Curiosity-Driven Development of Tool Use Precursors

Sébastien Forestier (sebastien.forestier@inria.fr)
INRIA Bordeaux Sud-Ouest
Bordeaux, France

Pierre-Yves Oudeyer (pierre-yves.oudeyer@inria.fr
INRIA Bordeaux Sud-Ouest
Bordeaux, France

Abstract

This is the abstract.

Keywords: curiosity-driven learning; tool use; goal babbling; overlapping waves;

Introduction

(Csikszentmihalyi, 1990) (Kidd, Piantadosi, & Aslin, 2012) (Baranes & Oudeyer, 2010) (Baranes & Oudeyer, 2013) (Moulin-Frier, Nguyen, & Oudeyer, 2014) (Gottlieb, Oudeyer, Lopes, & Baranes, 2013) (Moulin-Frier, Rouanet, Oudeyer, & others, 2014) (Oudeyer, Kaplan, & Hafner, 2007) (Ijspeert, Nakanishi, Hoffmann, Pastor, & Schaal, 2013) (Baranes & Oudeyer, 2009) (Schmidhuber, 1991) (Oudeyer & Smith, 2014) (Cangelosi et al., 2010) (Santucci, Baldassarre, & Mirolli, 2013) (Ugur, Nagai, Sahin, & Oztop, 2015) (Schmerling, Schillaci, & Hafner, 2015) (Oudeyer, 2007) (Mugan & Kuipers, 2009a) (Metzen & Kirchner, 2013) (Sutton et al., 2011) (Mugan & Kuipers, 2009b) (Vigorito & Barto, 2010) (Sutton, Precup, & Singh, 1999) () ()

Methods

Environment

We simulate a 2D robotic arm using tools to move an object into different boxes in the environment. In each trial, we execute a motor command given by the agent, we evaluate its consequences on the sensory dimensions and we give him this sensory feedback. Finally the environment is resetted to its intial state.

The next sections precisely describe the different items of the environment and their interactions. See Fig. 1 for an exemple of the state of the environment.

Robotic arm The 2D robotic arm has 3 joints plus a gripper located at the end-effector. Each joint can rotate from $-\pi$ *rad* to π *rad* around its initial position, mapped to a standard interval of $[-1,1]$. The length of the 3 parts of the arm are 0.5, 0.3 and 0.2 so the total length of the arm is 1 unit. The initial position of the arm is vertical with each joint at 0 *rad* and its base is fixed at position $[0,0]$. The gripper g has 2 possible positions: *open* ($g \geq 0$) and *closed* ($g < 0$) and its initial position is *open* (with $g = 0$). The robotic arm thus has 4 degrees of freedom represented by a vector in $[-1,1]^4$. A trajectory of the arm will be represented as a sequence of such vectors.

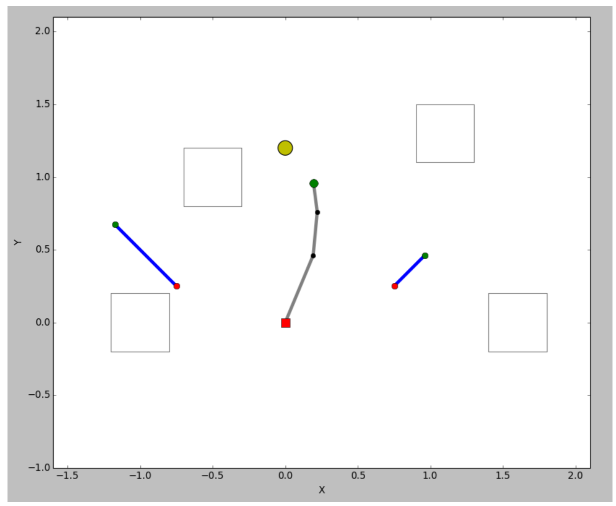

Figure 1: Play Environment

Objects and tools A yellow sphere can be moved into one of the 4 fixed squared boxes. The initial position of the sphere is $(0,1.2)$ and is thus unreachable directly with the gripper. One of two sticks can be grasped in order to reach the object. A small stick of length 0.3 is located on the right of the arm, with initial position $(0.75,0.25)$ and initial angle $\frac{\pi}{4}$ from the horizontal line. A long stick of length 0.6 is located on the left of the arm, with initial position $(-0.75,0.25)$ and initial angle $\frac{3\pi}{4}$ from the horizontal line as in Fig. 1. If the gripper closes near the end of one of the sticks (closer than 0.1), it is considered grasped and will follow the gripper's position and the angle of the arm's last part until the gripper opens. Similarly, if the other end of a stick reaches the sphere (within 0.1), the object will follow the end of the stick. Four boxes are fixed at positions $(-1,0)$, $(-0.5,1)$, $(1.1,1.3)$ and $(1.6,0)$ and have size 0.4, so that the two boxes on the left can be reached by the two sticks, and the two boxes on the right can only be reached by the long stick. At the end of the trial, the object is considered to be in one of the box if its center is in the box.

Motor control We use Dynamical Movement Primitive (Ijspeert et al., 2013) to control the arm's movement as this framework permits the production of a diversity of arm's trajectories with few parameters. Each of the 4 arm's degrees-of-freedom (DOF) is controlled by a DMP with a starting and a goal position equal to the rest position of the joint. Each DMP is parameterized by one weight on each of 3 basis functions whose centers are distributed homogeneously throughout the movement duration. The weights are bounded in the interval $[-200, 200]$ (mapped to the standard interval $[-1, 1]$) which allow each joint to fairly cover the interval $[-1, 1]$ during the movement. Each DMP outputs a series of 50 positions that represents a sampling of the trajectory of one joint during the movement. The arm's movement is thus parameterized by 12 weights which are represented by the motor space $M = [-1, 1]^{12}$

Sensory feedback At the end of the movement, the robot gets sensory feedback from the different items of the environment. It gets the trajectory of its hand and gripper, the trajectory of the end of the sticks, the end position of the object, and whether the object is in each box. The trajectory of the hand and of the end point of the sticks are represented by sequences of x and y positions at different time points: steps 12, 25, 37 during the movement of 50 steps (6D for the hand and for each stick). Similarly, the trajectory of the gripper is a sequence of 1 or -1 depending whether the gripper is open or closed (3D). For each box, the robot receives a 1 if the object is in the box at the end of the movement, and 0 otherwise (4D). The sensory information thus contains 9 values for the trajectory of the hand and gripper, 6 for the trajectory of the end of each stick, 2 for the end position of the object and 4 for the boxes. The total sensory space has 27 dimensions.

Model's architecture

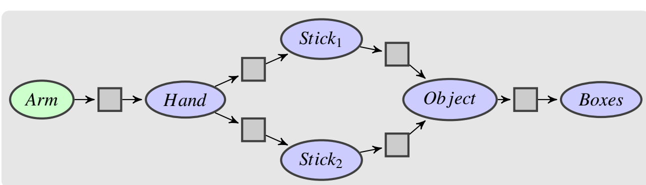

Figure 2: Hierarchy of sensorimotor models

Results

Discussion

CoGNiTiVe ScIeNcE

Figure 3: This is a figure.

Acknowledgments

Place acknowledgments (including funding information) in a section at the end of the paper.

References

Baranes, A., & Oudeyer, P.-Y. (2009). R-iac: Robust intrinsically motivated exploration and active learning. *Autonomous Mental Development, IEEE Transactions on*, *1*(3), 155–169.

Baranes, A., & Oudeyer, P.-Y. (2010). Intrinsically motivated goal exploration for active motor learning in robots: A case study. In *Intelligent robots and systems (iros), 2010 ieee/rsj international conference on* (pp. 1766–1773).

Baranes, A., & Oudeyer, P.-Y. (2013, January). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems*, *61*(1), 49–73. doi: 10.1016/j.robot.2012.05.008

Cangelosi, A., Metta, G., Sagerer, G., Nolfi, S., Nehaniv, C., Fischer, K., ... others (2010). Integration of action and language knowledge: A roadmap for developmental robotics. *Autonomous Mental Development, IEEE Transactions on*, *2*(3), 167–195.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row. Retrieved from `http://books.google.fr/books?id=V9KrQgAACAAJ`

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013, November). Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. doi: 10.1016/j.tics.2013.09.001

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., & Schaal, S. (2013). Dynamical movement primitives: learning attractor models for motor behaviors. *Neural computation*, *25*(2), 328–373.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One*, *7*(5), e36399.

Metzen, J. H., & Kirchner, F. (2013). Incremental learning of skill collections based on intrinsic motivation. *Frontiers in Neurorobotics*, *7*. doi: 10.3389/fnbot.2013.00011

Moulin-Frier, C., Nguyen, S. M., & Oudeyer, P.-Y. (2014). Self-organization of early vocal development in infants and machines: the role of intrinsic motivation. *Frontiers in Psychology*, *4*.

Moulin-Frier, C., Rouanet, P., Oudeyer, P.-Y., & others. (2014). Explauto: an open-source Python library to study autonomous exploration in developmental robotics. In *ICDL-Epirob-International Conference on Development and Learning, Epirob.*

Mugan, J., & Kuipers, B. (2009a). Autonomously Learning an Action Hierarchy Using a Learned Qualitative State Representation. In *IJCAI* (pp. 1175–1180).

Mugan, J., & Kuipers, B. (2009b). Autonomously learning an action hierarchy using a learned qualitative state representation.

Oudeyer, P.-Y. (2007). What is intrinsic motivation? A typology of computational approaches. *Frontiers in Neurorobotics*, *1*. doi: 10.3389/neuro.12.006.2007

Oudeyer, P.-Y., Kaplan, F., & Hafner, V. V. (2007, April). Intrinsic Motivation Systems for Autonomous Mental Development. *IEEE Transactions on Evolutionary Computation*, *11*(2), 265–286.

Oudeyer, P.-Y., & Smith, L. (2014). How evolution may work through curiosity-driven developmental process.

Santucci, V. G., Baldassarre, G., & Mirolli, M. (2013). Which is the best intrinsic motivation signal for learning multiple skills? *Frontiers in Neurorobotics*, *7*. doi: 10.3389/fnbot.2013.00022

Schmerling, M., Schillaci, G., & Hafner, V. V. (2015). Goal-directed learning of hand-eye coordination in a humanoid robot. In *5th joint ieee international conferences on development and learning and on epigenetic robotics (icdl-epirob).*

Schmidhuber, J. (1991). A possibility for lmplementing curiosity and boredom in model-building neural controllers.

Sutton, R. S., Modayil, J., Delp, M., Degris, T., Pilarski, P. M., White, A., & Precup, D. (2011). Horde: A scalable real-time architecture for learning knowledge from unsupervised sensorimotor interaction. In *The 10th international conference on autonomous agents and multiagent systems-volume 2* (pp. 761–768).

Sutton, R. S., Precup, D., & Singh, S. (1999). Between mdps and semi-mdps: A framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, *112*(1), 181–211.

Ugur, E., Nagai, Y., Sahin, E., & Oztop, E. (2015, June). Staged development of robot skills: Behavior formation, affordance learning and imitation with motionese. *Autonomous Mental Development, IEEE Transactions on*, *7*(2), 119-139. doi: 10.1109/TAMD.2015.2426192

Vigorito, C. M., & Barto, A. G. (2010). Intrinsically motivated hierarchical skill learning in structured environments. *Autonomous Mental Development, IEEE Transactions on*, *2*(2), 132–143.