

Autonomous development and learning in artificial intelligence and robotics: Scaling up deep learning to human-like learning

Pierre-Yves Oudeyer

► **To cite this version:**

Pierre-Yves Oudeyer. Autonomous development and learning in artificial intelligence and robotics: Scaling up deep learning to human-like learning. Behavioral and Brain Sciences, Cambridge University Press (CUP), 2017, 40, pp.1-6. 10.1017/S0140525X17000243 . hal-01653604

HAL Id: hal-01653604

<https://hal.inria.fr/hal-01653604>

Submitted on 5 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Autonomous development and learning in artificial intelligence and robotics: Scaling up deep learning to human-like learning

Pierre-Yves Oudeyer

Inria and Ensta Paris-Tech, France, 33405 Talence, France

pierre-yves.oudeyer@inria.fr

<http://www.pyoudeyer.com>

***Abstract:** Autonomous lifelong development and learning is a fundamental capability of humans, differentiating them from current deep learning systems. However, other branches of artificial intelligence have designed crucial ingredients towards autonomous learning: curiosity and intrinsic motivation, social learning and natural interaction with peers, and embodiment. These mechanisms guide exploration and autonomous choice of goals, and integrating them with deep learning opens stimulating perspectives.*

Deep learning (DL) approaches made great advances in artificial intelligence, but are still far away from human learning. As argued convincingly by Lake et al., differences include human capabilities to learn causal models of the world from very little data, leveraging compositional representations and priors like intuitive physics and psychology. However, there are other fundamental differences between current DL systems and human learning, as well as technical ingredients to fill this gap, that are either superficially, or not adequately, discussed by Lake et al.

These fundamental mechanisms relate to *autonomous development and learning*. They are bound to play a central role in artificial intelligence in the future. Current DL systems require engineers to manually specify a task-specific objective function for every new task, and learn through off-line processing of large training databases. On the contrary, humans learn autonomously open-ended repertoires of skills, deciding for themselves which goals to pursue or value, and which skills to explore, driven by intrinsic motivation/curiosity and social learning through natural interaction with peers. Such learning processes are incremental, online, and progressive. Human child development involves a progressive increase of complexity in a curriculum of learning where skills are explored, acquired, and built on each other, through particular ordering and timing. Finally, human learning happens in the physical world, and through bodily and physical experimentation, under severe constraints on energy, time, and computational resources.

In the two last decades, the field of Developmental and Cognitive Robotics (Cangelosi and Schlesinger, 2015; Asada et al., 2009), in strong interaction with developmental psychology and neuroscience, has achieved significant advances in computational

modeling of mechanisms of autonomous development and learning in human infants, and applied them to solve difficult AI problems. These mechanisms include the interaction between several systems that guide active exploration in large and open environments: curiosity, intrinsically motivated reinforcement learning (Schmidhuber, 1991; Oudeyer et al., 2007; Barto, 2013), and goal exploration (Baranes and Oudeyer, 2013), social learning and natural interaction (Chernova and Thomaz, 2014; Vollmer et al., 2014), maturation (Oudeyer et al., 2013) and embodiment (Pfeifer et al., 2007). These mechanisms crucially complement processes of incremental online model building (Nguyen and Peters, 2011), as well as inference and representation learning approaches discussed in the target article.

Intrinsic motivation, curiosity and free play. For example, models of how motivational systems allow children to choose which goals to pursue, or which objects or skills to practice in contexts of free play, and how this can impact the formation of developmental structures in lifelong learning, have flourished in the last decade (Baldassarre and Mirolli, 2013; Gottlieb et al., 2013). In depth models of intrinsically motivated exploration, and their links with curiosity, information-seeking and the “child-as-a-scientist” hypothesis (see Gottlieb et al., 2013 for a review), have generated new formal frameworks and hypotheses to understand their structure and function. For example, it was shown that intrinsically motivated exploration, driven by maximization of learning progress (i.e. maximal improvement of predictive or control models of the world, see Schmidhuber, 1991; Oudeyer et al., 2007;) can self-organize long-term developmental structures, where skills are acquired in an order and with timing that shares fundamental properties with human development (Oudeyer and Smith, 2016). For example, the structure of early infant vocal development self-organizes spontaneously from such intrinsically motivated exploration, in interaction with the physical properties of the vocal systems (Moulin-Frier et al., 2014). New experimental paradigms in psychology and neuroscience were recently developed and support these hypotheses (Kidd, 2012; Baranes et al., 2014).

These algorithms of intrinsic motivation are also highly efficient for multitask learning in high-dimensional spaces. In robotics, they allow efficient stochastic selection of parameterized experiments and tasks, enabling to collect data and update skill models incrementally, through automatic and online generation of a learning curriculum. Such active control of the growth of complexity, enables robots with high-dimensional continuous action spaces to learn omnidirectional locomotion on slippery surfaces, and versatile manipulation of soft objects (Baranes and Oudeyer, 2013), or hierarchical control of objects through tool use (Forestier and Oudeyer, 2016). Recent work in deep reinforcement learning has included some of these mechanisms to solve difficult reinforcement learning problems, with rare or deceptive rewards (Bellemare et al., 2016; Kulkarni et al., 2016), as learning multiple (auxiliary) tasks in addition to the target task

simplifies the problem (Jaderberg et al., 2016). However, there are many unstudied synergies between models of intrinsic motivation in developmental robotics and Deep RL systems, e.g. curiosity-driven selection of parameterized problems (Baranes and Oudeyer, 2013), of learning strategies (Lopes and Oudeyer, 2012), and combinations between intrinsic motivation and social learning (Nguyen and Oudeyer, 2013), have not yet been integrated with deep learning.

Embodied self-organization. The key role of physical embodiment in human learning has also been extensively studied in robotics, and yet it is out of the picture in current deep learning research. The physics of bodies and their interaction with their environment, can spontaneously generate structure guiding learning and exploration (Pfeifer and Bongard, 2007). For example, mechanical legs reproducing essential properties of human leg morphology, generate human-like gaits on mild slopes without any computation (Collins et al., 2005), showing the guiding role of morphology in infant learning of locomotion (Oudeyer, 2016). Yamada et al. (2010) developed a series of models showing that hand-face touch behaviours in the foetus, and hand looking in the infant, self-organize through interaction of a non-uniform physical distribution of proprioceptive sensors across the body with basic neural plasticity loops. Work on low-level muscle synergies also showed how low-level sensorimotor constraints could simplify learning (Flash and Hochner, 2005).

Human learning as a complex dynamical system. Deep learning architectures often focus on inference and optimization. Although these are essential, developmental sciences suggested many times that learning happens through complex dynamical interaction among systems of inference, memory, attention, motivation, low-level sensorimotor loops, embodiment, and social interaction. Although some of these ingredients are part of current DL research, (e.g. attention and memory), the integration of other key ingredients of autonomous learning and development opens stimulating perspectives for scaling up to human learning.

References

- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y. & Yoshida, C. (2009). Cognitive developmental robotics: A survey. *IEEE Transactions on Autonomous Mental Development* 1(1) :12-34.
- Baldassare, G. & Mirolli, M. (2013) *Intrinsically motivated learning in natural and artificial systems*. Springer-Verlag, Berlin.

- Baranes, A. & Oudeyer, P.-Y. (2013) Active learning of inverse models with intrinsically motivated goal exploration in robots *Robotics and Autonomous Systems* 61(1):49-73.
- Baranes, A.F., Oudeyer, P.Y. & Gottlieb, J. (2014) The effects of task difficulty, novelty and the size of the search space on intrinsically motivated exploration. *Front. Neurosci.* 8 :1–9.
- Barto, A. (2013) Intrinsic motivation and reinforcement learning. In: Baldassarre, G., Mirolli, M. eds., *Intrinsically Motivated Learning in Natural and Artificial Systems*. Springer 17–47.
- Bellemare, M., Srinivasan, S. Ostrovski, G., Schaul, T., Saxton, D. & R. Munos (2016). Unifying count-based exploration and intrinsic motivation. *Advances in Neural Information Processing Systems* 30.
- Cangelosi, A. & Schlesinger, M. (2015). *Developmental robotics : From babies to robots*. MIT Press.
- Chernova, S. & Thomaz, A.L. (2014) Robot learning from human teachers. *Synthesis lectures on artificial intelligence and machine learning*. Morgan & Claypool.
- Collins, S., Ruina, A., Tedrake, R. & Wisse, M. (2005). Efficient bipedal robots based on passive-dynamic walkers. *Science* 307(5712) :1082-85.
- Flash, T., Hochner, B. (2005). Motor primitives in vertebrates and invertebrates. *Current Opinion in Neurobiology* 15(6) :660-66.
- Forestier, S. & Oudeyer, P.-Y. (2016) Curiosity-driven development of tool use precursors: a computational model. In: *Proceedings of the 38th Annual Conference of the Cognitive Science Society*.
- Gottlieb, J., Oudeyer, P.-Y., Lopes, M. & Baranes, A. (2013) Information seeking, curiosity and attention: Computational and neural mechanisms *Trends in Cognitive Science* 17(11) :585-96.
- Jaderberg, M., Mnih, V., Czarnecki, W. M., Schaul, T., Leibo, J. Z., Silver, D. & Kavukcuoglu, K. (2016) Reinforcement learning with unsupervised auxiliary tasks. *arXiv preprint arXiv:1611.05397*.

- 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.
- Kulkarni, T.D., Narasimhan, K.R., Saeedi, A. & Tenenbaum, J.B. (2016) Hierarchical deep reinforcement learning: integrating temporal abstraction and intrinsic motivation. Available at : <https://arxiv.org/abs/1604.06057>.
- Lopes, M. & Oudeyer, P.-Y. (2012) The strategic student approach for life-long exploration and learning. In: *IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL)*. IEEE 1–8.
- Moulin-Frier, C., Nguyen, M. & Oudeyer, P.-Y. (2014) Self-organization of early vocal development in infants and machines: The role of intrinsic motivation. *Front. Cogn. Sci.* 4 :1–20. Available at: <http://dx.doi.org/10.3389/fpsyg.2013.01006>.
- Nguyen-Tuong, D. & Peters, J. (2011). Model learning for robot control: A survey. *Cognitive Processing* 12(4) :319-40.
- Nguyen, M. & Oudeyer, P-Y. (2013) Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation learner, *Paladyn Journal of Behavioural Robotics* 3(3):136–46.
- Oudeyer, P.-Y., Kaplan, F. & Hafner, V. (2007) Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput* 11(2) :265–86.
- Oudeyer, P-Y. & Smith. L. (2016) How evolution may work through curiosity-driven developmental process *Topics in Cognitive Science* 1-11
- Oudeyer, P-Y. (2016) What do we learn about development from baby robots? *WIREs Cognitive Science* doi: 10.1002/wcs.1395, Wiley. Available at : <http://www.pyoudeyer.com/oudeyerWiley16.pdf>.
- Oudeyer, P-Y., Baranes, A. & Kaplan, F. (2013) Intrinsically motivated learning of real-world sensorimotor skills with developmental constraints. In: *Intrinsically Motivated Learning in Natural and Artificial Systems* eds. Baldassarre G. & Mirolli M. Springer.
- Pfeifer, R., Lungarella, M. & Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science* 318(5853) :1088-93.

Schmidhuber, J. (1991) Curious model-building control systems. Proceedings of the International Joint Conference on Neural Network 2 :1458–63.

Vollmer, A-L., Mühlig, M., Steil, J.J., Pitsch, K., Fritsch, J., Rohlfing, K. & Wrede, B. (2014). Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning, PLoS ONE 9(3): e91349.

Yamada, Y., Mori, H. & Kuniyoshi, Y. (2010) A fetus and infant developmental scenario: Selforganization of goal-directed behaviors based on sensory constraints. In :10th International Conference on Epigenetic Robotics.145-52.