Proximo-Distal Competence Based Curiosity-Driven Exploration

Adrien Baranes, Pierre-Yves Oudeyer
INRIA Bordeaux - Sud-Ouest
351 Cours de la Liberation, 33405 Talence, France
{Adrien.Baranes, Pierre-Yves.Oudeyer}@inria.fr

Learning constraints and guiding mechanisms are involved since the very beginning of the infant development. Allowing a progressive and open-ended scaffolding of new skills, these mechanisms have been described as crucial, by psychologists and neuroscientists. Developmental heuristics presented here are directly inspired by the ability to control the growth of complexity of both exploration and learning in human children. More precisely, we focus on *intrinsic motivations* guiding mechanisms, responsible of spontaneous exploration, and on *maturational evolution* of the neural and muscular systems, that progressively allow the organism to control novel muscles, and thus, to increase its number of degrees of freedom (Lungarella and Berthouze, 2002). Therefore, we present a system using both self-motivation, and neuro-physiological maturation in an integrated computational mechanism, that aim to guide a robot, to gradually explore and learn its sensorimotor space.

1. Competence Based Intrinsic Motivation System

Previous work, presented in (Oudeyer et al., 2007), (Baranes and Oudeyer, 2009) introduced IAC and R-IAC as two knowledge-based (Oudeyer and Kaplan, 2008) computational models of intrinsic motivation in which a robot was motivated to explore sensorimotor subspaces where its predictions of the consequences of its actions increased maximally fast. These algorithms were shown to allow for self-organized developmental trajectories (Oudeyer et al., 2007) as well as efficient active learning of sensorimotor forward models. (Oudeyer and Kaplan, 2008) introduced the competence based intrinsic motivation framework, in which measures of interest are related to properties of the achievement of self-determined goals rather than to the properties of forward model predictions. In this poster, we introduce a competence based version of the R-IAC system (Baranes and Oudeyer, 2009). This system is composed of a forward model (to be learnt), a goal selection system which chooses goals with a probability proportional to the expected progress in their mastery, and a controller/planner which allows the robot to reach a selected goal reusing the forward model. In analogy to R-IAC, the space of potential goals, i.e. of sensorimotor configurations to be reached, is split into subregions in each of which the mastery progress is monitored. Hence, the mastery progress, defined as the derivative of the evolution of errors in reaching goals in a particular subregion, is used as the measure of interestingness of given goals. Thanks to this measure of mastery progress, this algorithm allows the robot to explore and attempt goals of gradually increasing complexity. Additionally, we couple this mechanism with physiological maturation constraints as we will now explain.

2. Physiological Constraints

Over the first years, the physiological development of infants represents a very important constraint for the exploration and learning process. An important aspect of the maturation of the neural system is the myelinization process, which leads to the extensive development of the corticospinal tract. This internal constraint allows a gradual development of the infants ability to control the distal musculature, from the trunk to hands (proximo-distal vector), and from the center of the body outwards (cephalo-caudal vector) (Kuipers, 1981). For instance, in the reaching task, the fact that young infants predominately use the musculature of the proximal arm and trunk, symplify the learning problem by reducing the functional degrees-of-freedom of the arm. In this poster, we propose to introduce such maturational constraints, progressively unfreezing the robot degrees of freedom, and their interaction with the intrinsic motivation system, described previously.

3. First Experiments

3.1 Competence Based Curiosity

Considering a robot controlled in a *configuration space* **C**, and evolving in an *operational space* **O**, monitoring the mastery of reaching goals can be treated at different level. Firstly, the ***macro-level*** considers the mastery to reach a precise goal state: in the case of an arm control task, a goal state can be described as a vector of precise value, in the *operational space* $\mathbf{O}$ (for instance, the position of the arm extremity). Secondly, the mastery study can be performed at a ***micro-level***, which is defined as the competence to perform micro-movements toward a precise goal in $\mathbf{O}$, from states in the operational space. Therefore, in the ***macro-level*** vision, we are able to monitor the competence level, to precisely reach a goal, and, in the ***micro-level***, we are able to monitor the competence level, to perform micro/primitives actions.

In the first serie of experiment presented here, we analyse the behavior of the ***micro-level*** approach, while the full version of the poster will introduce the ***macro-level***.

3.2 Proximo-Distal Evolution

The proximo-distal and cephalo-caudal maturation of humans can be described as the release of each controllable joint, following a sequence which depends on its morphology. This sequence can be implemented as a graph, whose each node represents an available joint, and each link is assigned with a weight representing a needed maturational level, to evolve to the next joint. Here, the needed maturational level is described as depending on two notions: (1) The *maturational* needed age, which has to be handcrafted, as a genetically coded value, it could be fixed proportionnaly to the number of learning experiments. (2) The *global competence level*, which depends on the global competence progress, allowing the passage to the next joint only if it is stabilized.

3.3 Experiment

The following experiment involved a simulated single *eye*, with pan/tilt rotations capabilities controled by joints velocities $(\dot{q}_{11}, \dot{q}_{12}) \in \mathbf{C}$, and a 2-joints *arm*, controled by $(\dot{q}_{21}, \dot{q}_{22}) \in \mathbf{C}$, possessing a visible extremity (simulating a hand). The couple $(x, y) \in \mathbf{O}$ represents the hand position in the *eye* referential, and $v \in \mathbf{O}$, a Boolean value meaning the presence of the hand, in the camera sight. The global system replies to the mapping $(q, \dot{q})_t \mapsto (v, x, y)_{t+1}$ with $q = (q_{11}, q_{12}, q_{13}, q_{14})$ and $\dot{q} = (\dot{q}_{11}, \dot{q}_{12}, \dot{q}_{13}, \dot{q}_{14})$. By choosing goal values $(v, x, y) \in \mathbf{O}$ and trying to reach it, the system is able to learn the forward model $(q, \dot{q})_t \mapsto \delta(v, x, y)_{t+1}$ (where $\delta(v, x, y)_{t+1} = (v, x, y)_{t+1} - (v, x, y)_t$), and motor skills (policies) $\pi(v, x, y, q) = \dot{q}$ of different complexities. In the studied configuration, we can point different kind of skill complexity, like the ones where the hand is not in the sight of the camera ($(v, x, y) = 0$), which can be considered as easy space subregions, or the ones where it is, which contains more skills to learn. The proposed experiment consists of observing the learning behavior of three approaches of goal selection $(v, x, y) \in \mathbf{O}$. The first one, which represents a uniform selection inside the whole space, is called *Random*. The second approach, called *RIAC Competence* represents the implementation of the Competence Based Intrinsic Motivation heuristic (without physiological constraints). Finally, the *Proximo-Distal Competence Based Curiosity* (PDCC) heuristic (using physiological constraints) is evaluated considering two stages, the system beginning by just moving its camera (values $(\dot{q}_{11}, \dot{q}_{12})$), and freeing its arm (values $(\dot{q}_{21}, \dot{q}_{22})$) on the second stage.

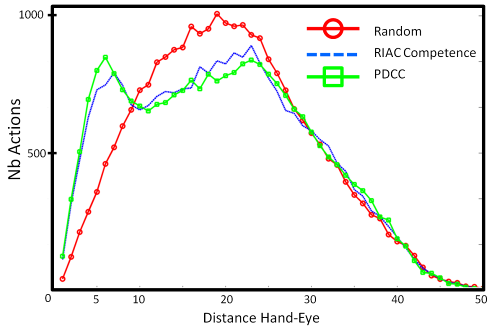

Figure 1: Histograms of Hand-Eye distances

The previous figure shows that RIAC Competence-based guides the eye to focus on the hand, more than the random guiding approach, and that the PDCC approach guides it toward the hand more than the two others. This allows us to argue that using both physiological constraint and competence based approaches can guide the system to avoid a too important focalization on too simple areas and guide it toward skills of intermediate or high complexity.

References

Baranes, A. and Oudeyer, P.-Y. (2009). R-iac: Robust intrinsically motivated active learning. In *Proc. of the IEEE International Conference on Learning and Development.*

Kuipers, H. (1981). *Anatomy of the descending pathways.* American Physiological Society.

Lungarella, M. and Berthouze, L. (2002). Adaptivity via alternate freeing and freezing of degrees of freedom. In *Proc. of the 9th Intl. Conf. on Neural Information Processing.*

Oudeyer, P.-Y. and Kaplan, F. (2008). How can we define intrinsic motivations ? In *Proc. Of the 8th Conf. On Epigenetic Robotics.*

Oudeyer, P.-Y., Kaplan, F., and Hafner, V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11(2):pp. 265–286.