Towards hierarchical curiosity-driven exploration of sensorimotor models
Sébastien Forestier, Pierre-Yves Oudeyer

To cite this version:
Sébastien Forestier, Pierre-Yves Oudeyer. Towards hierarchical curiosity-driven exploration of sensorimotor models. 5th International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob), Aug 2015, Providence, RI USA, United States. Proceedings of the 5th International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob), 2015, 10.1109/DEVLRN.2015.7346146. hal-01250424

HAL Id: hal-01250424
https://hal.archives-ouvertes.fr/hal-01250424
Submitted on 4 Jan 2016

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.
Introduction

Curiosity-driven exploration mechanisms have been proposed to allow robots to actively explore high-dimensional sensorimotor spaces in an open-ended manner [1]. In such setups, intrinsic motivations based on competence progress show good results - the learner explores its sensory space with a bias toward regions which are predicted to yield a high competence progress. However, 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. We rely on the SAGG-RIAC series of architectures [2] and describe some ways to extend those architectures to the exploration of a hierarchy of sensorimotor skills. We developed a simulated robotic setup to evaluate the different architectures, where a robot has to push an object to different locations.

Curiosity-Driven Exploration

We use and extend the Explauto library [3] that aims at studying autonomous exploration. In the Explauto framework, a sensorimotor model is learned together with an interest model that guides future exploration.

- Sensorimotor model
This model stores the experimented motor commands and their associated sensory experience and builds a mapping between the motor space and the sensory space. We use the fast nearest neighbor algorithm to build this mapping but more powerful regression methods could be used instead.

- Interest model
The interest model estimates how interesting it is to explore given parts of the sensory space. We use the SAGG-RIAC architecture (Stiff Adaptable Goal Generation - Robust Intelligent Active Curiosity [2]) with an intrinsic motivation that pushes the agent to explore regions where the progress of the competence to reach self-generated goals in the sensory space is the higher.

- Exploration
The agent selects a goal in its sensory space according to the interest model and infers motor parameters to get closer to this point with the sensorimotor model. It adds some exploration noise to discover new motor configurations, executes the command and observes the sensory experience. Finally, it updates the sensorimotor model with the new association, and the interest model with the competence to reach the goal.

Hierarchical Exploration

To study hierarchical exploration, we design a setup where the agent learns two sensorimotor models, with the second one that reuse the first one. In our setup, the first sensorimotor model in the hierarchy is a relation between the 20 motor parameters of a robotic arm and the 9 parameters of the 3D trajectory of the robotic hand. The second sensorimotor model is a relation between the trajectory of the hand and the 2D position of a block at the end of the movement. Control architectures will have to learn directly a relation between the motor parameters and the position of the block.

Learning Architectures

- Motor Babbling Control
Exploration of a mapping between $M$ and $S_m$ with random motor actions.

- Goal Babbling Control
Exploration of a mapping between $M$ and $S_g$ with a competence-based intrinsic motivation (SAGG-RIAC).

- Simplicist First
Exploration of a mapping between $M$ and $S_g$ for the first half of the trials and of a mapping between $S_g$ and $S_m$ for the second half. Both explorations are made with SAGG-RIAC.

- Top-Down Guidance
Exploration of a mapping between $S_g$ and $S_m$. When asking to explore a hand’s movement $s_h$, this goal $s_h$ is used to guide the exploration of a mapping between $M$ and $S_g$, which is given certain amount of iterations to try motor configurations to reach that goal (with a block-heuristic optimization technique [4]).

Poppy Torso in the V-rep simulator

The robotic setup is the left arm of the Poppy robot [5], with 5 degrees of freedom, simulated in the V-rep simulator based on the Bullet physics engine.

- We use Dynamical Movement Primitive [6] 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 motor 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 7 basis functions whose centers are distributed homogeneously along the movement of duration $t_0$. $M$ is the 2D space of the motor parameters.

- $S_g$ is a 2D space representing the 3D trajectory of the hand. We also use the DMP framework to project each of the $X$ and $Y$ movements on a sensor DMP with 1 basis functions.

- A zero wide block is located near the robot’s hand and can be moved in different complex ways, e.g. with the hand pushing on the top of the block or on a side. $S_m$ is the 2D space representing the position of the block at the end of the simulation. The initial position of the block is $(X = 0.25 m, Y = 0.15 m)$. X axis points in front of the robot, Y axis is on its left.

Experiments

We ran different trials for each condition (number of trials, random walk, environment). Each trial is made of 1000 learning iterations. Each iteration takes about 3s on an Intel Core i7 3.5 GHz Xeon 2678 CPU. Exploration is evaluated every 500 iterations with 2 different measures. Statistical tests are performed after 5000 iterations to compare conditions.

Measures

- Exploration measure
This measures how diverse outcomes have been found by the agent in its goal space. It makes no assumptions about the interest of exploring different regions of the goal space from the viewpoint of the engineer. To compute the quantity of exploration we consider the goal space as an unbounded grid where we count the cells that have been reached during training.

- Low Robustness
If in our setup, the environment is highly stochastic in the sense that when the agent tries the same motor command different times, the variability in the generated arm movement and the collisions between the arm and the block leads to high variances in the end position of the block. Thus the agent might not succeed in reaching again a previously reached position, and in other words, a high exploration does not mean a high learning. We define the following measure to tackle this problem.

- Competent Exploration Measure
This measures how competent is the robot on the explored part of its goal space. We give the centers (black dots) of explored cells as goals to the agent, and we count the cells where the agent manages to get close to the goal (within 2cm).

Results

- Exploration measure
- Competent Exploration measure

Discussion

- Results show that exploration is better in the Top-Down Guidance condition than in the Simplicist First condition, with both measures. This implies that the guidance from the top-level model is useful to improve the exploration of the lower-level model on parts of the task space that are more interesting for the exploration of the higher one.

- Results show comparable exploration for the Top-Down Guidance and the Goal Babbling control condition, 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. An intuition for this result is that when exploring around motor parameters that lead to a movement of the block, the control architecture modifies directly joint trajectories leading to less accurate collisions with the block than hierarchical architectures that try to modify hand’s Cartesian trajectory. For this reason, hierarchical architectures would produce less diverse but more reproducible interactions than control architectures.

- In this setup, the intermediate hand’s movement representation is reused by only one higher-level model, but learning intermediate representations should be even more beneficial for exploration in complex learning hierarchies where more than one models are reusing the learned representations.

- The next step is to integrate social interaction with a human peer that will give demonstrations on the task space and the hierarchy of models to explore.

References


