Improving Extrinsically Motivated Developmental Robots through Intrinsic Motivations

Rodrigo Salgado, Abraham Prieto, Francisco Bellas and Richard J. Duro

Abstract—**This paper presents the initial results obtained in the application of the MotivEn motivational system in developmental robotics. The key feature of MotivEn is that it uses extrinsic motivations as the primary drive to guide the robot development, including intrinsic motivations to improve learning. We analyze here through an experiment with a simulated robot that this both types of motivations are required.**

I. Introduction

Cognitive architectures (CAs) in developmental robotics are based, as opposed to classical symbolic approaches, on cognitive processing theories where the key feature is the emergence of cognitive capabilities through embodied interaction with the real world in an incrementally more complex fashion [1]. Developmental robots are guided by intrinsic motivations, as can be seen in the most relevant CAs of the field, like the iCub CA [2], ERA [3], SASE [4] or HAMMER [5]. Following [6], intrinsic motivations drive the acquisition of knowledge and skills in a heterostatic fashion, that is, in the absence of a goal state that can be reached, whereas extrinsic motivations guide development towards a homeostatic goal state that can be satiated.

Thus, developmental robotics CAs use motivations like novelty, curiosity, knowledge acquisition or skill improvement, as an open-ended mechanism to obtain reliable cognitive capabilities [7]. A core idea behind developmental robotics is that, once consolidated, such emergent capabilities can be used by the robot in later developmental stages to fulfill extrinsic motivations, like those provided by a human user or other acquired during life [1]. In fact, the developmental robotics field arose as a new perspective for obtaining real autonomous robots that can operate life-long, escaping from the limited vision of classical robotics focused in reaching specific goals [2]. Thus, in a certain way, developmental robotics CAs have intentionally avoided the use of explicit extrinsic motivations.

Once the validity of the intrinsically motivated robots has been widely shown [7], the question that arises is how the acquired knowledge and skills can be managed by a CA to reach specific goal states if they emerge "freely". That is, the absence of extrinsic motivations to guide robot operation leads to a set of questions that must be faced: how can it be guaranteed that the intrinsically motivated robot is acquiring the cognitive capabilities it will require in the future? How can this emergent knowledge be reused to reach a goal state?

In this paper, we propose to consider a developmental robot as an open-ended system that is guided by extrinsic motivations, that is, it has to fulfill some goals in a homeostatic fashion. Moreover, the robot must have intrinsic motivations, as usual in developmental robotics, which are required to accomplish the robot development in the most reliable way. To analyze the response of this extrinsically guided developmental robot, we have developed a motivational engine, called MotivEn, and we have implemented some basic examples in simulated experiments as a part of the EU's H2020 DREAM project [8]. Here, we will briefly describe the main elements and operation of MotivEn (more details in [9]), and we will show the initial results obtained in a simulated experiment of autonomous goal acquisition when considering intrinsic motivations or not as a drive to improve the goal achievement.

II. Motivational Engine (MotivEn)

Given a perceptual state $\tau(t)$, MotivEn evaluates each future state $\tau(t+1)$, proposed by the action chooser of a general CA, based on three main motivational components:

1) *Blind intrinsic motivation (I_b):* it is active when there is no information on whether a goal exists or of how to get to it from the current (perceptual) region. It guides the robot behavior towards the discovery of unvisited states in the perceptual space and operates as an exploratory intrinsic process (similar to novelty [1]).

2) *Extrinsic motivation (E):* it tries to guide the behavior of the robot towards maximizing the reward, that is, towards what is typically called the goal of the scenario. The robot must create a value function (*VF*) as an internal representation of the utility function that associates an expected future utility to any point of the state space. Before MotivEn can start using that *VF*, the *CA* has to learn it. This process in MotivEn is based on the learning scheme used in the Multilevel Darwinist Brain cognitive architecture [10], which is an on-line neuroevolutionary approach where the *VF* is represented through an ANN.

A key aspect of MotivEn is the *certainty* of the *VF*, an area in the state space that represents its reliable coverage. To compute it, we have defined a certainty value $C(\tau)$ for each point of the *VF*, which corresponds to a measure of how many really evaluated trace points are near that state space point covered by the *VF*, modulated by the number of times it guided the robot towards achieving the goal [9]. The

All authors are with the Integrated Group for Engineering Research, University of A Coruña, Ferrol, Spain (phone: +34981337400, e-mail: {rodrigo.salgado, abprieto, fran, richard}@ udc.es).

This work has been partially funded by the EU's H2020 research and innovation programme under grant agreement No 640891 (DREAM) and by the Xunta de Galicia and European Regional Development Funds under grant GRC 2013-050.

Preprint submitted to 6th International Conference on Development and Learning and on Epigenetic Robotics. Received April 15, 2016.

combination of the expected reward and the certainty of a state produces the resulting extrinsic evaluation of a state which is what will finally guide the robot.

3) Certainty-based intrinsic motivation (I_{cb}): it is in charge of guiding the robot to improve the *VF* model. This is achieved by improving the sampling of the traces used to learn the *VF*. This motivation seeks to expand the sampled area by encouraging moving close to the boundaries of the certainty region of the value functions [9].

In summary, for each candidate state τ(t+1) proposed by the action chooser, MotivEn analyzes to what extent it falls into the certainty area of the *VF*. If the certainty value is beyond a threshold value (high reliability), the *extrinsic motivation* is used. If it is below the threshold (low reliability), the *certainty-based intrinsic motivation* is applied. Finally, it the state falls out of the VF certainty area, the *blind intrinsic motivation* is used to guide the robot.

III. Operation Example

In this example we show how the *certainty-based intrinsic motivation* improves the development of an extrinsically guided robot. To do it, we have created a simulated scenario (Fig. 1) that contains a blue box, a green button, a red ball and a virtual barrier (opaque to the sensors of the robot), which divides the arena into two parts. The robot is initially placed in the left part of the scenario. The red ball is placed on the right side of the scenario, and the robot cannot initially sense it. The robot is able to move around this environment and to reach the different objects. If it reaches the button, it is automatically pushed and the barrier disappears, so it can sense the ball. If it reaches the ball, it automatically picks it up. Finally, if it reaches the box when it is carrying the ball, it receives a reward (right plot). When this happens, the ball is returned to its original position, the virtual barrier is restored and the robot is placed in a random location of the left part.

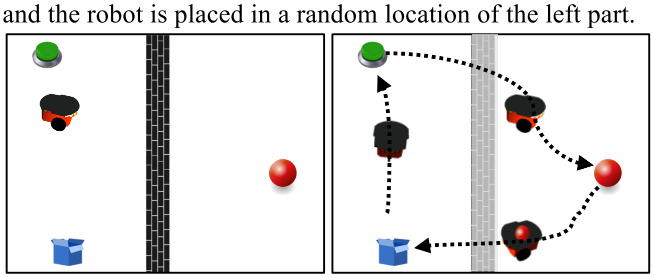

Fig. 1. Simulated scenario used in the experiment (left) and a typical execution trace followed by the robot when reaching the goal (right)

For this experiment, the response of the scenario to the actions of the robot is known a priori (the CA does not have to learn the world models), but the robot has no idea where the reward is and how to reach it. Consequently, it will have to discover it and then learn to get to it from anywhere in the environment, that is, learn the *VF*. To this end, three distance sensors were set up for the robot: to the green button (*g*), to the red ball (*r*) and to the blue box (*b*). As for the actions (*a*), the robot can change its orientation by an angle between -90º and 90º, and later move straight a fixed distance according it.

With this setup, we have executed ten runs of the MDB cognitive architecture using MotivEn to guide the action selection. The top plot of Fig. 2 displays the number of goals achieved (during the last 500 time steps) in a representative run of 20000 time steps. The red line corresponds to an execution where the *certainty-based intrinsic motivation* (I_{cb}) was disabled, while the blue line corresponds to an execution with the 3 motivational components of MotivEn. As it can be observed by the higher number of goal achievements, the presence of an intrinsic motivation that promotes the *VF* improvement (I_{cb}) is clearly more successful than using the extrinsic motivation directly. The bottom plot of Fig. 2 shows the time steps where the *blind intrinsic motivation* (I_b) was used in these two configurations. When I_{cb} was disabled (red line), I_b was continuously used because the *VF* certainty was low, leading to a more random operation. When enabling I_{cb} (blue line), the application of I_b was notably reduced, meaning that the *VF* was increasingly more reliable.

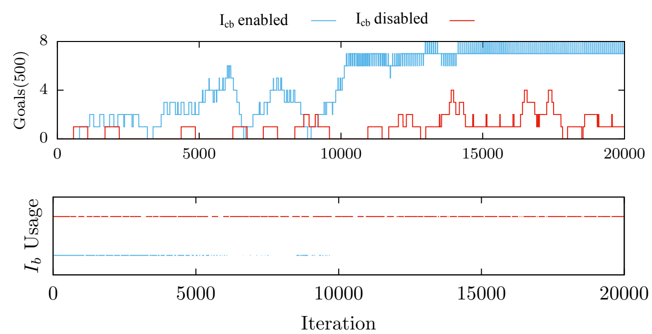

Fig. 2. Goals achieved by the robot (top) and use of *blind intrinsic motivation* (bottom) in a representative execution of the simulated experiment

IV. Conclusions

In this paper we have presented some initial results in the application of MotivEn, an integrated approach to the combination of intrinsic and extrinsic motivations in order to autonomously acquire knowledge in a developmental robot.

References

[1] A. Cangelosi, M. Schlesinger M., "Developmental Robotics: From Babies to Robots", *MIT Press*, 2015

[2] D. Vernon, C. von Hofsten, L. Fadiga, "A Roadmap for Cognitive Development in Humanoid Robots", vol 11, *Cognitive Systems Monographs*. Springer, 2011

[3] A.F. Morse, J. de Greeff, T. Belpeame, A. Cangelosi, "Epigenetic Robotics Architecture (ERA)", *IEEE Transactions on Autonomous Mental Development 2, 4*, pp. 325-339. 2010

[4] J. Weng, "On Developmental Mental Architectures". *Neurocomputing, vol. 70, no. 13-15*, pp. 2303-2323, 2007

[5] Y. Demiris, A. Meltzoff, "The Robot in the Crib: A Developmental Analysis of Imitation Skills in Infants and Robots", *Infant and child development 17(1):43-53*, 2008

[6] G. Baldassarre, "What are intrinsic motivations? A biological perspective," Proceedings *ICDL 2011*, IEEE Press, 2011, pp. 1-8

[7] G. Baldassarre, M. Mirolli, "Intrinsically Motivated Learning Systems: An Overview", *Springer*, 2013

[8] EU's H2020 DREAM project, "http://www.robotsthatdream.eu"

[9] R. Salgado, A. Prieto, P. Caamaño, F. Bellas, R. J. Duro, "Motivational Engine with Sub-goal Identification in Neuroevolution Based Cognitive Robotics", *LNCS 9648*, pp 659-670, 2016

[10] F. Bellas, R.J. Duro, A. Faina, D. Souto, "Multilevel Darwinist Brain (MDB): Artificial Evolution in a Cognitive Architecture for Real Robots". *IEEE Trans. on Auton. Mental Develop.* 2(4), 340-354, 2010

Preprint submitted to 6th International Conference on Development and Learning and on Epigenetic Robotics. Received April 15, 2016.