Know your body through intrinsic goals

Francesco Mannella[1]*, Vieri Santucci[1], Eszter Somogyi[2], Lisa Jacquey[2], Kevin J. O'Regan[2], Gianluca Baldassarre[1]

[1] *Institute of cognitive sciences and technologies, National Research Council - CNR, Rome, Italy*
[2] *Laboratoire Psychologie de la Perception (UMR 8158), Paris Descartes - CPSC, Paris, France*

Correspondence*:
Francesco Mannella
Institute of cognitive sciences and technologies
Via San Martino della Battaglia, 44
00185 Roma, Italia
francesco.mannella@gmail.com

ABSTRACT

The first 'object' that newborn children start start to play with is their own body. This activity allows them to autonomously form a sensorimotor map and a repertoire of actions that constitutes the core of future cognitive and motor development. In this work we propose a computational model incorporating the hypothesis that this acquisition of early knowledge is not guided by random motor-babbling, but rather by goals autonomously generated and set on the basis of intrinsic motivations. During the initial motor-babbling, the system forms representations of sensory events. When the agent realises the possibility to re-activate those representations through its motor behaviour, it will be intrinsically motivated to improve its competence in obtaining those specific events. More precisely, the discovered events become intrinsic goals that guide both the learning and the selection of motor actions. The model is based on five components: (1) a competitive neural network, supporting the acquisition of abstract representations based on experienced changes in the sensory input; (2) a selector that on the basis on competence-based intrinsic motivations (CB-IMs) determines the pursued goal and which motor resources will be trained to obtain that goal; (3) an echo-state neural network that controls the movements of the robot and supports the acquisition of the motor skills; (4) a predictor of the accomplishment of the pursued goal, used to measure the improvement of the system competence; (5) the generator of the CB-IM signal that biases the activity of the selector. The model is tested as the controller of a simulated simple planar robot composed of two kinematic 3DoF arms exploring own body in a 2D environment. Sensory information from self-touching is used by the system to form goals and guide skill learning. Results are presented, together with their possible implications for ongoing empirical experiments with human babies. Moreover, the model will be discussed in relation to possible applications to design new open-ended learning robotic architectures able to act in unstructured environments.

Keywords: developmental robotics, developmental psychology, intrinsic motivations, goals, body schema

1 INTRODUCTION

The first "object" that newborn children start to play with is their own body. This activity, that starts even in the fetus and continues for many years after birth (Bremner et al., 2008), determines the formation of a "body schema" (cit.), a sensorimotor map and a repertoire of actions that constitute the core of future cognitive and motor development. [TODO, Kevin]

If we observe children in their first months, they seem to alternate phases of rest to phases of random activity of the limbs that gradually became more controlled and targeted during their development. Differently from this empirical evidence, in this work we propose a computational model incorporating the hypothesis that early knowledge in children is not acquired through random motor-babbling, but guided by self-generated goals, autonomously set on the basis of intrinsic motivations (IMs).

The concept of IMs was introduced in animal psychology during the 1950s and then extended in human psychology (Berlyne, 1950; White, 1959; Berlyne, 1960; Deci and Ryan, 1985; Ryan and Deci, 2000) to describe a set of motivations that were incompatible with the Hullian theory of drives (Hull, 1943) where motivations were strictly connected to the satiation of primary needs. Different experiments (e.g. Harlow, 1950; Montgomery, 1954; Kish, 1955; Glow and Wtnefield, 1978) showed how exploration, novel or surprising neutral stimuli and even the possibility to affect the environment are able to modify the behaviour of the agents driving the acquisition of knowledge and skills in the absence of tasks directly established by biological fitness. Further neurophysiological research (e.g. Chiodo et al., 1980; Horvitz, 2000; Redgrave and Gurney, 2006) showed how IMs can be linked to neuromodulators activity, and in particular to dopamine. These results highlighted the role of IMs in enhancing neural plasticity and driving the learning of new skills. Following biological inspiration, IMs has been also introduced in machine learning (e.g. Barto et al., 2004; Schmidhuber, 2010) and developmental robotics (e.g. Oudeyer et al., 2007; Baldassarre and Mirolli, 2013) to foster the autonomous development of artificial agents and the open-ended learning of repertoires of skills. Depending on their functions and mechanisms, different typologies of IMs have been identified (Oudeyer and Kaplan, 2007; Santucci et al., 2013; Barto et al., 2013) and broadly into two main groups (Baldassarre et al., 2014): (1) knowledge-based IMs (KB-IMs), divided in (1a) novelty based IMs related to novel non-experienced stimuli, and (1b) prediction-based IMs, related to the violation of the agent's predictions; and (2) competence-based IMs (CB-IMs) related to action, i.e. to the agent's competence to change the world and accomplish self-defined $goals$. In their first implementations in computational research, IMs have been used to generate the learning signal for autonomous skills acquisition (Oudeyer et al., 2007; Hart and Grupen, 2011; Mirolli et al., 2013; Kompella et al., 2015). Recent research has started to use IMs for the autonomous generation and/or selection of $goals$ which can then drive the acquisition of skills (Merrick, 2012; Baranes and Oudeyer, 2013; Santucci et al., 2016) and the optimisation of learning processes in high-dimensional action spaces with redundant robot controllers (Baranes and Oudeyer, 2013; Rolf and Steil, 2014).

Together with IMs, $goals$ are a crucial element for the presented model. Here, in consonance with computational and empirical perspectives (Russell and Norvig, 2003; Thill et al., 2013), goals are intended as agent's internal representations of a world/body state or event (or of a set of them), with these properties: (a) the agent can keep the representation active even in the absence of the corresponding state or event; (b) the representation has the power to focus the behaviour of the agent towards the accomplishment of the goal and to generate a learning signal when the world state matches the goal ("goal-matching").

Our hypothesis is that goals and IMs play an important role even in the early phases of knowledge acquisition, i.e. in the first months after birth. In particular, the infant initial motor-babbling results in the formation of sensory events representations. When the baby discovers (through further random activity) the possibility to re-activate the same representations through its behaviour, this will generate a CB-IM signal for obtaining those specific sensory events. More precisely, the discovered events become intrinsic goals that guide both the learning and the selection of motor actions.

Given the important role of IMs in enhancing exploration, a different hypothesis on the development of early knowledge could exclude the involvement of an high-level construct such as goals.
TODO] Kevin, you should develop here the hypothesis that we discussed with you in the past (what we initially called the ‘oudeyer hypothesis’).
If we remember some key points of it were as follows.
The agent is endowed with a repertoire of actions.
The agent is able to recognise a certain number of ‘interesting’ outcomes.
The problem of its development is to learn a repertoire of action-outcome contingencies on the basis of intrinsic motivations.
The acquisition process works as follows.
The agent explores the environment with its actions.
When an interesting outcome is detected, the agent memorise in its repertoire the action-outcome contingency just experienced.
The action and the perceptual experience involved in the contingecy now receive a high motivation for exploration, so that the agent tends to produce similar actions and obtain similar contingencies, so learning other action outcome contingencies that are progressively stored in their repertoire.
When these action-outcome contingencies become progressively more refined the motivation for them decreases and exploration progressively shifts to other areas of the action-outcome space.

In this paper we only investigate our hypothesis, and leave the alternative one for future comparison. In particular, we implement a model (sec. **??** and sec. **??** tested as the controller of a simulated planar robot composed of two kinematic 3DoF arms exploring its own body in a 2D environment (sec. **??**). Sensory information from self-touch activity is used by the system to form goals and drive skill learning. Results of the experiment are presented (sec. 3) together with their possible implications for ongoing empirical experiments with human babies (sec 3). The final section of the paper (sec: 4) discusses relevant related literature and possible future development of the presented model.

2 THE MODEL

3 RESULTS

sectionPredictions of the model

4 DISCUSSION AND FUTURE WORKS

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

The Author Contributions section is mandatory for all articles, including articles by sole authors. If an appropriate statement is not provided on submission, a standard one will be inserted during the production process. The Author Contributions statement must describe the contributions of individual authors referred to by their initials and, in doing so, all authors agree to be accountable for the content of the work. Please see here for full authorship criteria.

FUNDING

This project has received funding from the European Union?s Horizon 2020 Research and Innovation Program under Grant Agreement no. 713010 (GOAL-Robots - Goal-based Open-ended Autonomous Learning Robots).

ACKNOWLEDGMENTS

The authors want to thanks ...

REFERENCES

Baldassarre, G. and Mirolli, M. (2013). *Intrinsically motivated learning in natural and artificial systems* (Berlin: Springer)

Baldassarre, G., Stafford, T., Mirolli, M., Redgrave, P., Ryan, R., and Barto, A. (2014). Intrinsic motivations and open-ended development in animals, humans, and robots: an overview. *Frontiers in Psychology – Cognitive Science* 5, E1–5

Baranes, A. and Oudeyer, P.-Y. (2013). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems* 61, 49 – 73

Barto, A., Mirolli, M., and Baldassarre, G. (2013). Novelty or surprise? *Frontiers in Psychology – Cognitive Science* 4, E1–15

Barto, A. G., Singh, S., and Chantanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills. In *Proceedings of the Third International Conference on Developmental Learning (ICDL)*. 112–119

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratiory behaviour. *British Journal of Psychology. General Section* 41, 68–80. doi:10.1111/j.2044-8295.1950.tb00262.x

Berlyne, D. E. (1960). *Conflict, Arousal and Curiosity* (McGraw Hill, New York)

Bremner, A. J., Mareschal, D., Lloyd-Fox, S., and Spence, C. (2008). Spatial localization of touch in the first year of life: Early influence of a visual spatial code and the development of remapping across changes in limb position. *Journal of Experimental Psychology: General* 137, 149

Chiodo, L. A., Antelman, S. M., Caggiula, A. R., and Lineberry, C. G. (1980). Sensory stimuli alter discharge rate of dopamine (da) neurons: evidence for two functional types of da cells in the substantia nigra. *Brain research* 189, 544–549

Deci, E. L. and Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior* (Springer Science & Business Media)

Glow, P. H. and Wtnefield, A. H. (1978). Response-contingent sensory change in a causally structured environment. *Animal Learning & Behavior* 6, 1–18

Harlow, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of Comparative and Physiological Psychology* 43, 289–294

Hart, S. and Grupen, R. (2011). Learning generalizable control programs. *IEEE Transactions on Autonomous Mental Development* 3, 216–231

Horvitz, J. C. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* 96, 651–656

Hull, C. (1943). *Principles of behavior: An introduction to behavior theory* (New York: Appleton-century-crofts)

Kish, G. B. (1955). Learning when the onset of illumination is used as reinforcing stimulus. *Journal of Comparative and Physiological Psychology* 48, 261–264

Kompella, V. R., Stollenga, M., Luciw, M., and Schmidhuber, J. (2015). Continual curiosity-driven skill acquisition from high-dimensional video inputs for humanoid robots. *Artificial Intelligence*

Merrick, K. E. (2012). Intrinsic motivation and introspection in reinforcement learning. *IEEE Transactions on Autonomous Mental Development* 4, 315–329

Mirolli, M., Santucci, V. G., and Baldassarre, G. (2013). Phasic dopamine as a prediction error of intrinsic and extrinsic reinforcements driving both action acquisition and reward maximization: A simulated robotic study. *Neural Networks* 39, 40 – 51

Montgomery, K. C. (1954). The role of the exploratory drive in learning. *J Comp Physiol Psychol* 47, 60–4

Oudeyer, P.-Y. and Kaplan, F. (2007). What is intrinsic motivation? a typology of computational approaches. *Front Neurorobot* 1, 6. doi:10.3389/neuro.12.006.2007

Oudeyer, P.-Y., Kaplan, F., and Hafner, V. (2007). Intrinsic motivation system for autonomous mental development. In *IEEE Transactions on Evolutionary Computation*. 703–713

Redgrave, P. and Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nat Rev Neurosci* 7, 967–75. doi:10.1038/nrn2022

Rolf, M. and Steil, J. J. (2014). Efficient exploratory learning of inverse kinematics on a bionic elephant trunk. *IEEE Transactions on Neural Networks and Learning Systems* 25, 1147–1160

Russell, S. J. and Norvig, P. (2003). *Artificial Intelligence: A Modern Approach* (Upper Saddle River, NJ: Prentice Hall). Second edition

Ryan, R. M. and Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology* 25, 54–67

Santucci, V. G., Baldassarre, G., and Mirolli, M. (2013). Which is the best intrinsic motivation signal for learning multiple skills? *Frontiers in Neurorobotics* 7

Santucci, V. G., Baldassarre, G., and Mirolli, M. (2016). GRAIL: a Goal-discovering Robotic Architecture for Intrinsically-motivated Learning. *IEEE Transactions on Cognitive and Developmental Systems* 8, 214–231

Schmidhuber, J. (2010). Formal theory of creativity, fun, and intrinsic motivation (1990-2010). *Autonomous Mental Development, IEEE Transactions on* 2, 230 –247

Thill, S., Caligiore, D., Borghi, A. M., Ziemke, T., and Baldassarre, G. (2013). Theories and computational models of affordance and mirror systems: an integrative review. *Neuroscience & Biobehavioral Reviews* 37, 491–521

White, R. (1959). Motivation reconsidered: the concept of competence. *Psychological Review* 66, 297–333