Studying How Innate Motivations Can Drive Skill Acquisition in Cognitive Robots [†]

Alejandro Romero [1,*], Francisco Bellas [2], Jose A. Becerra [2] and Richard J. Duro [2]

[1] Integrated Group for Engineering Research, Universidade da Coruña, 15403 Ferrol, Spain
[2] CITIC Research Center, Universidade da Coruña, 15071 A Coruña, Spain
[*] Correspondence: alejandro.romero.montero@udc.es
[†] Presented at the 2nd XoveTIC Conference, A Coruña, Spain, 5–6 September 2019.

Abstract: In this paper, we address the problem of how to bootstrap a cognitive architecture to opportunistically start learning skills in domains where multiple skills can be learned at the same time. To this end, taking inspiration from a series of computational models of the use of motivations in infants, we propose an approach that leverages two types of cognitive motivations: exploratory and proficiency based, the latter modulated by the concept of interestingness as an implementation of attentional mechanisms. This approach is tested in an illustrative experiment with a real robot.

Keywords: Cognitive Developmental Robotics; open-ended learning; motivational system; skill learning

1. Introduction

With the aim of designing robots with a higher degree of autonomy, the field of Cognitive Developmental Robotics (CDR) takes inspiration from models of cognitive human development. Robots are endowed with cognitive architectures which, starting from basic innate knowledge provided by the designer, are able to generate new knowledge, mainly models and skills, in a fully autonomous way throughout their "lives". Being able to learn in such an open-ended manner implies dealing with an unlimited sequence of a priori unknown tasks in unknown domains [1].

Consequently, the problem is not that of providing a robot with competences to perform particular tasks in known environments, but to provide the robot with mechanisms that allow it to figure out what tasks to carry out, and how, to achieve its objectives in the situations it faces. In other words, it needs to self-discover and self-select goals. It is important to emphasize here that a goal determines a task the robot must carry out (to reach the goal) and, consequently, a skill it must learn in order to be able to achieve it.

On the other hand, the robot also needs to determine how valuable any goal is (what is its utility) and, by extension, what may the expected utility of any point in state space be with regards to that goal. The mechanisms in charge of this are generally called motivational mechanisms or value systems. This work is framed within the problem of creating adequate motivational systems for autonomous robots, specifically, within the MDB cognitive architecture [2], to efficiently learn and purposefully behave in open-ended settings, and focusing on the initial stages of skill learning.

2. Unrewarded Skill Acquisition and Interestingness

At initial stages of interaction with an unknown world, the robot can only rely on what it has been innately endowed with by the designer, and it must use it to progressively acquire new skills that will allow it to become more proficient. Consequently, designing an appropriate set of innate drives is key to the adequate performance of the robot.

In the approach chosen within the motivational engine of the MDB [3], inspired by the observations of child cognitive development, we propose that two types of drives constitute the minimum set of cognitive drives required for this process. On the one hand, the robot needs to explore its state space in order to find utility. This exploration must be efficient and, consequently, some type of cognitive drive related to exploration must be included. In particular, in the experiments we present in the next section, we have made use of a drive related to novelty. However, to learn a skill, it is also necessary to train and become proficient at it. That is, the robot needs to be motivated to concentrate its interaction with the environment on cases that can lead to learning the skill. That is, to establish a virtual goal in that point and learn its utility model. We will call this a Proficiency based type of motivation. In particular, as skills are usually learned in order to be able to produce some effect on the environment, we will make use of an effectance based motivation in the experiments.

To induce training, we incorporate the concept of interestingness within the related Proficiency based motivation as a virtual utility value that can change in time as the robot becomes more proficient at achieving the corresponding goal. Thus, when an effect is produced by chance for the first time, the point in state space where that occurred becomes interesting (its interestingness level increases). This is reflected within the motivational engine as a virtual utility value when the goal is achieved and within the attention mechanism of the robot by increasing the saliency of the state-space point in the process of choosing where to go next. However, interestingness is also modulated by the proficiency in achieving the goal: the more proficient the robot is, the less interesting the virtual goal becomes. Once the robot is very proficient, the skill for achieving the goal will have been acquired and it can be sent to Long Term Memory (LTM) for storage and future recall.

3. Real Robot Experiment

The Baxter robot is placed in front of a white table with three different objects it can detect: a brown box, a red ball and a small plastic jar which lights up when it is grabbed. The robot can detect the distance to the objects by using their color and shape.

The execution of the experiment, illustrated in the images of Figure 1, can be described as follows: the robot started its operation without any explicit goal nor skill apart from the two innate motivations mentioned above. Consequently, it started moving its right arm guided by the novelty motivation. Eventually, this novelty seeking motivation leads it to hitting and pushing an object, in this case the ball (see Figure 1 (a)), thus generating a change in the perceptions of the robot that it will interpret as an effect of its actions on the environment. This increases the interestingness value of the point in state space where the change occurred and establishes it as a virtual goal to be achieved. As the robot becomes more proficient, the robot loses interest in moving the ball and goes back to seeking novelty. At this point the value function (VF) obtained for the push-ball skill, shown in Figure 2 (a), is stored in the LTM of the MDB for future use.
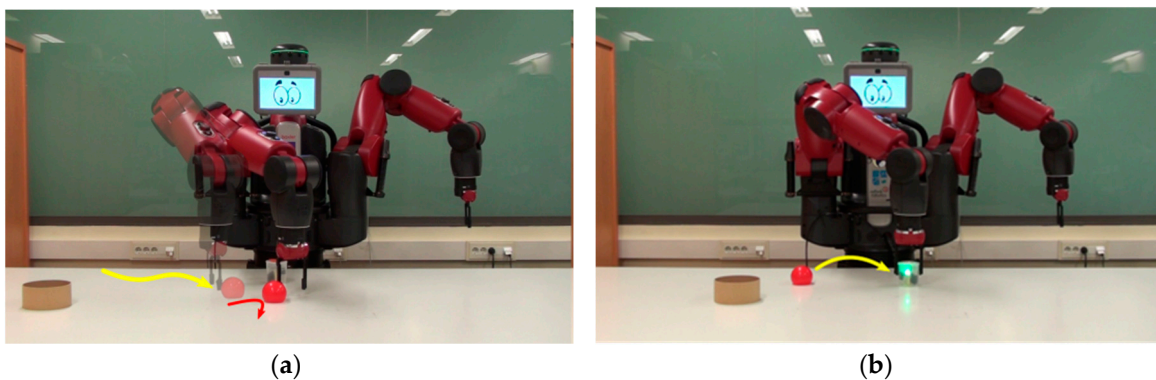
Figure 1. Experimental setup with the Baxter robot. (**a**) Pushing skill; (**b**) Grasping skill.

As the robot continues to explore, some object may end up between its gripper pads triggering the close gripper reflex action. This action really does not cause any effect in any of the objects except

for the jar. When it is the jar the one the gripper closes on, it lights up. This obviously is an effect and, as in the previous case, an interestingness value is assigned (see Figure 1b). Again, the proficiency based motivation starts guiding the robot response and a second VF learning process is launched. As the grasping skill associated to this VF improves, the interestingness value decreases until the corresponding VF (Figure 2b) has been correctly learnt and is stored in the LTM. The process continues with a new exploratory stage and, if pertinent, new activations of the effectance drive that will allow learning new skills.
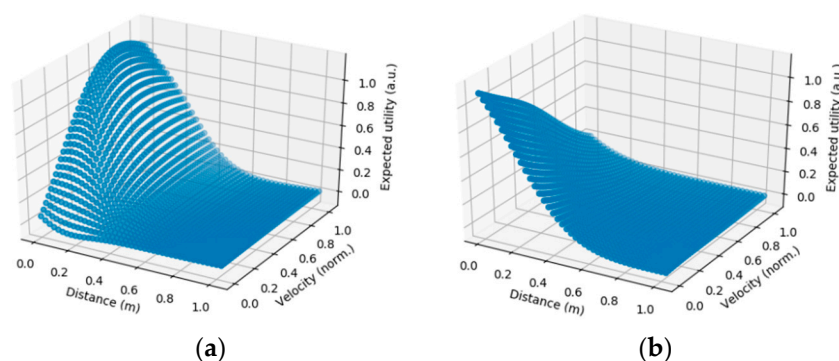
Figure 2. 3D representation of the skills learned in terms of distance and speed of the gripper. (**a**) VF associated to the push-object skill; (**b**) VF associated with the grabbing skill.

Author Contributions: Conceptualization, R.J.D. and A.R.; Methodology, A.R., F.B. and R.J.D.; Software, A.R. and J.A.B.; Validation, A.R., F.B. and R.J.D.; Writing—original draft preparation, A.R.; Writing—review and editing, A.R., F.B. and R.J.D; Visualization, A.R.; Supervision, R.J.D and F.B.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Doncieux, S.; Filliat, D.; Diaz-Rodriguez, N.; Hospedales, T.; Duro, R.; Coninx, A.; Roijers, D.; Girard, B.; Perrin, N.; Sigaud, O. Open-ended learning: a conceptual framework based on representational redescription. *Front. Neurorobot.* **2018**, *12*, 59.
2. Bellas, F.; Duro, R.J.; Faina, A.; Souto, D. Multilevel Darwinist Brain (MDB): Artificial Evolution in a Cognitive Architecture for Real Robots. *IEEE Trans. Auton. Ment. Dev.* **2010**, *4*, 340–354.
3. Romero, A., Prieto, A., Bellas, F., Duro, R.J. Simplifying the creation and management of utility models in continuous domains for cognitive robotics, *Neurocomputing* **2019**, *353*, 106–118.