<https://hal.science/hal-00631114>

Submitted on 11 Oct 2011

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.

Continuous On-Line Learning and Planning in a Pick-and-Place Task Demonstrated Through Body Manipulation.

Antoine de Rengervé, Julien Hirel, Pierre Andry, Mathias Quoy and Philippe Gaussier
ETIS, CNRS ENSEA University of Cergy-Pontoise, F-95000 Cergy-Pontoise, France
{antoine.rolland-de-rengerve,hirel,andry,quoy,gaussier}@ensea.fr

Abstract—When a robot is brought into a new environment, it has a very limited knowledge of what surrounds it and what it can do. One way to build up that knowledge is through exploration but it is a slow process. Programming by demonstration is an efficient way to learn new things from interaction. A robot can imitate gestures it was shown through passive manipulation. Depending on the representation of the task, the robot may also be able to plan its actions and even adapt its representation when further interactions change its knowledge about the task to be done. In this paper we present a bio-inspired neural network used in a robot to learn arm gestures demonstrated through passive manipulation. It also allows the robot to plan arm movements according to activated goals. The model is applied to learning a pick-and-place task. The robot learns how to pick up objects at a specific location and drop them in two different boxes depending on their color. As our system is continuously learning, the behavior of the robot can always be adapted by the human interacting with it. This ability is demonstrated by teaching the robot to switch the goals for both types of objects.

I. INTRODUCTION

Autonomous robots that start with no knowledge about the world and have to learn everything should not only depend on the unguided exploration of their possibilities for their development. They also should be able to learn from human caregivers or other robots by imitating them [1] or interacting with them. A simple visuo-motor associative learning and a motor controller can provide the basis for letting low level immediate imitation emerge as a result of simple mechanisms [2]. In a first phase, the system learns the body configuration as visuo-motor associations. Due to an ambiguity of perception caused by limited visual capabilities, the robot can confuse the hand of the human teacher with its own. The controller, following the homeostatic principle, tries to maintain an equilibrium between the visual and motor information according to the body learning. To do so, it moves the robot hand at the motor position that corresponds to the believed visual position of the robot hand. The robot consequently reproduces the observed gestures of the human. With almost the same mechanisms and the same controller, the robot can also learn by observation how to reproduce a demonstrated task encoded as a sequence of visual positions [3]. Visual demonstration cannot always be enough to teach what is expected from the learner. During their first years, infants are also taught by “embodying” them i.e. by using passive manipulation. This process enables caregivers to

guide the infant’s perception and attention so that the infant quickly learns what is taught [4].

Programming by demonstration [5] provides robots with the ability to learn how to perform a task as a result of a demonstration provided by a human teacher. This demonstration can be done through passive manipulation. Dynamic Motion Primitives (DMP) [6] is a framework for learning to perform demonstrated tasks. With little demonstration data from a human demonstrator, the system can determine through regression the adequate parameters for mixing the DMP to generate the correct movement. In [7], the learning system uses a statistical model based on Gaussian Mixture that is adapted to fit the data from training demonstrations using passive manipulation. Gaussian Mixture Regression can then provide a probabilistic estimation of the current state of the robot and of the adequate motor command. These systems have proved to be efficient for performing different tasks. They also enable generalization to different initial conditions and different goal positions.

One of the main issues is how to encode and represent the knowledge acquired through demonstration. The chosen model should enable the robot to continuously learn and integrate new knowledge into the representation of the task. Besides, this model should provide the basis for future action planning. Models with both learning by demonstration and planning have been developed in different works [8] [9] as programming by demonstration can speed up the learning of the knowledge needed for planning. In [10], path planning and Gaussian processes have been combined to use planning capabilities and obstacle avoidance while the robot is reproducing a demonstrated grasp-and-pour task.

In sec. II, an existing bio-inspired model for planning in mobile robot navigation task is used to learn the movements of a robotic arm. The neural network is adapted to work with proprioceptive states. A visual categorization system for goals corresponding to different types of object is implemented. Associations between object goals and proprioceptive state transitions are learned and adapted thanks to reinforcement learning. This system was implemented on a real robot to carry out an experiment of pick-and-place with objects, sorting them depending on their colors (sec. III). Initially, the robot learns the actions to be done after one demonstration for each type of object. As it is continuously learning, new demonstrations are

integrated in the model and can be used to adapt the behavior of the robot to match new conditions in the task i.e. a switch of the target boxes for each type of object (sec. IV).

II. ADAPTATION OF A BIO-INSPIRED COGNITIVE MAP MODEL TO ARM CONTROL

We previously developed a bio-inspired model of the hippocampus and prefrontal cortex [11]. In our model, spatial states would be encoded in the entorhinal cortex and the hippocampus would learn to associate successive states, thus creating transitions cells. The prefrontal cortex would use information about these transitions to create a cognitive map and use it to plan actions toward the fulfillment of specific goals. This model was later used on a real robot navigating in an indoor environment [12]. The aim of this paper is to reuse the principles of the model, result of a long series of experiments with mobile robots, and adapt them for planning in the proprioceptive space of a robotic arm. In addition, we take full advantage of the possible interactions between human and robot to speed up the learning process and allow for easy adaptation of the conditions of a pick-and-place task. As there are no separate phases of learning and reproduction, the system can adapt at any time. The model is shown in Fig. 1.

According to our model, some brain structures could be involved in both navigation and arm control. The categorization of states would involve different areas in navigation (the entorhinal cortex with place cells) and in arm control (motor cortex for proprioceptive states). However, in both cases the hippocampus would be used to detect new events as changes in the encoded states. The prefrontal cortex would also play a role in planning strategies. The cerebellum, more involved in ballistic control of the movement and prediction of motor consequences, is not considered here.

The hippocampal part of the model focuses on generating attractors that correspond to transitions between learned proprioceptive states. When the robotic arm moves, its proprioception (joint angles, gripper IR and force sensor values) is categorized into distinct states. Each state is associated to a prototypical joint/gripper configuration. The categorization of the states is based on a recruitment of new states depending on a vigilance threshold. This algorithm takes inspiration from the Adaptive Resonance Theory [13]. The states are encoded on the weights of the input links of recruited neurons. The computation of the activity (1) and learning (2) for those neurons is given by the following equations:

$$A_i = 1 - \frac{1}{N} \sum_j |w_{ij} - e_j| \quad (1)$$

$$\Delta w_{kj} = e_j \cdot H\left(\sum_j H(\lambda - |\Theta_i - \theta_i|)\right) \quad (2)$$

where N is the dimension of the input layer, A_i is the activity of the i^{th} neuron of the output layer and w_{kj} is the weight of the connection between an output neuron k that has not been recruited yet and the j^{th} neuron of the input layer. H is the Heavyside function. The activity of a neuron is similar to a measure of similarity

with the encoded proprioceptive configuration. The most activated neuron is considered as the current state of the system. The recruitment of a new neuron depends on the difference between each current joint value θ_i and the corresponding joint value Θ_i of the best recognized configuration. If this difference is below a given vigilance threshold λ , it triggers the recruitment of a new neuron i.e. the encoding of a new state. The choice of the vigilance threshold λ determines the granularity of the states in the proprioceptive space.

The possible transitions between states are learned when the arm moves from one state to the next. This is done in an associative memory that receives information about the newly entered state and the last state. The transition activity is then transmitted to the cognitive map where a graph of the different transitions is created. The cognitive map encodes information about all possible paths by linking subsequent transitions.

In the original navigation architecture [12], the robot could find resources in its environment. To each type of resource corresponds a drive that represents the motivation of the robot to look for that kind of reward. From an animal's perspective the reward could be food and the drive hunger, for instance. Upon detection of a resource the current transition in the cognitive map is associated to the relevant drive. Here the robot possesses two different contextual goals associated to two types of objects (red or green). The focal vision of the robotic system categorizes the scene using color detection. It can differentiate between two types of objects: a red can and a green can. When a red object is presented to the visual system, the corresponding goal will be activated and stay active until another object is presented. If the new object is green, the first goal will be inhibited and the second goal will stay active. The visual system can also recognize when a large amount of white is present in the visual field. This is a way to use visual interaction to give rewards to the system and use reinforcement learning to associate goals with transitions. When experimenters present a white object, they reward the current behavior of the robot and associate it with the active contextual goal. The learning of the association between goals and transitions is given by the following equation:

$$\frac{dW_{ij}}{dt} = r(t) \cdot x_j^{drive}(t) \cdot (\alpha \cdot (1 - W_{ij}(t)) \cdot x_j^{trans}(t) - \gamma) \quad (3)$$

where α is a learning rate, γ a decay rate and $r(t)$ is the reward signal (0 or 1).

In this experiment we use a simple visual categorization system. We assume that the categorization of objects into goals and reward has already been learned. However, more natural and plausible approaches could be taken in the future. The reward signal can be linked to social cues. Zukowgoldring and Arbib [4] stress the importance of some key words like "look" or "now, you do it" combined with gestures that are used by caregivers to raise the attention of their infants. Also, a robot head that can display and recognize facial expressions could be added to the robotic setup of this experiment to make the interaction between human and robots easier. With such as

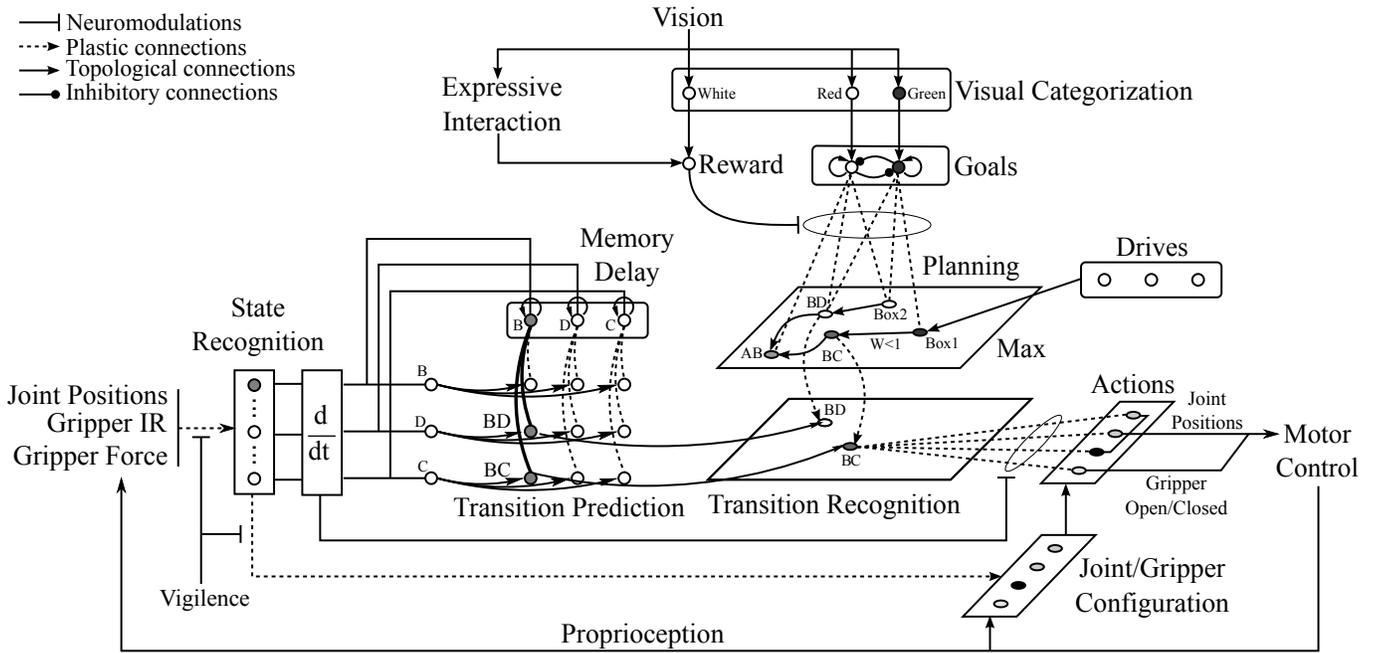


Fig. 1. General architecture for state categorization, transition learning, goal categorization, planning and action selection with the robotic arm with emphasis on the synaptic learning used in this paper.

system, the human could directly give reinforcing or punishing signals through interaction, by displaying positive or negative emotions. This kind of interaction has already successfully been used for the social referencing of objects with good performances [14]. The facial expression and direction of gaze of the teacher can also provide the necessary cues for guiding the attention of the visual system to interesting objects [15] and help categorize them. The robot would then be able to learn from interaction with a human partner which objects are interesting and which different types of object should be categorized into different goals.

When a goal is active, its activity will be propagated in the map with weights $w < 1$. As a result, transition neurons in the cognitive map have an activity proportional to their distance (in terms of number of transitions) to the goal. Since a max operator on the input is used in the computation of the activity of each neuron, the map gives the shortest path to the strongest or closest goal. In any given state, the hippocampal system predicts all possible transitions from the current position. The activity in the cognitive map is used to bias the selection of a particular transition. The transition that has the highest potential, ie. that corresponds to the shortest path to get the reward, is selected. During the learning process, the system creates proprioceptive states and memorizes the joint/gripper configuration associated to each state. When a transition between states is performed and learned, the position associated to the second state is associated to the transition. As a result, when the competition between transition occurs and a particular transition is selected, the motor configuration of the target state is given as the target position to the motor control system.

The motor control architecture (Fig. 2) receives the action from the cognitive map. If the arm is being manipulated by the human (by holding the wrist) this action is inhibited. Otherwise the action is transmitted to a force controller which computes the motor command of the robot by comparing the current and target configurations. When the arm is not being manipulated and the cognitive map transmits no action, a memory stores the value of the last action (or proprioceptive configuration if the arm was manipulated) to maintain the arm at that position. While the motor system receives no commands, a stress level increases until it finally reaches a sufficient value to inhibit the memory. A low-priority reactive behavior will then generate an attractor to a given rest position.

III. INTERACTIVE LEARNING OF A PICK-AND-PLACE TASK

The presented model was implemented with a 6DOF Katana robotic arm. At first, the robot has no knowledge about its environment. The task is a pick-and-place task (see Fig. 3 for the robotic setup and Fig. 4 for the different learning and

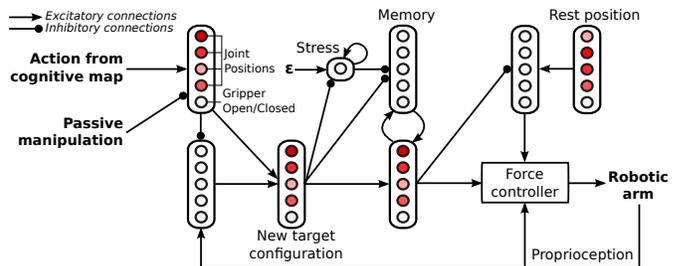


Fig. 2. Detailed architecture for motor control.



Fig. 3. Pick-and-place task learned and adapted through passive manipulation. Example of demonstration for the red-can-in-box-2 condition and reproduction by the robot. The camera looks at the pick-up place to recognize the type of object by color detection (red or green).

reproduction phases). An object (a colored can) is present in the reaching space of the robot at a given location (i.e. the pick-up place). The robot must grasp the can and drop it into one of two boxes, also at static positions in reaching space. A monocular camera is used as the visual system of the robot and is focused on the pick-up place. We assume that the visual attention of the robot is always directed toward that place.

At the beginning, the reactive behavior maintains the arm in rest position. The human teacher demonstrates how to grasp an object through passive manipulation of the arm. The first object is a red can. During the movement, proprioceptive states are recruited to encode the trajectory according to the process described in sec. II. The succession of states leads to transitions and the cognitive map being learned. After the arm has been manipulated by the human and brought to the object, the infrared sensors on the gripper detect the presence of an object in its range, triggering a grasping reflex. Force sensors on the gripper detect when an object is hold and determine the state of the gripper as opened or closed on an object. The arm is then manipulated again and brought above one of the drop boxes, learning the trajectory between the pick-up place and the drop box as shown by the human. A button located on the gripper commands the opening of the gripper. The button is used to simulate the interacting human forcing the opening of the gripper of the robot (which cannot be done due to mechanical limitations).

Once the object has been dropped into the box, a positive reinforcement is given by simply presenting a white sheet of paper in front of the camera. So when the robot releases the can, the goal associated to the dropped can (red or green) is linked to the current transition (i.e. changing the state of the gripper from closed to open when above this box). The robot then waits until the teacher demonstrates how to move from the box to its initial position or until the stress level activates the movement to the rest position. Once the loop is connected, the robot can plan the sequence of transitions to move to the position where it previously grasped the object. The robot starts reproducing the task without the help of a

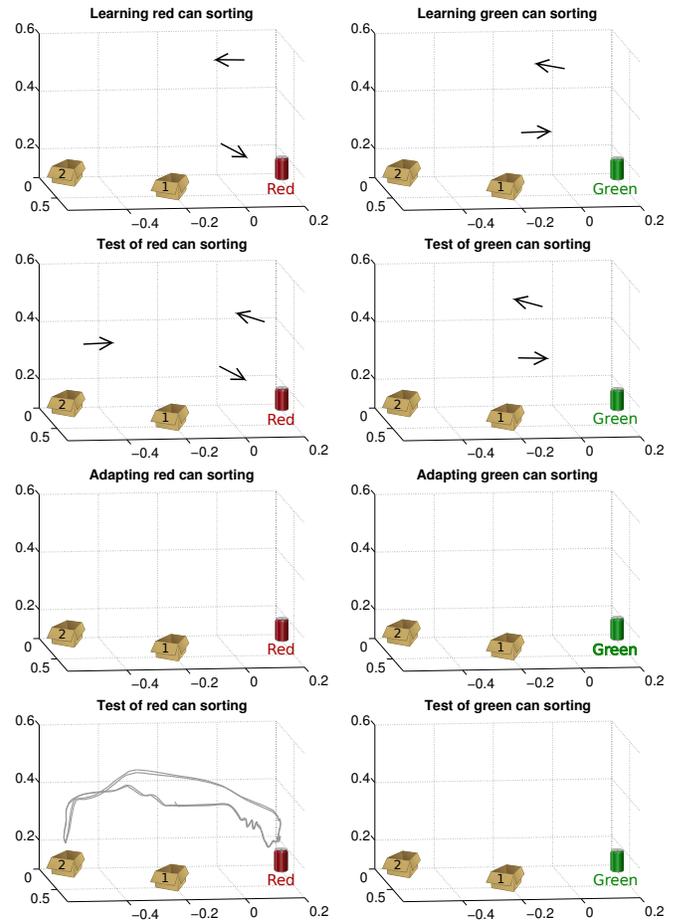


Fig. 4. Trajectory of the arm end effector in the 3D Cartesian space. The experiment can be explained as different phases, each one corresponding to a row. The first phase is the learning of the pick-and-place task for the red can and the green can. The black thick line corresponds to the trajectory during passive manipulation of the arm. The second phase is the validation of the reproduction of the task after learning. The third phase is the adaptation of the learning to new conditions: the reversal of the boxes for red and green can. The movement of the arm is corrected using passive manipulation to show the expected gestures. The learning of the new conditions needs several reinforced corrections to be learned. The last row shows the adapted reproduction after relearning.

human. Since one of the goals (depending on the color of the last seen object) is always active, the robot will constantly try to reach that goal and thereof to drop objects at the learned places (i.e. into on of the boxes).

From its starting position, the cognitive map gives the shortest path in terms of states in order to reach the motivated transition (dropping the object into the correct box). The first step is to go to the pick-up place. If an object is detected, it is grasped and the arm proceeds to the next phase, otherwise the robot waits for an object to be presented. Here we present the red can again and let the robot reproduce what it has just been taught. The arm reaches the position above the box where it opens its gripper and goes back to the pick-up location where a new can has been placed. The action is rewarded by the human for another two reproductions, reinforcing this behavior. When the behavior has been sufficiently reinforced for the red can,

the same learning process is used for the green can. The human first demonstrates how to drop the object in the second box and then lets the robot do the task, reinforcing this behavior for the next 2 reproductions. The robot is then given a series of red and green cans to verify that it correctly learned to place them in different boxes.

IV. CONTINUOUS LEARNING AND ADAPTATION TO NEW CONDITIONS

The robot successfully learned to reproduce the task. The system is still learning and can be adapted thanks to new interactions. For that we place a red can which the robot picks, but when it moves to drop it in the first box, the human teacher grabs the arm and moves it to the second box. When the object has been dropped into the new box, the behavior is rewarded. Due to the learning equation (3) the action of dropping the object in the new box is associated to the goal for red objects and the association between this goal and the old box is decaying a little. The stability of a behavior depends on the number of given rewards. The initial behavior was reinforced three times. After the rewarding of the first demonstration of the new behavior, the transitions leading to the new box are still less activated in the cognitive map than the transitions associated to the initial box. With another demonstration and reward, the goal/transition association begins to take precedence with a short lead. The robot is able to autonomously reproduce the correct behavior. A last reward when it has managed to correctly perform the task increases the lead of the new behavior (Fig. 5). After this learning the robotic arm places both types of objects in the second box. Finally, the procedure is repeated for the green objects and the robot is taught to place them in the first box. As a consequence, the

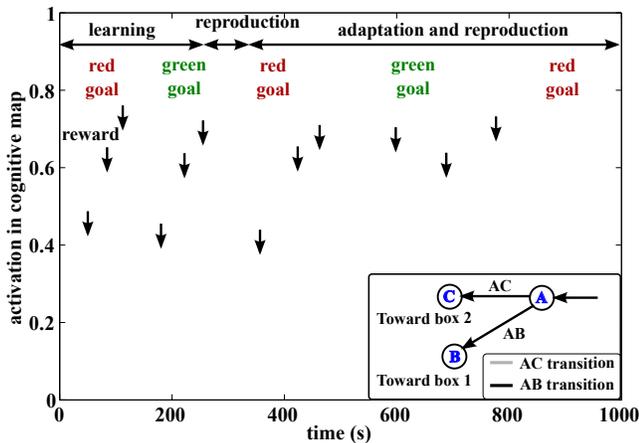


Fig. 5. The states A, B and C display the bifurcation of the paths leading to the two different boxes (See Fig. 6 for the 3D positions of these states). The graph represents the activity of the two transitions AB and AC in the cognitive map. With each reward for dropping in a box, the potentials of the transitions leading to this box increase while others tends to decrease. When the red and green goals are first presented, there is no activity in the cognitive map because no transition has been associated to these goals yet. In the second half of the experiment, the destination boxes are switched for both goals.

sorting of the objects in the two boxes is reversed compared to what was initially learned. The system benefits from the generalization properties of the planning architecture. During the adaptation to changing conditions in the task, only the part of the movement to the box was shown. After dropping the can, the robot was able to recognize its configuration and determine the path back to the pick-up place using what was previously learned for other conditions.

Figure 6 shows the position of the learned states in the workspace and the transitions learned between states at the end of the task. Interestingly, more states were recruited near the pick-up place than for the movement between the boxes. The 3D Cartesian space representation is not the space in which the states are encoded. Close positions in the Cartesian space can be distant in the multi-dimensional joint space of the arm. Besides, during the demonstration of the grasping, the different proximo-distal articulations were manipulated to reach the object with accuracy and with the correct orientation for the gripper. During the demonstration of the simplified left/right and up/down movements to the boxes, only the proximal joints were manipulated as there were no particular movements to be shown for the other joints. As a result, the recruitment process depending on each joint variation from the recognized prototypical joint/gripper configuration (sec. II) explains the different densities of states.

Figure 7 shows the most active state of the system when the end effector is at a given 3D position during reproduction. The color of the point on the trajectory corresponds to the color of the most active learned state. A state can be recognized before the arm end effector is at the learned position for the states. The system consequently anticipates the next move without finishing the current command to reach the center of the states, which leads to small shortcuts being taken in the reproduced trajectories as opposed to what was learned. The effect is more

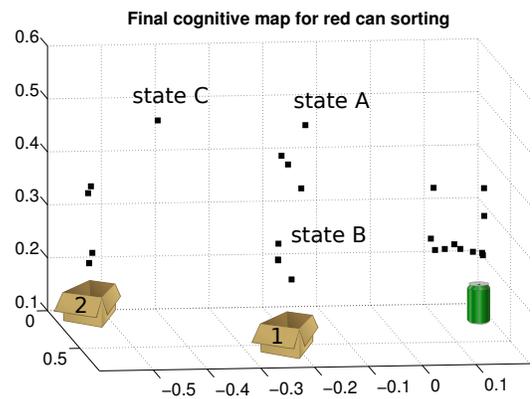


Fig. 6. Propagation of goal activity in the cognitive map for the red can goal at the end of the experiment. The learned states (squares) and the possible transitions (lines) are represented by the positions of the arm end effector in the 3D Cartesian space. The transitions are not symmetrical. The plain part of the line shows the starting state and the dashed part shows the final state of the transition. The darker and thicker lines corresponds to the transitions that have higher potentials in the cognitive map and thus are preferentially chosen.

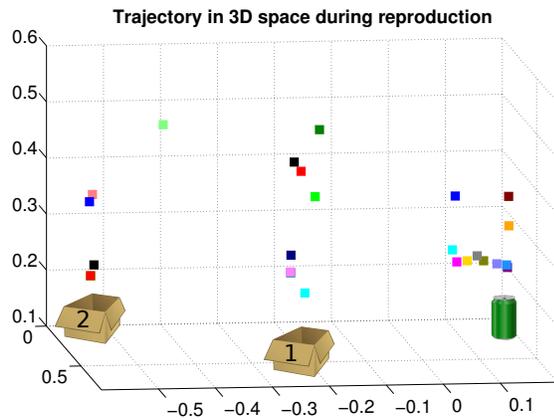


Fig. 7. Trajectory of the end effector in the 3D Cartesian space. The learned states are represented by colored squares. The color of the trajectory corresponds to the recognized states at the moment.

important when there are few states like in the middle of the trajectory from the pick-up place to the boxes. Figure 6 gives an example of the propagation of the activity for the different transitions in the cognitive map when one of the goals is active.

V. CONCLUSION

In this paper, we presented a bio-inspired model that enables a robot to build a representation of a task demonstrated through passive manipulation. This representation is based on proprioceptive states that are recruited and constitute the motor primitives of the system. It can be used by the system to plan its actions depending on visually categorized goals. Our system can continuously develop its model of the task through several interactions with humans. It was tested with a real robot on a changing pick-and-place task.

In this work, we described only one cognitive map built through interaction rather than exploration of the state space. Planning with large DOF systems can quickly lead to dimensionality problems that can be addressed by the use of hierarchical maps [16]. We consider the integration of different layers in our cognitive map model. One cognitive map would still be mostly acquired by interacting with humans, whereas another could be completed by exploring the environment (by babbling). This level of representation would be more extensive and only called upon in cases when the robot cannot infer what to do from what it was interactively taught. Future work will also be dedicated to the integration of different control systems for the imitation of observed actions and the imitation of actions experimented through “embodying”. This raises a major action selection issue as the different systems might give different motor commands in parallel.

Finally, an interesting addition to the system would be for the robot to indicate to the human that it does not know what to do. Several works [17] [18] have stressed the importance of the behavior of the robot in the way the interacting human will teach it. The stress value of the motor control system only increases when the planning system cannot propose any action, it is thus a way to detect that the current knowledge is

too limited to perform the task. The signal is used to trigger a return to the rest position but could also activate a visual cue indicating to the human the need for assistance. Signaling to the human that the robot needs help, e.g. by directing its gaze toward the caregiver, would lead to another interaction with the aim of addressing the current situation. If the human accepts to help the robot and demonstrates a solution, new information can be integrated in the representation of the task.

ACKNOWLEDGMENT

This work was supported by the INTERACT (referenced ANR-09-CORD-014) and NEUROBOT french ANR projects, the CNRS and DGA.

REFERENCES

- [1] P. Bakker and Y. Kuniyoshi, “Robot See, Robot Do : An Overview of Robot Imitation,” in *AISB96 Workshop on Learning in Robots and Animals*, 1996, pp. 3–11.
- [2] P. Andry, P. Gaussier, J. Nadel, and B. Hirsbrunner, “Learning Invariant Sensorimotor Behaviors: A Developmental Approach to Imitation Mechanisms,” *Adaptive Behavior*, vol. 12, no. 2, pp. 117–140, 2004.
- [3] A. de Rengerve, S. Boucenna, P. Andry, and P. Gaussier, “Emergent imitative behavior on a robotic arm based on visuo-motor associative memories,” in *Proc. IEEE/RSJ Int Intelligent Robots and Systems (IROS) Conf.*, 2010, pp. 1754–1759.
- [4] P. Zukowgoldring and M. Arbib, “Affordances, effectivities, and assisted imitation: Caregivers and the directing of attention,” *Neurocomputing*, vol. 70, no. 13-15, pp. 2181–2193, 2007.
- [5] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, “Survey: Robot Programming by Demonstration,” in *Handbook of Robotics*. MIT Press, 2008, vol. chapter 59.
- [6] A. J. Ijspeert, J. Nakanishi, and S. Schaal, “Learning attractor landscapes for learning motor primitives,” in *advances in neural information processing systems 15*. cambridge, ma: mit press, 2003, pp. 1547–1554.
- [7] S. Calinon, F. Guenter, and A. Billard, “On Learning, Representing and Generalizing a Task in a Humanoid Robot,” *IEEE transactions on systems, man and cybernetics*, vol. 37, no. 2, pp. 286–298, 2007.
- [8] R. Dillmann, O. Rogalla, M. Ehrenmann, R. Zöllner, and M. Bordegoni, “Learning Robot Behaviour and Skills Based on Human Demonstration and Advice: The Machine Learning Paradigm,” 1999.
- [9] S. Ekvall and D. Kragic, “Robot learning from demonstration: a task-level planning approach,” *International Journal of Advanced Robotic Systems*, 2008.
- [10] M. Schneider, R. Cubek, T. Fromm, and W. Ertel, “Combining Gaussian Processes and Conventional Path Planning in a Learning from Demonstration Framework,” in *Proceedings of the Eurobot Conference 2010*, Rapperswil-Jona (CH), 2010.
- [11] J. P. Banquet, P. Gaussier, J. C. Dreher, C. Joulain, A. Revel, and W. Gunther, “Space-time, order and hierarchy in fronto-hippocampal system: A neural basis of personality,” in *Cognitive Science Perspectives on Personality and Emotion*. Elsevier Science BV, 1997, pp. 123–189.
- [12] N. Cuperlier, M. Quoy, and P. Gaussier, “Neurobiologically inspired mobile robot navigation and planning,” *Front Neurobotics*, vol. 1, p. 3, 2007.
- [13] G. A. Carpenter and S. Grossberg, “Adaptive Resonance Theory,” 2002.
- [14] S. Boucenna, P. Gaussier, L. Hafemeister, and K. Bard, “Autonomous development of social referencing skills,” in *From Animals to Animals 11*, 2010, vol. 6226, pp. 628–638.
- [15] S. Boucenna, P. Gaussier, and L. Hafemeister, “Development of joint attention and social referencing,” to be published.
- [16] H. Liu, Y. Li, H. Wen, J. Xia, and T. Chu, “Hierarchical roadmap based rapid path planning for high-dof mobile manipulators in complex environments,” in *Proc. IEEE Int Robotics and Biomimetics (ROBIO) Conf.*, 2009, pp. 189–195.
- [17] Y. Nagai, A. Nakatani, and M. Asada, “How a robot’s attention shapes the way people teach,” in *Proceedings of the 10th International Conference on Epigenetic Robotics*, 2010, pp. 81–88.
- [18] A. L. Thomaz and C. Breazeal, “Teachable robots: Understanding human teaching behavior to build more effective robot learners,” *Artif. Intell.*, vol. 172, pp. 716–737, 2008.