

# Towards Hierarchical Curiosity-Driven Exploration of Sensorimotor Models

Sébastien Forestier  
ENS Rennes, Bruz  
INRIA/ENSTA-ParisTech, Bordeaux  
Email: sebastien.forestier@inria.fr

Pierre-Yves Oudeyer  
INRIA/ENSTA-ParisTech, Bordeaux  
Email: pierre-yves.oudeyer@inria.fr

## I. INTRODUCTION

Curiosity-driven exploration mechanisms have been proposed to allow robots to actively explore high dimensional sensorimotor spaces in an open-ended manner [1], [2]. In such setups, competence-based intrinsic motivations show better results than knowledge-based exploration mechanisms which only monitor the learner's prediction performance [2], [3]. With competence-based intrinsic motivations, the learner explores its sensor space with a bias toward regions which are predicted to yield a high competence progress. Also, throughout its life, a developmental robot has to incrementally explore skills that add up to the hierarchy of previously learned skills, with a constraint being the cost of experimentation. Thus, a hierarchical exploration architecture could allow to reuse the sensorimotor models previously explored and to combine them to explore more efficiently new complex sensorimotor models. Here, we rely more specifically on the R-IAC and SAGG-RIAC series of architectures [3]. These architectures allow the learning of a single mapping between a motor and a sensor space with a competence-based intrinsic motivation. We describe some ways to extend these architectures with different tasks spaces that can be explored in a hierarchical manner, and mechanisms to handle this hierarchy of sensorimotor models that all need to be explored with an adequate amount of trials. We also describe an interactive task to evaluate the hierarchical learning mechanisms, where a robot has to explore its motor space in order to push an object to different locations. The robot can first explore how to make movements with its hand and then reuse this skill to explore the task of pushing an object.

## II. METHODS

### A. Experimental setup

The experimental setup is composed of a robotic arm with 4 degrees of freedom from the Poppy robot [4] simulated with the V-REP simulator based on the Bullet physics engine. We use Dynamical Movement Primitive [5] to control the arm's movement as this framework allows the production of a diversity of arm's trajectories with few parameters. Each arm's DOF is controlled by a DMP with a starting and a goal position equal to the rest position of the motor. Each DMP is parameterized by one weight on each of 5 basis functions whose centers are distributed homogeneously throughout the movement of duration 4s. Each DMP outputs a series of angle positions that represents a sampling of the trajectory of one motor during the movement. A block is placed near the robot's hand and can be moved in two dimensions in different complex ways, e.g. with the hand pushing on the top of the block or on a side, in one or two steps.

We use and extend the Explauto library [6] to study autonomous exploration in developmental robotics. In the Explauto framework, a sensorimotor model mapping a motor space to a goal space is learned together with an interest function over the goal space. The interest function evaluates how much interesting exploring a given part of the goal space is. The robot selects a point in its goal space according to the interest function and infers DMP parameters allowing to get close to this point based on the past sensorimotor experience. It also adds some exploration noise to discover new motor configurations near the ones already known. Then it executes the motor trajectory corresponding to the parameters, and observes the consequences in the goal space.

Let  $M$  be the  $20D$  space of the motor parameters. Let  $S_h$  be the  $9D$  space representing the  $3D$  trajectory of the hand with each dimension projected on a DMP with 3 basis functions.  $S_o$  is the  $2D$  space representing the position of the object at the end of the simulation.

As we are interested in how an architecture has explored the different locations where the object can be pushed to, we define an exploration measure by discretizing  $S_o$  as a grid and counting the number of cells reached. However, the simulated environment is quite stochastic, so that some motor trajectories entail a high variance on the end position of the block and so are less robust for the robot. We thus define also a measure of competent exploration, which counts the number of explored cells where the agent manages to put the object within a threshold distance of the cell ( $2cm$ ) when asked to.

### B. Exploration architectures

We define two different hierarchical architectures to allow exploring the sensorimotor hierarchy.

**Control:** We compare our architectures to the control one where the robot learns directly a mapping between  $M$  and  $S_o$ , with a competence-based intrinsic motivation (GOAL-BABBLING) or a fully random motor babbling (MOTOR-BABBLING).

**Simplest First:** The agent explores for the first half of its trials the mapping between  $M$  and  $S_h$  and for the second half a mapping between  $S_h$  and  $S_o$ . When exploring the second one, the agent chooses a hand movement  $s_h$  in order to reach an object position  $s_o$ . To try to produce  $s_h$ , it then uses the first mapping to select the best motor configuration  $m$  to obtain this movement  $s_h$ .

**Top-Down Guidance:** We let the highest model in the hierarchy draw goals in  $S_o$ , then infer an interesting hand’s movements in  $S_h$  to try in order to reach that goal. The lowest model in the hierarchy is now given a certain amount of iterations to try motor configurations to obtain this hand’s movement. It is searching solutions with a black-box optimization technique [7].

## III. RESULTS

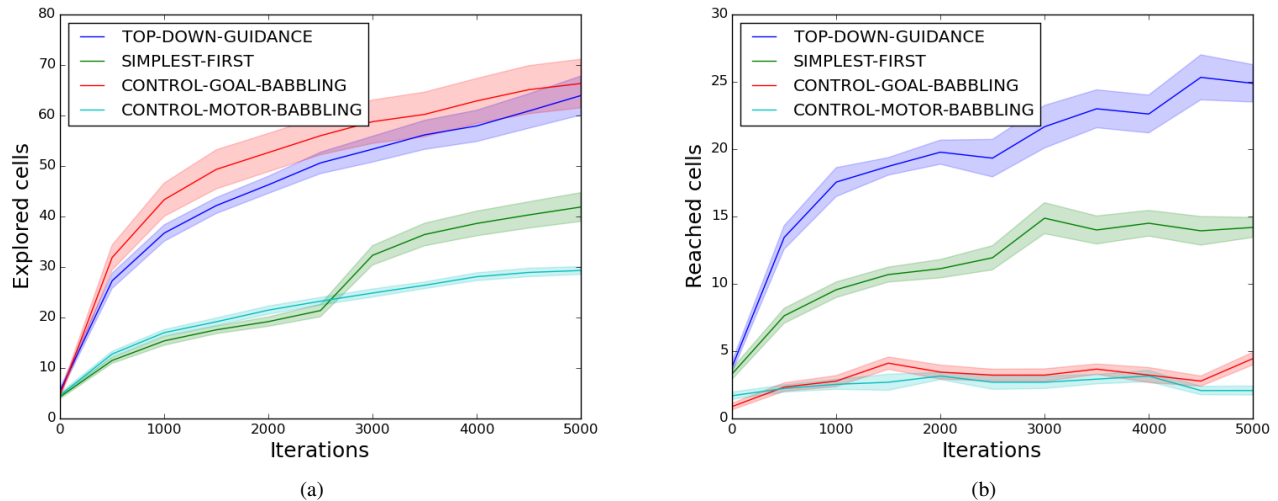


Fig. 1: Comparison of the different exploration architectures. (a) Exploration measure. (b) Competent exploration measure.

## IV. DISCUSSION

Results show comparable amounts of exploration for the top-down guidance and the control (goal babbling) conditions, whereas the competent exploration measure shows that the architectures learning an intermediate hand’s movement representation allow the agent to put again the object on much more diverse locations. The intuition for this result is that when exploring around given motor parameters, the control architecture will modify some motor’s angle trajectories and will produce more unstable interactions with the object than a hierarchical architecture that modify directly hand’s Cartesian trajectory. Learning intermediate representations should be more beneficial with complex learning hierarchies where more than one model reuses the learned representations.

## REFERENCES

- [1] J. Schmidhuber, “A possibility for Implementing curiosity and boredom in model-building neural controllers,” 1991.
- [2] V. G. Santucci, G. Baldassarre, and M. Mirolli, “Which is the best intrinsic motivation signal for learning multiple skills?” *Frontiers in Neurobotics*, vol. 7, 2013.
- [3] A. Baranes and P.-Y. Oudeyer, “Active learning of inverse models with intrinsically motivated goal exploration in robots,” *Robotics and Autonomous Systems*, vol. 61, no. 1, pp. 49–73, Jan. 2013.
- [4] M. Lapeyre, P. Rouanet, J. Grizou, S. Nguyen, F. Depaetere, A. Le Falher, and P.-Y. Oudeyer, “Poppy Project: Open-Source Fabrication of 3D Printed Humanoid Robot for Science, Education and Art,” in *Digital Intelligence 2014*, Nantes, France, Sep. 2014, p. 6.
- [5] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, “Dynamical movement primitives: learning attractor models for motor behaviors,” *Neural computation*, vol. 25, no. 2, pp. 328–373, 2013.
- [6] C. Moulin-Frier, P. Rouanet, P.-Y. Oudeyer, and others, “Explauto: an open-source Python library to study autonomous exploration in developmental robotics,” in *ICDL-Epirob-International Conference on Development and Learning, Epirob*, 2014.
- [7] N. Hansen, “The CMA evolution strategy: a comparing review,” in *Towards a new evolutionary computation*. Springer, 2006, pp. 75–102.