

# Guest Editorial Active Learning and Intrinsically Motivated Exploration in Robots: Advances and Challenges

Manuel Lopes, Pierre-Yves Oudeyer

► **To cite this version:**

Manuel Lopes, Pierre-Yves Oudeyer. Guest Editorial Active Learning and Intrinsically Motivated Exploration in Robots: Advances and Challenges. IEEE Transactions on Autonomous Mental Development, IEEE, 2010, 2 (2), pp.65–69. <inria-00541788>

**HAL Id: inria-00541788**

**<https://hal.inria.fr/inria-00541788>**

Submitted on 1 Dec 2010

**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.

# Guest Editorial

## Active Learning and Intrinsically Motivated Exploration in Robots: Advances and Challenges

### I. TWO CONVERGING APPROACHES TO EXPLORATION IN ROBOTS

**L**EARNING techniques are increasingly being used in today's complex robotic systems. Robots are expected to deal with a large variety of tasks using their high-dimensional and complex bodies, to manipulate objects and also, to interact with humans in an intuitive and friendly way. In this new setting, not all relevant information is available at design time, and robots should typically be able to learn, through self-experimentation or through human-robot interaction, how to tune their innate perceptual-motor skills or to learn, cumulatively, novel skills that were not preprogrammed initially. In a word, robots need to have the capacity to *develop* in an open-ended manner and in an open-ended environment, in a way that is analogous to human development which combines genetic and epigenetic factors. This challenge is at the center of the developmental robotics field [7], [35]–[37]. Among the various technical challenges that are raised by these issues, *exploration* is paramount. Self-experimentation and learning by interacting with the physical and social world is essential to acquire new knowledge and skills.

Exploration in physical robotic agents poses challenging problems due to the high-dimensional and complex dynamics of the body-environment system (especially when other agents are part of the environment), and to the open-ended nature of the sensory-motor spaces of real environments. Typically, in those spaces, the lack of adequate constraints in exploration strategies can result in at best very slow learning, and most often in no consistent learning that can even be dangerous or destructive.

This special issue<sup>1</sup> provides a discussion forum and presents novel contributions on *intentional exploration*, i.e., internal mechanisms and constraints that explicitly foster organized exploration. Of central interest are mechanisms that push agents to select actions that allow them to gain maximal knowledge or maximal control/competence over their bodies and environments. To this, end different fields suggested different approaches, ranging from operational heuristics that indirectly target maximal information gain, to theoretically

optimal policies in relation to various criteria. The field of statistical learning theory studied approaches such as optimal experimental design and *active learning* [1]–[6], [19], [39], and the field of developmental and cognitive robotics studied approaches based on *intrinsic motivation* [7], [9], [23], [24].

Active learning was initially developed for classification and regression learning problems where the cost of querying a given data point for its label or output value is high. Therefore, the goal is to find strategies to minimize the number of queried data to reach a given level of performance, thus maximizing the usefulness of each new experiment. This goal can be instantiated in various settings [5], such as pool-based querying, sequential acquisition, experimental design, among others. A large diversity of criteria can be used to evaluate the utility of given sampling candidates, such as the maximization of prediction errors [22], the local density of already queried points [40], the maximization of the decrease of global model variance [2], expected improvement [11], or maximal uncertainty of the model [41], among others. There have been active-extensions to most of the existing learning methods, e.g., logistic regression [12], support vector machines [13], and gaussian processes [14]–[16]. It has been proved that, under mild conditions, an active learner outperforms a passive learner [6], [17], provided that we have some information on the structure of the problem, even when we cannot estimate in advance that improvement [18]. However, the extension of these guarantees to more general models is still an open question.

Another approach to exploration came from an initially different problem, that of understanding how robots could achieve cumulative and open-ended learning. This raised the question of the task-independent mechanisms that may allow a robot to get interested in practicing skills that were not specified at design time. Two communities of researchers, the first one in reinforcement learning [9], [19], [23], [24], [42], the second one in developmental robotics [7], [20], [43]–[45], formalized, implemented, and experimented several mechanisms based on the concept of intrinsic motivation (sometimes called curiosity-driven learning) grounded into theories of motivation, spontaneous exploration, free play, and development in humans [10], [51], [52], as well as in recent findings in the neuroscience of motivation [46]–[48].

As argued in [19], [24], and [38], architectures based on intrinsically motivated learning can be conceptualized as active learning mechanisms which, in addition to allowing for the self-organized formation of behavioral and developmental complexity, can also allow an agent to efficiently learn a model

Digital Object Identifier 10.1109/TAMD.2010.2052419

<sup>1</sup>Two IEEE technical committees sponsor this special issue, *CIS TC on Autonomous Mental Development* and the *RAS TC on Robot Learning*, to incorporate contributions from a large diversity of approaches.

of the world by parsimoniously designing its own experiments/queries. Yet, in spite of these similarities between work in active learning and intrinsic motivation, these two strands of approaches often differ in their underlying assumptions and constraints, leading to sometimes very different active learning algorithms. Many active learning models make combinations of the following assumptions:

- it is possible to learn a model of the complete world within the lifetime of the learning agent;
- the world is learnable everywhere;
- the noise is homogeneous.

Given those assumptions, heuristics based on the exploration of parts of the space where the learned model has maximal uncertainties or where its prediction are maximally wrong are often very efficient. Yet, these assumptions typically do not hold in real-world robots in an unconstrained environment: the sensorimotor spaces, including the body dynamics and its interactions with the external world, are simply much too large to be learned entirely in a lifetime even with optimal exploration strategies; there are typically subspaces which are unlearnable due to inadequate learning biases or unobservable variables; noise can be strongly heterogeneous. Thus, different authors claimed that typical criteria used in traditional active learning approaches, such as the search for maximal uncertainty or prediction errors, might get trapped or become inefficient in situations that are common in open-ended robotic environments [7], [9]. To approach those challenges, different criteria were elaborated, such as the search for maximal reduction in prediction error, maximal compression progress, or maximal competence progress [7], [9]. Only very recently have these approaches been applied to robotic problems, and even more recently if we consider examples with *real* robots. Nevertheless, examples that consider robotic problems already exist for a large variety of problems: mapping [21], [22], reinforcement learning [23], body schema learning [25], imitation [26], [27], exploration of objects and body properties [7], manipulation [28], among many others.

This special issue tries to address several topics. How can traditional active learning heuristics be applied to robotics problems such as motor learning, affordance learning or interaction learning? How to select an active strategy? Are there general purpose methods, or are they task dependent? How can active and intrinsically motivated exploration enable life-long, task-independent learning and development? Is there a unified formalism to both approaches? How precisely do they model human active learning and exploration and its role in development? Can these approaches be used for social tasks, e.g., joint-work and human-robot interaction?

## II. CONTENTS OF THE SPECIAL ISSUE

This special issue contains a selection of six papers among 20 papers that were submitted.

The first paper [29] attacks the fundamental question of what are extrinsic and intrinsic motivations and rewards, and how they relate to each other, by formulating this question within a broad evolutionary perspective. The authors show that when internal reward systems are evolved for organisms that have to

solve a given (set of) task(s) in environments that change and vary, then typically one ends up with reward systems which do not directly and solely encode task-specific rewards, but also encode rewards which push the organism to search explicitly for novelty. This gives several important theoretical and practical insights. First, it shows that, even in situations where there is a predefined task, the *optimal* reward function that should be given to an agent equipped with machine learning algorithms, such as those used for reinforcement learning problems, shall not always be a direct “translation” of the task fitness function. It might be very beneficial to include task-independent reward mechanisms such as the search for novelty. This is an important contribution since it provides a firm theoretical basis to understand how and when task-independent intrinsically motivated exploration can be helpful to solve particular tasks, as illustrated in the various existing experiments in developmental robotics and intrinsically motivated reinforcement learning. Furthermore, this evolutionary perspective also provides the tools to understand the relation between work on intrinsic motivation in developmental robotics with other work recently presented in evolutionary robotics research [49], [50], which showed that when evolving robot controllers for solving a predefined tasks in a changing environment, it can be very useful to use a fitness function which does push the system to search for novelty explicitly and independently from the task. Finally, the evolutionary perspective provided in [29] allows us to understand that there is probably a continuum between what we call “intrinsic” and “extrinsic” motivation, and that they might be entangled in a complex manner in real organisms. This work also provides the practical insight that evolving the reward function of learning robots, even for task-specific setups, might theoretically provide more efficient reward functions than those that can be built by hand. This practical insight is explored and confirmed in a follow-up paper [30], which shows how evolutionary programming method can allow to find in practice those theoretically near-optimal reward functions.

The third paper shows how to control the motion of a virtual eye to optimally gather information in the environment [31]. The problem of detecting desired objects in an image is approached with an Infomax perspective. Images are searched by relying on probabilistic models to decide where and how to sample next, thus avoiding the sampling of the whole image. It is shown how active models of attentional control can speed up importantly the achievement of tasks such as image recognition or interpretation.

An example of using active learning in a human-robot interaction scenario is presented by [32]. Some previous approaches suggested different methods for using active learning in a learning by demonstration scenario [26], [27]. The authors go further in this topic by providing a real-robot implementation, and a careful validation and comparison of the efficiency of different approaches in terms of learning the task and the interaction with humans.

The work presented in [33] compares different motivational approaches such as novelty-seeking, interest, and competence-seeking motivation. It aims at providing a system that is able to incorporate the different mechanisms in a single formalism, and thus to compare quantitatively, in given contexts, various

technical approaches to the elaboration of intrinsic motivation systems in robots.

The final work addresses the problem of how intrinsic motivation can be used to guide the cumulative learning of a hierarchical structure of skills [34]. Active learning is coupled with the introduction of incremental option learning, which are two complementary mechanisms allowing to decrease the time necessary to learn the causal structure of the environment as well as associated skills.

### III. DISCUSSION AND FUTURE PERSPECTIVES

The selection of papers in this issue addresses many important challenges of active learning and intrinsically motivated exploration. In this section, we highlight several open problems which remain to be further addressed.

#### A. Systematic Experiments With Real Robots and Scalability to High-Dimensional Spaces

The use of real robotic platforms is limited in this series of articles as much as in the broader literature on active learning and intrinsic motivation. An important reason for this is that there are many practical difficulties in setting up experiments with real robots such that statistically significant results can be obtained. In particular, not only the physical experiments should be repeated to show how stable is the learning and exploration process, but they should also be performed under different environments and with various exploration strategies for comparisons. As illustrated in the accepted papers, simulations are often used to address this issue, and when physical robots are used, their sensorimotor space is typically tightly constrained to allow for repeated experiments. Yet, this methodological stance biases research to avoid a potentially very challenging problem: how do active learning and intrinsically motivated exploration scale up in practice in real continuous high-dimensional and large volume spaces? It is actually surprising to observe that this question has been very little addressed so far given that among the initial motivations for elaborating active learning and intrinsic motivation systems, the scalability of exploration in such spaces was central. Indeed, there are both practical and theoretical challenges ahead of research. One of them is meta-exploration. All active learning and intrinsic motivation systems are based, in a way or another, on a measure of “interestingness” of particular zones of the sensorimotor/state space, and are used to push systems to focus on zones of maximal interest. But the computation of this measure of interest for a given zone of the sensorimotor space requires that this zone be at least explored/sampled a little bit. In other words, we have a meta-exploration/meta-exploitation dilemma for achieving efficiently the “smart” exploration process. How this meta-exploration/meta-exploitation dilemma could be solved (or maybe cannot be solved without particular constraints) in high-dimensional spaces remains to be understood.

#### B. Comparison, Mixing, and Autonomous Selection of Active Exploration Criteria

Results have already shown that there is no active learning criteria that is optimal for all situations. Taking into account the large variety of exploration criteria, it is important to understand

more formally when and why certain criteria are more suited than others. A related challenge is to devise methods that would allow an autonomous system, e.g., a robot for example, to dynamically mix and select those criteria depending on the current situation. The development of formal guaranties of convergence, and speed of convergence, of the methods, may provide a rule to select among them.

#### C. Integration of Intrinsically Motivated Learning in Complete Autonomous Creatures

A motivation for introducing the use of intrinsic motivation systems in developmental robotics was to address aspects of the problem of open-ended sensorimotor and cognitive development. As argued many times in the literature (e.g., [53]), flexible and adaptive behavior and development in autonomous creatures can hardly be built or understood by only focusing on individual components. Development and behavior self-organize as a result of the complex dynamical interaction between all elements of the cognitive apparatus, of the body, and of the physical and social environment [54]. Yet, most research on active learning and intrinsic motivation have studied these exploration mechanisms in an isolated manner. Further work should try to understand how they might interact with other kinds of motivations systems (e.g., extrinsic motivational systems), within a complete sensorimotor and cognitive architecture, characterized by various constraints such as motor primitives or maturation of the perceptual system, within a body that has peculiar morphological properties and performs morphological computation [55], and within a physical and social environment which also strongly influences the organism.

#### D. Active Learning and Intrinsic Motivation in Multiagent Systems

An aspect to explore further is the use of active learning and intrinsic motivation in human–robot interaction and other multiagent systems. Results have shown that it is possible to learn in a multiagent scenario even without knowledge of the other’s actions and even without knowing that there is another agent [56], [57]. In such settings, exploration is fundamental to ensure convergence. More interesting settings would be based on an intuitive interaction between robots, or robot and humans, based on joint–attention and multimodal communication.

#### E. Evaluation Methodology

As active learning and intrinsic motivation have different origins and motivations, their similarities in terms of technical approaches should be counterbalanced with their dissimilarities in terms of evaluation methodologies. This means that depending on the point of view, they might or might not be considered as identical scientific topics. Both approaches can be evaluated on a single task with a standard loss function. Yet, if we consider a developmental perspective and evaluate the approaches taking into account many tasks and life-long learning, then intrinsic motivation can no longer be seen as only a method for building fast and accurate models of the world, but also a crucial component of a general development program. Under this perspective, exploration guides not only the learning of a single task but a learning process that simultaneously acquires sensory-motor

representations, skills, and general task solutions. This includes hierarchical structures of information and control, among many other complex representations. In that case, a comprehensive evaluation framework, including both quantitative and qualitative analysis, is still to be elaborated.

#### F. Theoretical Understanding of the Nature and Origins of Intrinsic Motivation and Its Relation to Exploration

Intrinsic motivation is a complex concept imported from psychology, and computational modelling can play a role in understanding it better through attempts of formalization and operationalization. Yet, as discussed in [8], there is wide diversity of potential formalization of intrinsic motivation, some of which appearing compliant with for example Berlyne's definition [52], but yet not fostering spontaneous exploration. A paradigmatic example is that of cognitive homeostasis (e.g., [58] and [60]), which is a mechanism which explicitly push autonomous systems to search for low prediction errors and in a task-independent manner. Another example is mechanisms that push organism to be "interested" in synchrony or contingency independently of the particular task or sensorimotor channels at play (e.g., [59]). Thus, the articulation between intrinsic motivation and spontaneous exploration should be further investigated. The complexity of the concept of intrinsic motivation is also illustrated by [29], which shows that searching for hard and fast features for discriminating intrinsic motivation from extrinsic motivation might be impossible. At the same time, [29] provides a very original outline of an hypothesis for understanding the evolutionary origins of intrinsic motivation in animals, humans in particular. It would be particularly stimulating to see this hypothesis going back to and being assessed by psychology, neuroscience, and evolutionary biology.

MANUEL LOPES, *Guest Editor*

School of Computing and Mathematics  
University of Plymouth  
Plymouth, PL4 8AA U.K.  
E-mail: manuelcabidolopes@gmail.com

PIERRE-YVES OUDEYER, *Guest Editor*  
INRIA

Talence, 33640 France  
E-mail: pierre-yves.oudeyer@inria.fr

#### REFERENCES

- [1] D. Angluin, "Queries and concept learning," *Mach. Learn.*, vol. 2, pp. 319–342, 1988.
- [2] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, "Active learning with statistical models," *J. Artif. Intell. Res.*, vol. 4, pp. 129–145, 1996.
- [3] K. Chaloner and I. Verdinelli, "Bayesian experimental design: A review," *J. Statist. Sci.*, vol. 10, 1995.
- [4] M. Seeger and H. Nickisch, "Compressed sensing and bayesian experimental design," in *Proc. Int. Conf. Mach. Learn.*, 2008.
- [5] B. Settles, Active Learning Literature Survey Univ. Wisconsin-Madison, Madison, WI, 2009, CS Tech. Rep. 1648.
- [6] R. Castro and R. Novak, "Minimax bounds for active learning," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, May 2008.
- [7] P.-Y. Oudeyer, F. Kaplan, and V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 265–286, Apr. 2007.
- [8] P.-Y. Oudeyer and F. Kaplan, "How can we define intrinsic motivation?," in *Proc. 8th Int. Conf. Epigenetic Robot.: Modelling Cogn. Develop. Robot. Syst.*, Lund, Sweden, 2001.
- [9] J. Schmidhuber, "Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts," *Connect. Sci.*, vol. 18, no. 2, pp. 173–187, Jun. 2006.
- [10] R. W. White, "Motivation reconsidered: The concept of competence," *Psychol. Rev.*, vol. 66, pp. 297–333, 1959.
- [11] D. Jones, M. Schonlau, and W. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.*, vol. 13, no. 4, pp. 455–492, 1998.
- [12] A. Schein and L. H. Ungar, "Active learning for logistic regression: An evaluation," *Mach. Learn.*, vol. 68, pp. 235–265, 2007.
- [13] G. Schohn and D. Cohn, "Less is more: Active learning with support vector machines," in *Proc. 17th Int. Conf. Mach. Learn.*, Seattle, WA, 2000.
- [14] A. Kapoor, K. Grauman, R. Urtasun, and T. Darrell, "Active learning with Gaussian processes for object categorization," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Crete, Greece, 2007.
- [15] A. Krause and C. Guestrin, "Nonmyopic active learning of Gaussian processes: An exploration-exploitation approach," in *Proc. 24th Int. Conf. Mach. Learn.*, Corvallis, OR, 2007.
- [16] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *J. Mach. Learn. Res.*, vol. 9, pp. 235–284, 2008.
- [17] M. F. Balcan, S. Hanneke, and J. Wortman, "The true sample complexity of active learning," in *Proc. Conf. Learn. Theory*, Helsinki, Finland, 2008.
- [18] S. Dasgupta, "Analysis of a greedy active learning strategy," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2005, pp. 337–344.
- [19] J. Schmidhuber, "A possibility for implementing curiosity and boredom in model-building neural controllers," in *Proc. From Animals to Animats: 1st Int. Conf. Simulation Adapt. Behav.*, Cambridge, MA, 1991, pp. 222–227.
- [20] P. Oudeyer, F. Kaplan, V. Hafner, and A. Whyte, "The playground experiment: Task-independent development of a curious robot," in *Proc. AAAI Spring Symp. Develop. Robot.*, Stanford, CA, 2005, pp. 42–47.
- [21] R. Martinez-Cantin, N. de Freitas, E. Brochu, J. Castellanos, and A. Doucet, "A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot," *Autonom. Robot. – Special Issue Robot Learn.*, pt. B, 2009.
- [22] S. Thrun, "Exploration in active learning," in *Handbook of Brain Science and Neural Networks*, M. Arbib, Ed. Cambridge, MA: MIT Press, 1995, pp. 381–384.
- [23] S. Singh, A. Barto, and N. Chentanez, "Intrinsically motivated reinforcement learning," in *Proc. 18th Annu. Conf. Neural Info. Process. Syst.*, Vancouver, BC, Canada, 2004.
- [24] A. Barto, S. Singh, and N. Chentanez, "Intrinsically motivated learning of hierarchical collections of skills," in *Proc. 3rd Int. Conf. Develop. Learn.*, San Diego, CA, 2004, pp. 104–111.
- [25] R. Cantin-Martinez, M. Lopes, and L. Montesano, "Body schema acquisition through active learning," in *Proc. IEEE Int. Conf. Robot. Autom.*, Anchorage, AK, 2010.
- [26] M. Lopes, F. S. Melo, and L. Montesano, "Active learning for reward estimation in inverse reinforcement learning," in *Proc. Eur. Conf. Mach. Learn.*, Bled, Slovenia, 2009.
- [27] S. Chernova and M. Veloso, "Interactive policy learning through confidence-based autonomy," *J. Artif. Intell. Res.*, vol. 34, pp. 1–25, 2009.
- [28] S. Hart, S. Sen, and R. Grupen, "Intrinsically motivated hierarchical manipulation," in *Proc. 2008 IEEE Conf. Robot. Autom.*, Pasadena, CA, 2008.
- [29] S. Singh, R. L. Lewis, A. G. Barto, and J. Sorg, "Intrinsically motivated reinforcement learning: An evolutionary perspective," *IEEE Trans. Autonom. Mental Develop.*, vol. 2, no. 2, pp. 70–82, 2010.
- [30] S. Niekum, A. Barto, and L. Spector, "Genetic programming for reward function search," *IEEE Trans. Autonom. Mental Develop.*, vol. 2, no. 2, pp. 83–90, Jun. 2010.
- [31] N. J. Butko and J. R. Movellan, "Infomax control of eye movements," *IEEE Trans. Autonom. Mental Develop.*, vol. 2, no. 2, pp. 91–107, Jun. 2010.
- [32] M. Cakmak, C. Chao, and A. L. Thomaz, "Designing interactions for robot active learners," *IEEE Trans. Autonom. Mental Develop.*, vol. 2, no. 2, pp. 108–118, Jun. 2010.
- [33] K. Merrick, "A comparative study of value systems for self-motivated exploration and learning by robots," *IEEE Trans. Autonom. Mental Develop.*, vol. 2, no. 2, pp. 119–131, Jun. 2010.
- [34] C. M. Vigorito and A. G. Barto, "Intrinsically motivated hierarchical skill learning in structured environments," *IEEE Trans. Autonom. Mental Develop.*, vol. 2, no. 2, pp. 132–143, Jun. 2010.

- [35] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, "Autonomous mental development by robots and animals," *Science*, vol. 291, pp. 599–600, 2001.
- [36] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connect. Sci.*, vol. 15, no. 4, pp. 151–190, 2003.
- [37] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: A survey," *IEEE Trans. Autonom. Mental Develop.*, vol. 1, no. 1, pp. 12–47, 2009.
- [38] A. Baranes and P.-Y. Oudeyer, "R-IAC: Robust Intrinsicly motivated exploration and active learning," *IEEE Trans. Autonom. Mental Develop.*, vol. 1, no. 3, pp. 155–169, 2009.
- [39] V. Fedorov, *Theory of Optimal Experiments*. New York: Academic, 1972.
- [40] S. Whitehead, A Study of Cooperative Mechanisms for Faster Reinforcement Learning Univ. Rochester, Rochester, NY, Tech. Rep. TR-365.
- [41] S. Thrun and K. Möller, J. Moody, S. Hanson, and R. Lippmann, Eds., "Active exploration in dynamic environments," in *Proc. Adv. Neural Info. Process. Syst. 4*, Denver, CO, 1992.
- [42] R. S. Sutton, "Integrated architectures for learning, planning, and reacting based on approximating dynamic programming," in *Proc. 7th Int. Conf. Mach. Learn.*, Washington DC, 2001, pp. 216–224.
- [43] X. Huang and J. Weng, C. Prince, Y. Demiris, Y. Marom, H. Kozima, and C. Balkenius, Eds., "Novelty and reinforcement learning in the value system of developmental robots," in *Proc. 2nd Int. Workshop Epigenetics Robot.: Modeling Cogn. Develop. Robot. Syst.*, Lund, Sweden, 2002, vol. 94.
- [44] D. Blank, D. Kumar, L. Meeden, and J. Marshall, "Bringing up robot: fundamental mechanisms for creating a self-Motivated, self-organizing srchitecture," *Cybern. Syst.*, vol. 36, no. 2, 2005.
- [45] M. Schembri, M. Mirolli, and G. Baldassare, "Evolution and learning in an intrinsically motivated reinforcement learning robot," in *Proc. 9th Eur. Conf. Artif. Life*, Lisbon, Portugal, 2007, pp. 294–303.
- [46] W. Schultz, P. Dayan, and P. Montague, "A neural substrate of prediction and reward," *Science*, vol. 275, pp. 1593–1599, 1997.
- [47] S. Kakade and P. Dayan, "Dopamine: Generalization and bonuses," *Neural Netw.*, vol. 15, pp. 549–559, 2002.
- [48] P. Redgrave and K. Gurney, "The short-latency dopamine signal: A role in discovering novel actions?," *Nature Rev. Neurosci.*, vol. 7, pp. 967–975.
- [49] S. Doncieux and J. B. Mouret, "Behavioral diversity measures for evolutionary robotics," in *Proc. IEEE Congr. Evol. Comput.*, Barcelona, Spain, 2010.
- [50] J. Lehman and K. Stanley, "Abandoning objectives: Evolution through the search for novelty alone," *Evol. Comput. J.*, 2010, to be published.
- [51] E. Deci and R. Ryan, *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum, 1985.
- [52] D. Berlyne, *Conflict, Arousal and Curiosity*. New York: McGraw-Hill, 1960.
- [53] R. A. Brooks, "How to build complete creatures rather than isolated cognitive simulators," in *Architectures for Intelligence*, K. Van Lehn, Ed. New York: Erlbaum, 1991, pp. 225–239.
- [54] L. B. Smith and E. Thelen, "Development as a dynamic system," *Trends Cogn. Sci.*, vol. 7, pp. 343–348.
- [55] R. Pfeifer, M. Lungarella, and F. Iida, "Self-organization, embodiment, and biologically inspired robotics," *Science*, vol. 318, pp. 1088–1093, 2007.
- [56] F. Melo and M. Lopes, "Convergence of independent adaptive learners," in *Proc. Progress Artif. Intell.: 13th Portuguese Conf. Artif. Intell., Lecture Notes Artif. Intell.*, 2007, vol. 4874, pp. 555–567.
- [57] D. S. Leslie and E. J. Collins, "Generalised weakened fictitious play," *Games Econ. Behav.*, vol. 56, no. 2, pp. 285–298, 2006.
- [58] P. Andry, P. Gaussier, J. Nadel, and B. Hirsbrunner, "Learning invariant sensorimotor behaviors: A developmental approach to imitation mechanisms," *Adapt. Behav.*, vol. 12, pp. 377–385, 2003.
- [59] M. Rolf, M. Hanheide, and K. J. Rohlfing, "Attention via synchrony. Making use of multimodal cues in social learning," *IEEE Trans. Autonom. Mental Develop.*, vol. 1, pp. 55–67.
- [60] H. von Foerster, "On constructing a reality," in *Environmental Research Design*. Stroudsburg, PA: Dowden, Hutchinson, Ross, 1973, vol. 2, pp. 35–46.



**Manuel Lopes** received the Ph.D. degree in electrical and computer engineering, and has participated in various international research projects in the areas of robotics and cognitive systems.

He is currently a Lecturer in Humanoid Robotics & Intelligent Systems at the Cognitive Robotics and Neural Systems Research Centre of the University of Plymouth, U.K. His current research interests are the combination of robotics and machine learning to provide better machines and to model cognitive behavior. He aims at developing robots that can learn complex skills through a development perspective by combining learning by exploration, by reward, and by imitation.



**Pierre-Yves Oudeyer** received the M.Sc. degree in theoretical computer science from Ecole Normale Supérieure de Lyon, France, and received the Ph.D. degree in artificial intelligence from the University Paris VI, France.

After eight years as a permanent researcher at Sony CSL Paris, he is now a Research Scientist at INRIA, France, where he heads the FLOWERS team. He has published a book, more than 60 papers in international journals and conferences, and received several prizes for his work in developmental robotics and on the origins of language. He is interested in the mechanisms that allow humans and robots to develop perceptual, motivational, behavioral, and social capabilities to become capable of sharing cultural representations.

Dr. Oudeyer is Editor of the IEEE CIS NEWSLETTER ON AUTONOMOUS MENTAL DEVELOPMENT, and Associate Editor of IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT, *Frontiers in Neurobotics*, and of the *International Journal of Social Robotics*.