September 02, 2014

Developmental Learning of Sensorimotor Models for Control in Robotics

By <u>Pierre-Yves Oudeyer</u>
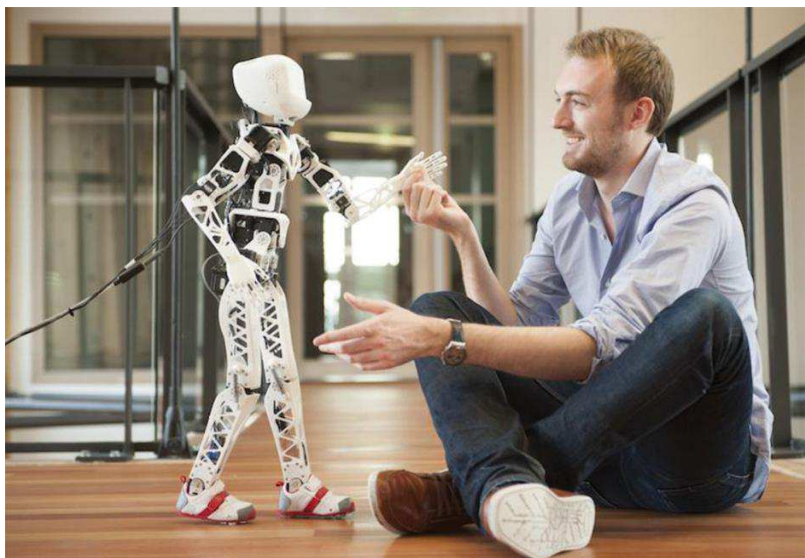
The design of humanoid robots for rich, safe interactions with everyday environments and humans requires techniques for online learning of sensorimotor and social control skills. The open-source 3D-printed humanoid robot Poppy shown here allows systematic experimentation in the integration of such learning methods with adequately designed morphologies (http://www.poppy-project.org; photo Inria/H. Raguet).

Robotics is evolving today toward applications in which robots enter into everyday life and thus need to be able to interact with non-engineers; robots that assist people with disabilities in their homes are an example [5]. For such applications, robots must be able to manipulate and use objects in the human environment that cannot be known and modelled beforehand by engineers. For reasons of safety and energy efficiency, the use of soft materials and soft actuators is flourishing in these applications [12]; difficult control problems arise, however, because the dynamics of these soft materials are difficult to model analytically. For these reasons, a crucial goal has become the development of learning methods that allow robots to acquire models of the dynamics of their own bodies and their interactions with new objects.

Researchers working on these learning methods face several challenges: (1) the sensorimotor spaces to be modelled are high-dimensional and strongly nonlinear; (2) learning should occur incrementally through physical interaction, as it is impossible to anticipate all situations "in factory"; (3) data has to be acquired autonomously through sensorimotor experiments that are costly in time and energy. It is not possible to rely on pre-built databases containing learning examples, fed in a batch manner to statistical inference algorithms. A major challenge is to find ways to acquire data efficiently: Learning by randomly chosen experiments is bound to fail; recently, developmental learning methods, partially inspired by infant development, have been studied to guide the process of data acquisition [10].

Several strands can be distinguished among learning methods. Methods of active learning that allow the choice of experiments that maximize model improvement combine advanced empirical evaluation of information gain [3] with stochastic action-selection algorithms [8]. These methods have made possible, for example, the acquisition of omnidirectional locomotion skills in a compliant robot on slippery surfaces [3]. Other methods have exploited the ability of (non-engineer) humans to learn by imitation. Demonstration of a skill by a human is used as input to a stochastic optimization method in an effort to adapt to the physical particularities of the robot; the idea is to infer either which aspects of the demonstrated skills are relevant [2,4] or the hidden objective through inverse reinforcement learning [1]. The combination of active autonomous learning and learning by imitation was recently approached through methods allowing the active choice of learning strategies [9]. Developers of methods making up other strands have explored how processes of maturation, progressively freeing degrees of freedom in the motor and perceptual spaces, could accelerate learning in large spaces [7].

Finally, a topic of central importance is the co-design of body morphologies, controllers, and learning methods: The integrated design of body geometry, distribution of mass, and material properties can indeed considerably facilitate the acquisition of control skills [11]. This explains recently released robotic platforms that allow the combination of rapid prototyping of body morphologies (based on 3D printing) and control methods [6].

References
[1] P. Abbeel, A. Coates, M. Quigley, and A.Y. Ng, *An application of reinforcement learning to aerobatic helicopter flight*, Adv. Neural Inf. Process. Syst., 19 (2007), 1.
[2] B.D. Argall, S. Chernova, M. Veloso, and B. Browning, *A survey of robot learning from demonstration*, Robot. Auton. Syst., 57:5 (2009), 469–483.
[3] A. Baranes and P.-Y. Oudeyer, *Active learning of inverse models with intrinsically motivated goal exploration in robots*, Robot. Auton. Syst., 61:1 (2013), 49–73.
[4] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, *Robot programming by demonstration*, in Handbook of Robotics, B. Siciliana and O. Khatib, eds., Springer, New York, 2007, 1371–1389.
[5] B. Gates, *A robot in every home*, Scientific American, 296:1 (2007), 58–65.
[6] M. Lapeyre et al., *Poppy Project: Open-source fabrication of 3D printed humanoid robot for science, education and art*, Proceedings of Digital Intelligence, 2014; https://flowers.inria.fr/LapeyreetalDI2014.pdf.
[7] J. Law, P. Shaw, M. Lee, and M. Sheldon, *From Saccades to play: A model of coordinated reaching through simulated development on a humanoid robot*, IEEE Trans. Autonomous Mental Development, in press.
[8] M. Lopes, T. Lang, M. Toussaint, and P.-Y. Oudeyer, *Exploration in model-based reinforcement learning by empirically estimating learning progress*, Adv. Neural Inf. Process. Syst., 25 (2012), 206–214.
[9] M. Nguyen, and P.-Y. Oudeyer, *Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation learner*, Paladyn, J. Behavioral Robotics, 3:3 (2013),136–146.
[10] P.-Y. Oudeyer, A. Baranes, and F. Kaplan, *Intrinsically motivated learning of real-world sensorimotor skills with developmental constraints*, in Intrinsically Motivated Learning in Natural and Artificial Systems, G. Baldassarre and M. Mirolli, eds., Springer, New York, 2013, 303–365.
[11] R. Pfeifer, M. Lungarella, and F. Iida, *Self-organization, embodiment, and biologically inspired robotics*, Science, 318:5853 (2007), 1088–1093.
[12] D. Trivedi, C.D. Rahn, W.M. Kier, and I.D. Walker, *Soft robotics: Biological inspiration, state of the art, and future research*, Appl. Bion. and Biomech., 5:3 (2008), 99–117.

Pierre-Yves Oudeyer is research director at Inria in Talence, France; www.pyoudeyer.com. The SIAM Activity Group on Control and Systems Theory provided this article.