IEEE CIS

AMD Newsletter

Robotics
Psychology
Neuroscience
Machine Intelligence

The Newsletter of the Autonomous Mental Development Technical Committee

VOLUME 7, NUMBER 2 ISSN 1550-1914 October 2010

Editorial

A major event in the evolution of our field will happen next year: the two major conferences focused on computational developmental sciences, ICDL and the Epigenetic Robotics conference, will be jointly organized in Frankfurt in August 2011. After a decade of successful meetings in parallel involving various communities not always totally aware of each other's work, this joint event will be a great opportunity to observe in a single place the whole landscape of these interdisciplinary sciences. The call for papers should be publicized soon, and everyone is encouraged to submit contributions.

This month, the newsletter highlights a topic that has recently gathered a lot of interest in developmental robotics: intrinsic motivation and its relation to open-ended cumulative learning of skills. The dialog initiated by Gianluca Baldassarre and Marco Mirolli, on "What are the key open challenges for understanding autonomous cumulative learning of skills?", has received answers from Andy Barto, Kevin Gurney and Peter Redgrave, Juergen Schmidhüber, Kenji Doya, Jun Tani and myself. As emphasized in the synthesis, those contributions cover a wide range of complementary questions. It shows that both theory and practice around these issues are already well structured and progressing fast.

Intrinsic motivation and active learning is also the topic of the first special thematic issue of IEEE Transactions on Autonomous Mental Development, guest edited by Manuel Lopes and myself and just published. It establishes some bridges in machine learning and robotics between computational approaches to development and more classical approaches to active learning. The table of contents is provided in the newsletter together with the other recent issues of IEEE TAMD. Three other special issues are on their way: Representations and architectures for cognitive systems, Grounding language in action, and Computational modeling of neural and brain development. This shows the vitality and growth of the journal. Again, I strongly encourage new contributions to this journal, as it will help both further structuring and visibility of our field.

Reacting to the stimulating dialog on the symbol grounding problem in the April 2010 issue of the newsletter, John Weng proposes a novel call for dialog, entitled "Are natural languages symbolic in the brain?". It directly addresses a quite controversial but fundamental issue: Is the symbol grounding problem a real problem? Are symbols really fundamental for understanding human cognitive development, or are they just a conceptual invention of modern human culture? Interested researchers are welcome to submit a response (contact pierre-yves.oudeyer@inria.fr or weng@cse.msu.edu) by March 15th, 2011. The length of each response should be between 300 and 500 words (including references).

- Pierre-Yves Oudeyer, INRIA, Editor

Message from the Chair of IEEE AMD Technical Committee

I, as AMD TC Chair, participated in IEEE ICS AdCom meetings on June 17 and 18th, 2010 in conjunction with WCCI 2010, Barcelona, Spain, and gave the report of our activities. AMD TC has the largest number (76) of committee members among all TCs in IEEE CIS, but it was pointed out that the geographical balance (more than half from North America) and the gender balance (very few female members) should be improved. Therefore, we collectively need to address this issue and any proposals to improve the current situation will be very appreciated.

The main activity of our TC is to organize the International Conference on Development and Learning. ICDL 2010 has been a successful and stimulating conference, as summarized in this newsletter by Thomas Schultz and Benjamin Kuipers. ICDL 2011 will be held next summer, and a big news is that it will be a joint event with EpiRob 2011 (co-organization).

- Minoru Asada, New Chair of the AMD TC

Dialog Column

What are the Key Open Challenges for Understanding Autonomous Cumulative Learning of Skills?

Gianluca Baldassarre, Marco Mirolli
Laboratory of Computational Embodied Neuroscience, Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Rome, Italy

The capacity to autonomously learn a number of different skills in a cumulative fashion is one of the hallmarks of intelligence and is at the core of the Autonomous Mental Development endeavour (Weng et al. 2001). We ask our colleagues to identify the key open challenges for understanding autonomous cumulative learning in organisms and reproducing it in robots (although social processes are extremely important for human development, here we focus only on nonsocial, fully autonomous learning). As stressed by the recently funded EU Integrated Project 'IM-CLeVeR' – Intrinsically Motivated Cumulative Learning Versatile Robots (http://im-clever.eu), the problem of cumulative learning can be divided into two general sub-problems: (a) which are the signals that can drive cumulative learning? (b) which control and learning architectures can support the cumulative acquisition of skills?

Learning signals. Most research devoted to develop autonomous learning robots focuses on the solution of single tasks and hence uses task-specific learning signals. These signals have a strong limit for cumulative learning in that they can only drive the acquisition of new skills strictly related to the task(s) decided by the researcher. Inspired by the findings of both animal and human psychology, several researchers have started to explore the possibility of using non-task-specific learning signals generated by intrinsic motivations. Several fundamental issues remain open in this regard, for example: Which kind (s) of intrinsic motivations does cumulative learning need (for a useful taxonomy, see (Oudeyer & Kaplan, 2007))? Almost all of the proposed intrinsic learning signals are based on the stimuli that the learning system perceives and processes internally, i.e. on its knowledge (e.g. Schmidhuber, 1991; Oudeyer et al. 2007): does cumulative learning also need intrinsic learning signals based on what the system does, i.e. on its competence (see Barto et al. 2004; Schembri et al. 2007 for some first proposals)? What are the relationships between intrinsic motivations and other motivations? What are the brain mechanisms underlying them in real organisms (e.g., phasic dopamine: Redgrave & Gurney 2006)?

Architectures. Cumulative learning requires that new skills are acquired on the basis of the previous ones without disrupting them. Which kind of architecture can support open-ended cumulative learning? For example, research on hierarchical reinforcement learning (Barto & Mahadevan, 2003) has been developing systems based on a rather strict relationship between the functional hierarchy of skills and their structural/architectural organization. These approaches have been mainly developed at a theoretical level and for discrete problems: can they scale up to allow cumulative learning in robotic systems? How? On the other hand, it has been demonstrated that structural modularity and hierarchy are not strictly necessary for producing functionally modular and hierarchical behaviours (Botvinick & Plaut, 2004; Yamashita & Tani, 2008). But can these approaches avoid the problem of catastrophic forgetting when requested to scale to truly open-ended cumulative learning? How? Between these two extremes there is a range of possible solutions which raise a number of open challenges: How can skills be segmented and encoded in 'soft modules'? How many hierarchical levels are needed? What are the functions of these levels and their relations? How can skills be re-used, generalised, and composed?

References:

- Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M. & Thelen, E. (2001). Autonomous mental development by robots and animals. *Science*, 291, pp. 599–600.
- Oudeyer, P. & Kaplan, F. (2007). What is Intrinsic Motivation? A Typology of Computational Approaches. *Frontiers in Neurorobotics*, 1(1), pp. 1-14.
- Schmidhuber, J. (1991). A possibility for implementing curiosity and boredom in model-building neural controllers. In Arcady-Meyer J. & Wilson S.W., (eds), *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, Cambridge, MA : MIT Press, pp. 222–227.
- Oudeyer, P.; Kaplan, F. & Hafner, V. V. (2007), Intrinsic Motivation Systems for Autonomous Mental Development, *IEEE Transactions on Evolutionary Computation,*11(2), pp. 265–286.
- Barto, A., Singh, S. & Chentanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills. In Movellan, J.R., Chiba, A., Deak, G., Triesch, J., Bartlett, M. S. (eds.), *Proceedings of the Third International conference of Development and Learning*.
- Schembri, M., Mirolli, M. & Baldassarre, G. (2007). Evolving childhood's length and learning parameters in an intrinsically motivated reiforcement learning robot. In Berthouze, L., Dhristiopher, Prince G., Littman, M., Kozima, H. and Balkenius, C. (eds.), *Proceedings of the Seventh International Conference on Epigenetic Robotics*, Lund: Lund University Cognitive Studies, pp. 141–148.

Dialog Column

- Redgrave, P. & Gurney, K. (2006). The short-latency dopamine signal: a role in discovering novel actions? *Nature Reviews Neuroscience*, 7(12), pp. 967–975.
- Barto, A. G. & Mahadevan, S. (2003). Recent advances in hierarchical reinforcement learning, discrete event dynamic systems. *Discrete Event Dynamic Systems*, 13(4), pp. 341–379.
- Botvinick, M. & Plaut, D. C. (2004), Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *PsycholSogical Review*, 111, pp. 395–429.
- Yamashita, Y. & Tani, J. (2008). Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment. *PLoS computational biology*, 4(11), e1000220.

What are Intrinsic Rewards Signals?

Andy Barto
Department of Computer Science, University of Massachussets Amherst, Amherst, USA

A long-held bias of mine is that it is all about control. Knowledge—of whatever kind—is useful only to the extent that it facilitates control. The primacy of control is due to the evolutionary process: control skills conferred evolutionary fitness upon our ancestors; passive knowledge did not. A "world model" is useful only to the extent that it can be used to produce better control rules. This is why I have focused on skill acquisition instead of just model acquisition. The importance of models—and they are important—is through their role in acquiring, refining, and extending skills. Of course, I am not thinking of just sensorimotor skills but other kinds as well. Moreover, the "world" of a world model should be a circumscribed part of the agent's environment that is relevant to achieving its control objectives. No system can afford to model everything.

Evolution also informs my view of how widely reusable skills can be acquired, which is why I became interested in intrinsic motivation and what I believe is its key role in cumulative learning of skills. Recent research (Singh et al. 2010) clarified for me, indeed has changed my view, of what intrinsic motivation is and how it differs from other forms of motivation. Recognizing that motivation is more than just the nature of reward signals, it is nevertheless useful to think just about the origin of reward signals from an evolutionary perspective.

An animal's primary reward signals, i.e., those that do not need to be learned, are generated by mechanisms that have evolved over many generations across many ancestral environments. Evolution endowed the mechanisms that generate primary reward signals with the ability to encourage the learning of *ubiquitously useful* behaviour, that is, behaviour that has proven to be useful across wide swatches of ancestral environments. The classical primary rewards associated with food, sex, escape, etc. obviously have this status, but so do lots of others that facilitate acquiring knowledge useful for control.

Our conclusion from computational studies is that there is no clear-cut distinction between extrinsic and intrinsic reward signals (once it is recognized that all of them are generated by animal-internal machinery). All reward signals exist because of their importance for evolutionary fitness. What we call intrinsic reward signals are those that encourage behaviour whose influence on evolutionary fitness tends to require causal pathways that are particularly indirect and complex.

Therefore, I believe it is not useful to look for definitive features that distinguish intrinsic and extrinsic reward signals. However, the distinction is still useful because it alerts us to the possible benefits of defining reward functions that depend on a wider range of factors than those usually considered in computational reinforcement learning. Specifically, reward functions can depend on the state of an agent's internal environment, which includes remembered and learned information. This opens the door to a large space of reward functions with which we have relatively little computational experience. This space needs to be further explored, along with new algorithms for caching skills, adapting them to specific tasks, and exploiting them efficiently in planning and further learning.

References:

- Singh, S., Lewis, R.L., Barto, A.G., & Sorg, J. (2010). Instrinsically motivated reinforcement learning: An evolutionary perspective. *IEEE Transactions on Autonomous Mental Develop*evelopment, 2(2), pp. 70-82

Dialog Column

Discovering and Deploying New Actions: a Biological Perspective


Kevin Gurney, Peter Redgrave
Department of psychology, University of Sheffield, Sheffield, UK.

When, for the first time, you place a rat in a box with a lever that delivers food, nowhere in its nervous system is there a representation of the action of lever press. However, after some time the rat may be seen sitting in front of the lever, purposefully pressing it and collecting the food pellets as they are delivered. What has happened in the intervening period?

Typically the rat explores the environment, sometimes during which the lever is 'accidentally' pressed and an initially unpredicted food delivery occurs. We know that the unpredicted sensory event causes a phasic release of the neurotransmitter dopamine into the basal ganglia - one of the fundamental processing units of the vertebrate brain associated with action selection and reinforcement learning. The dopamine signal is therefore of the kind alluded to by Baldassarre and Mirolli, in that it is stimulus induced, and reflects the *knowledge* of the animal; in this case the novelty of, or lack of knowledge about, the event. However, we believe these very signals are those necessary in order to accomplish the learning of *competences*. In our proposal (Redgrave & Gurney, 2006) the phasic dopamine signal reinforces the selection of the behaviour immediately preceding the unexpected event (via plastic change in cortico-basal ganglia connections). This repetition bias in selection policy causes the presentation of motor and sensory signals (context and outcome) to occur at brain circuits that can store internal models of action-outcome contingency (e.g. pressing a lever in a cage cause food delivery). This process occurs until the contingency is predicted. Initially if the action is inefficient or unreliable in producing the outcome (e.g a tail swipe, rather than paw press on the lever) the animal explores the action-space until it discovers the action which reliably induces the outcome. Then, prediction ensues and policy is re-normalised, but a new competence has been learned. In this scheme, it is environmental feedback which supplies the signals required for competence learning. We therefore view intrinsic motivation based on knowledge or expectation (Oudeyer & Kaplan, 2007) as an integral part of competence learning.

Baldassarre and Mirolli also highlight the importance of hierarchies. We agree that a strict functional decomposition (tied to architecture) is probably not scalable to arbitrary levels of behavioural complexity, and that, architectural schemes with no pre-imposed semantics may show more promise. Indeed the brain appears to have a large scale, architectural hierarchy comprising multiple loops between cortical (or sub-cortical) inputs to basal ganglia, and return paths to those inputs areas (see for example Yin & Knowlton, 2006). We propose that careful study of the function of this architecture will be rewarded by an understanding of hierarchical, cumulative learning in animals.

References:

- Redgrave, P. & Gurney, K., 2006. The short-latency dopamine signal: a role in discovering novel actions? *Nature Reviews Neuroscience*, 7(12), pp. 967–975.
- Oudeyer, P. & Kaplan, F., 2007. What is Intrinsic Motivation? A Typology of Computational Approaches. *Frontiers in Neurorobotics*, 1(1), pp. 1-14.
- Yin, H.H. & Knowlton, B.J., 2006. The role of the basal ganglia in habit formation. *Nature Reviews. Neuroscience*, 7(6), pp. 464-476.

Knowledge and Competence: Two Sides of the Same Coin

Juergen Schmidhuber
IDSIA, Lugano, Switzerland.

According to the Formal Theory of Creativity and Intrinsic Motivation (IM), a creative agent (one that never stops generating non-trivial, novel, and surprising behaviours and data) must have two learning components: R and P, where R is a general reinforcement learner, and P is an adaptive predictor or compressor of the agent's growing data history (the record of the agent's interaction with its environment). The learning progress of P is the intrinsic reward for R. That is, R is motivated to invent "interesting" spatio-temporal patterns that P does not yet know but can easily learn with little computational effort. To maximize expected reward (in the absence of external reward) R will create more and more-complex behaviours that yield temporarily surprising (but eventually boring) patterns that make P quickly improve.

Dialog Column

I have argued that this principle explains science, art, music and humour. Ongoing work focuses on testing the theory using more powerful methods for prediction and reinforcement learning than were available when the basic principles were first introduced in 1991 (Schmidhuber 1991a; 1991b).

Baldassarre and Mirolli distinguish between knowledge and competence, and claim the latter is a relatively new ingredient of IM-based learning, and ask: "Does cumulative learning also need intrinsic learning signals based on what the system does?". In the light of the above, the answer is: of course. Actions and perceptions are intimately intertwined. Any measure of what the system does and whether it improves is based on sensory feedback partially caused by its actions. That's why even the very first intrinsic reward implementations (Schmidhuber, 1991a) measured not just "knowledge" but also "competence": they motivated the action learner (R) to create behaviours leading to novel but learnable patterns, measuring R's success through the improvements in the prediction of action-caused sensory feedback. So the predictor (P) learns "passive" action-independent knowledge while R continually gains "competence" in the sense that it learns previously unknown skills. The recent confusion about what is and what isn't IM is addressed in a recent survey (Schmidhuber, 2010) which also provides an alternative typology of IM.

A more recent IM-based system (Schmidhuber, 2002) provides an even better example of "knowledge" and "competence" combined in one, collapsing both P and R into single probabilistic programs that both predict and act, and are motivated to create new programs (possibly by re-using previously found code) leading to surprising and therefore intrinsically rewarding outcomes. This system incrementally searches in general program space, automatically including many of the concepts addressed by Baldassarre and Mirolli: hierarchical, cumulative learning (also see here the pioneering work of Mark Ring in Ring 1992, 1994), generalization, recursion, reuse and composition of code, etc. (see papers since 1990 under http://www.idsia.ch/~juergen/subgoals.html). The reinforcement learning or search algorithm in (Schmidhuber, 2002), however, is limited by the power of its program search. I believe that a key challenge will be to re-implement similar general systems using more recent and more powerful program-search techniques, in particular, the Optimal Ordered Problem Solver (Schmidhuber, 2004), or evolutionary techniques for recurrent neural networks (Gomez et al. 2008).

References:

- Schmidhuber, J. (1991a). A possibility for implementing curiosity and boredom in model-building neural controllers. In J. A. Meyer and S. W. Wilson, editors, *Proc. of the International Conference on Simulation of Adaptive Behavior: From Animals to Animats*, pp. 222-227. MIT Press/ Bradford Books.
- Schmidhuber, J. (1991b). Curious model-building control systems. *In Proc. International Joint Conference on Neural Networks*, Singapore, 2, pp. 1458-1463.
- Schmidhuber, J. (2010). Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990-2010). *IEEE Transactions on Autonomous Mental Development*, 2(3), pp. 230-247.
- Schmidhuber, J. (2002). Exploring the Predictable. In Ghosh, S. Tsutsui, eds., *Advances in Evolutionary Computing*, pp. 579-612, Springer.
- Ring, M. B. (1992). Two Methods for Hierarchy Learning in Reinforcement Environments. In J. A. Meyer, H. Roitblat, and S. W. Wilson, editors, *Proc. of the International Conference on Simulation of Adaptive Behavior: From Animals to Animats*, pp. 148-155. MIT Press.
- Ring, M. B. (1994). Continual Learning in Reinforcement Environments, *University of Texas at Austin dissertation*.
- Schmidhuber, J. (2004).. Optimal Ordered Problem Solver. *Machine Learning*, 54, pp. 211-254.
- Gomez, F., Schmidhuber, J., Miikkulainen, R. (2008). Accelerated Neural Evolution through Cooperatively Coevolved Synapses. *Journal of Machine Learning Research (JMLR)*, 9, pp. 937-965.

Open Challenges for Autonomous Cumulative Learning

Kenji Doya
Okinawa Institute of Science and Technology, Japan.

Baldassarre and Mirolli posit two major problems in autonomous cumulative learning: learning signals and modular architectures. The former is related to the computational question of "what is the objective function of an agent" and the latter to the implementational question of "how to avoid interference and to reuse learnt knowledge or skills." Probably the most knowledge-hungry creatures today on earth are web-search engines. They try to visit all the web pages accessible on the net and extract all the words and links to form an exhaustive knowledge of what is where and which is more useful.

Dialog Column

Such exhaustive fishing for information has proven to be useful and is proliferating, but it is a question whether that is also useful for autonomous robots or a good model of living creatures. Exploration of different states and actions, or information structure behind them, is important, but some animals like sea squirts quit exploration at some stage in life and even digest their own nervous system, and they are all fine. The killer problem in designing curiosity is how not to let agents try to learn all the detailed structures in the world, most of which have nothing to do with their life. This is why we decided to build robots that we call Cyber Rodents, which mimic the two major constraint of life: survival and reproduction. We have had some success in letting a colony of robots find out supplementary rewards, such as the reward for the sight of a battery pack or a potential mating (program copying) partner, through embodied evolution (Elfwing et al. in press). In this study, however, the potential sources of rewards were selected by the experimenters and it remains to be seen whether and how robots can find out more general reward signals that help non-random exploration.

The second issue of modularity also poses several open challenges. The primate visual system is composed of multiple cortical areas divided into two pathways: the ventral pathway for object recognition and the dorsal pathway for spatial processing. In many functional brain imaging studies, researchers elucidate distinct functions of these areas based on their differential levels of activation depending on the task instructions despite the same visual stimuli. There is a wealth of papers published in this paradigm, but surprisingly we do not know why and how a particular brain area can activate when its specific function is needed. This is a big question in neuroscience, but a big challenge in learning theory as well. How we can select the right set of modules as needed, and how can they be wired appropriately? How a set of tasks can be decomposed into subtasks, and how can they be allocated within available learning hardware? The theory of self-organization by competition and cooperation had a good success in reproducing feature detector properties of neurons in particular cortical areas, but a theory of self-organization at the level of learning modules is strongly needed (Uchibe & Doya, 2004).

References:

- Elfwing S, Uchibe E, Doya K, Christensen HI (in press). Darwinian embodied evolution of the learning ability for survival. *Adaptive Behavior*.
- Uchibe, E., Doya, K. (2004). Competitive-cooperative-concurrent reinforcement learning with importance sampling. *The Eighth International Conference on the Simulation of Adaptive Behavior*, pp. 287-296.

Self-Organization of Functional Structures in Neural Dynamical Systems through Cumulative Learning

Jun Tani
Riken Brain Science Institute, Saitama, Japan

Self-Organization of Functional Structures in Neural Dynamical Systems through Cumulative Learning
In regard to the issue of cumulative learning, it is said that mammals show consolidation learning (McClelland et al. 1994) in which recent experiences are stored first in short-term memory in the hippocampus and then are transferred to long-term memory in the cortex for memory consolidation. In this scenario, consolidation learning is considered to be much more difficult in the cortex than in the hippocampus because in the cortex such learning would face the problem of catastrophic interference: new knowledge has to be well structured and generalized with previously acquired knowledge stored in memory. However, consolidation learning in the hippocampus might actually be much easier because it might be more like rote learning that does not require structuring, functioning in much the same way as a photocopy does.

Referring to recent results by Botovinick and Plaut (Botvinick & Plaut, 2004) and Yamashita and Tani (Yamashita & Tani, 2008) which showed that functional hierarchy can be self-organized in some dynamic neural network models without explicit network structures provided, Baldassarre and Mirolli asked how such functional structures can develop in the course of cumulative learning in these models. Recent findings by Nishimoto et al. (Nishimoto & Tani, 2009) as well as by Nishimoto and Tani (unpublished data) provide a partial answer to the question. In their research into the incremental learning capability of a dynamic neural network model, the so-called multiple timescale recurrent neural network (MTRNN) (Nishimoto & Tani, 2009), a humanoid robot implemented with MTRNN was tutored for multiple object manipulation tasks in which each task consisted of a sequence of different combinations of behavioural primitives. The network model employed had a fast dynamics part and a slow dynamics part. When the MTRNN was trained incrementally with the sensory-motor sequence pattern for each task along with previously trained ones, it was observed that each primitive behavioural

Dialog Column

pattern was acquired one by one distributedly in the fast dynamics network part while combinatorial sequencing of these primitive behavioural patterns was acquired in the slow dynamics part where compositional functional structures were developed. An interesting observation was that the internal neural representation for the previously learned primitives as well as their control sequences changed gradually as new primitives were incrementally acquired. This observation may indicate how generalized functional structures can self-organize in distributed neural activation dynamics through cumulative learning. Because the generalization of previously learned knowledge and newly acquired knowledge would require that each memory item be embedded in a relational structure with other items rather than being independent, the additional learning of each piece of new knowledge could lead to re-organization of such relationship to some extent. This accounts for the aforementioned observation of the gradual changes seen in internal representation during incremental learning.

To conclude, the problem of cumulative learning prompts us to consider how functional structures with generalization can be achieved through the iterative re-organization of internal structures during the process of incremental learning. Although I expect that the self-organization capability of neural dynamical systems rather than the explicit algorithmic control of them can naturally solve this problem, this assertion needs to be verified in future studies.

References:

- McClelland, J.L., McNaughton, B.L. & O'Reilly, R. (1994). Why there are complementary learning systems in the Hippocampus and Neocortex. *Technical Report PDO*.CNS.94.1, Carnegie Mellon University.
- Botvinick, M. & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111, pp. 395–429.
- Yamashita, Y. & Tani, J. (2008). Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment. *PLoS Computational Biology*, 4(11), e1000220.
- Nishimoto, R. & Tani, J. (2009). Development of hierarchical structures for actions and motor imagery: a constructivist view from synthetic neuro-robotics study, *Psychological Research*, 73, pp. 545-558.

Developmental Constraints on Intrinsically Motivated Exploration

Pierre-Yves Oudeyer
INRIA, France

The sensorimotor space of typical developmental robots is so large (typically unbounded) and high-dimensional that the set of potentially learnable skills is much larger than what can be learnt during a lifetime. So, in an open-ended development setting, the question of how to explore and what to learn is crucial. The solution of exploring randomly or even driven by the search for novelty or uncertainty, as done in traditional active learning or in some curiosity-driven architectures, fails.

This is the reason why intrinsic motivation systems based on peculiar measures such as “compression improvement” or “learning progress” were introduced (Schmidhuber, 1991; Oudeyer et al. 2007). Driven by the maximal improvement of local forward models, i.e. by increase of knowledge about the consequences of one’s own actions, they allow robots to acquire open-ended skills, i.e. competences, as an automatic side effect. Architectures driven by the maximal increase of competences for reaching self-generated goals (improving knowledge only as a side effect) have also been explored and have shown to significantly improve the efficiency of exploration in high-dimensional redundant sensorimotor spaces (Baranes & Oudeyer, 2010a).

Yet, whatever the measures used for intrinsic motivation, there is a major meta-exploration obstacle. These measures essentially answer to the question: How much interesting is a given sensorimotor activity/sub-space? The answer is then used to decide whether it is worth trying again in the future or trying something else. But the measure can only be made after at least a minimum amount of local exploration! And we get into the same initial fundamental problem: it is impossible to sample this measure over the whole sensorimotor space.

Thus, intrinsically motivated exploration needs to be further constrained to harness the very large, potentially unbounded, volume and high-dimensionality of the space defined by the potential interactions between the body and its environment. Inspiration can be taken from infant development.

Dialog Column

Embodiment and morphological computation constraints. If bodies are equipped with appropriate geometries and materials, one can leverage the self-organizing properties of physics to project apparently extremely high-dimensional sensorimotor spaces into constrained low-dimensional manifolds. For example, learning dynamic biped locomotion can be transformed into a relatively low-dimensional problem (see Ly & Oudeyer, 2010; http://flowers.inria.fr/acroban.php).

Motor primitives. Human infants do not learn to control their whole body movements "pixel by pixel". Rather, they are born with muscle synergies, i.e. neurally embedded dynamical systems which generate parameterized coordinated movements, e.g. CPGs. These motor primitives can considerably decrease the size of the explorable space and transform complex low-level action planning problems in higher-level low-dimensional dynamical system tuning problems. This was essential for the acquisition of dynamic motor skills in the Playground Experiment (Oudeyer et al, 2007), and may be also approached in terms of "options" (Barto et al, 2004).

Maturational constraints. Human infants are not born with complete access to all their potential degrees of freedom. The neural system, mainly through myelination, progressively grows, opening for control new muscle synergies and increasing the range and resolution of sensorimotor signals. Initial experiments studying the coupling of intrinsic motivation with maturational constraints show promising results that should be further explored (Baranes & Oudeyer, 2010b).

Social learning. Last but not least, social interaction should be a central companion to intrinsic motivation. The interaction between those two guiding mechanisms is at the centre of educational research. This shall probably also become the case in developmental robotics.

References:

- Schmidhuber, J. (1991). A possibility for implementing curiosity and boredom in model-building neural controllers. In Arcady-Meyer J. & Wilson S.W., (eds), *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour*, Cambridge, MA : MIT Press, pp. 222–227.
- 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.
- Baranes, A., Oudeyer, P-Y. (2010a) Intrinsically Motivated Goal Exploration for Active Motor Learning in Robots: a Case Study, in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010)*, Taipei, Taiwan
- Ly, O., Oudeyer, P-Y. (2010) Acroban the Humanoid: Playful and Compliant Physical Child-Robot Interaction, in *ACM SIGGRAPH'2010 Emerging Technologies*.
- Barto, A., Singh, S. & Chentanez, N. (2004). Intrinsically motivated learning of hierarchical collections of skills. In *Movellan*, J.R., Chiba, A., Deak, G., Triesch, J., Bartlett, M. S. (eds.), *Proceedings of the Third International conference of Development and Learning*.
- Baranes, A., Oudeyer, P-Y. (2010b) Maturationally-Constrained Competence-Based Intrinsically Motivated Learning, in *Proceedings of IEEE International Conference on Development and Learning (ICDL 2010)*, Ann Arbor, Michigan, USA.

Reply and Summary: On the Open Challenges for Understanding Cumulative Learning


Gianluca Baldassarre, Marco Mirolli
Laboratory of Computational Embodied Neuroscience, Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Rome, Italy

We thank all authors who replied to our dialog proposal for their valuable contributions to understanding autonomous cumulative learning.

Barto starts by claiming that 'it's all about control': all that matters (for both organisms and robots) is the acquisition of skills. The acquisition of knowledge is important only in as much as it supports effective behaviour. We completely agree, and appreciate that Barto spelled out this point because it is often overlooked while we think it is paramount. Then Barto approaches the problem of what intrinsic reward signals are from an evolutionary perspective, concluding that the distinction between extrinsic and intrinsic reward signals is quantitative rather than qualitative: the former drive learning of behaviours directly relevant to fitness, whereas the latter tend to have a more indirect relation to fitness. While we agree that the evolutionary approach is very useful and that extrinsic and intrinsic motivations are on a continuum with respect to their relation to fitness, we do not exclude the possibility that useful qualitative distinctions can be found on other grounds.

Dialog Column

Schmidhuber presents his Formal Theory of Creativity and Intrinsic Motivation (IM), according to which IM corresponds to the learning progress of a general predictor given as a reward to a reinforcement learner, and claims that this theory captures all forms of IM. Hence, he criticizes the distinction between knowledge-based and competence-based IM on the basis that in his schema the predictor progress (knowledge) drives the learning of the controller (competence). This criticism stems from an under-appreciation of the distinction between *functions* and *mechanisms* of IM (Mirolli & Baldassare, forthcoming). We agree that knowledge-based intrinsic reinforcements (mechanism) can be used for learning competences (function). But the same function (i.e., driving skill acquisition) can be played also by *competence-based mechanisms,* e.g. measures of progress in skill acquisition (see Schembri et al. 2007 for an example). Which kind of IM are most useful for robots and which IM exist in real animals are interesting empirical questions that cannot be settled a priori.

Gurney and **Redgrave** present their theory of skill acquisition according to which phasic dopamine represents a reinforcement signal produced by unpredicted sensory events that drive the discovery and learning of new actions, rather than a reward prediction error driving instrumental conditioning, as commonly thought in recent neuroscientific literature. Hence, in this view dopamine represents an example of knowledge-based IM driving the acquisition of competences. We consider this hypothesis of the most importance as it represents one of the rare attempts to explain the brain mechanisms underlying action learning. Recently (Santucci et al. 2010) we have proposed a computational model that partially reconciles the two contrasting views on phasic dopamine by showing how learning signals triggered by unexpected events may drive the cumulative acquisition of skills of a reinforcement learning system.

Doya proposes that the key problem in designing efficient (intrinsic) learning signals is preventing that the agent tries to learn everything because the most of what an agent can learn is useless. We agree that this is a crucial issue. Indeed, this is one of the main reasons why we think that knowledge-based IM are not enough: most of the knowledge which might be acquired is irrelevant for the agent, so its acquisition should be canalized towards the acquisition of useful skills, e.g. through competence-based IM. Doya also suggests that a crucial challenge for cumulative learning is how to decompose tasks, allocate them to the systems modules, and select the appropriate modules when needed. We completely agree. Recently we have proposed a hierarchical reinforcement learning system in which the allocation and deployment of modules is autonomously achieved on the basis of their ability to maximize reinforcement in different contexts (Caligiore et al. 2010).

Tani approaches cumulative learning from a different perspective, in which *functional* hierarchical modules self-organize within dynamical neural systems without pre-imposed *structural* hierarchy and modularity. In his contribution he explains how this can lead to cumulative learning thanks to the gradual re-organization of pre-acquired behavioural primitives and control sequences. Although we are not sure that purely functional hierarchy and modularity are enough for prolonged cumulative learning, we think that Tani's work is very important, and consider the further exploration of this approach extremely valuable. In this respect, we suggest that an interesting direction for future research might consist in trying to develop systems that incrementally learn functional hierarchies of behaviours in an autonomous *unsupervised* fashion (currently most of Tani's works use supervised learning).

Finally, **Oudeyer** points to a problem of intrinsically motivated cumulative learning that has been so far overlooked: if the sensorimotor space is huge it is not possible to perform a satisfying sampling of the entire space, as required by IM mechanisms for deciding where to focus learning. He proposes that a solution might come from constraining exploration in different ways: through morphological and maturational constraints, the use of motor primitives, and social learning. We agree that this is indeed a great challenge for cumulative learning, and that those proposed by Oudeyer are important possible solutions for it. In this respect, we think that another fundamental solution (related to the motor-primitive solution) might come from the *development of abstractions:* increasingly more abstract representations of *both perceptions and actions* might allow to progressively reduce the sensorimotor space. Developing architectures and algorithms that build useful abstractions is a major challenge for cumulative learning.

References:

- Mirolli M., Baldassarre G. (forthcoming): Functions and mechanisms of intrinsic motivations: The knowledge versus competence distinction.
- Schembri, M., Mirolli, M., Baldassarre, G. (2007). Evolving childhood's length and learning parameters in an intrinsically motivated reinforcement learning robot. In *EpiRob7*, pp. 141-148.
- Santucci, V. G., Baldassarre, G., Mirolli, M. (2010). Biological cumulative learning requires intrinsic motivation: A simulated robotic study on the development of visually-guided reaching. In *EpiRob10*.
- Caligiore D., Borghi A. M., Mirolli M., Parisi D., Baldassarre G. (2010). A hierarchical modular reinforcement learning model that autonomously learns multiple tasks in continuous state and action spaces. In *EpiRob10*.

Dialog Initiation

Are Natural Languages Symbolic in the Brain?

Juyang Weng
Department of Computer Science and Engineering, Cognitive Science Program, Neuroscience Program, Michigan State University, East Lansing, Michigan, USA.

Many of the readers may think that I am asking a stupid question. However, I hope that this question can lead to meaningful further dialogs and hopefully some clarification following the dialog of ”The Symbol Grounding Problem Has Been Solved: Or Maybe Not?” initiated by Angelo Cangelosi and appeared in AMD Newsletter vol. 7, no. 1, April 2010.

Stevan Harnad wrote: “The symbol grounding problem is the problem of causally connecting symbols inside an autonomous system to their referents in the external world without the mediation of an external interpreter. ... So it is the words of a full-blown natural language (not all of them, but the ones that cannot be grounded by definition in the others) that need to be connected to their referents in the world. Have we solved that problem? Certainly not.” Madden et al. also stated: “the robot will need different kinds of symbols that correspond to the different types of meaningful things that can occur in the world.” Luc Steels claimed “I will take the same decomposition (as Cangelosi) to argue again why I believe the symbol grounding problem has been solved.” Stephen Cowley argued “Inner symbols use contingencies that lack any causal link with the population’s external symbols. Outsiders become insiders by drawing on developmental history. By hypothesis, this depends on co-action.” Aaron Sloman charged from the opposite direction: “They need information about things that are not, or even cannot be, sensed, e.g. things and events in the remote past or too far away in space” “... because deep concepts of explanatory sciences (e.g. ”electron”, ”gene”) cannot be grounded in the claimed way. Symbol grounding is impossible.” Vincent C. Müller was even wilder: “We need a problem that does not assume computationalism ... How does physics give rise to meaning? We do not even know how to start on the hard problem.”

The above authors seemed to agree that the symbol grounding problem is important to clarify. From the very diverse responses from the responders, it seems the way the problem originally posed by Searle’s Chinese Room Argument is misleading. I do not think that all the fields on intelligence — natural and artificial — have sufficiently understood the essence of the problem, let alone have got an effective and well understood solution.

Trying not to be misled by the Searle’s argument, let us discuss more fundamental issues:
1**. Output from the brain in life time**: The brain produces behaviours that gives us an impression (or illusion) that the brain has symbols in it. But are the outputs from the brain symbolic? What are such outputs? They are gestures we produce (visual signs), vibrations of vocal tract that we make (sound we utter), scribbles we write and other forms of muscle contractions. The outputs from the brain seem not really computer symbols (e.g., ASCII code of text) that is always the same and matched uniquely by a computer program. Every instance of a word is different every time we produce it, although we try to produce it the same way each time. Any objections?

2. **Input into the brain in life time**: The brain must receive stimuli in order to develop internal representations and produce outputs. What are such inputs produced by another human for inter-personal communication? They are retinal images of visual languages (e.g., American Sign Language ASL), sound waves as spoken languages, page images as written languages, and signals from somatic receptors as touch languages (e.g., braille). All other stimuli from the environments are also the physical properties that are converted by the receptors into spikes. The inputs into the brain seem not really computer symbols either (e.g., ASCII code of text). Any objections?

3. **Does the brain have rigid symbolic modules in it during its life time**? If it never receives a computer-like system or never produces a computer-like symbol, does a normal (infant or adult!) brain have symbols in it, even if we narrow down to only the function of words in human languages and their relation to meanings? Considering that the brain never senses a word exactly the same twice and never produces a word exactly the same twice, do you think that the brain have internal atomic symbols (or relax: symbolic modules each corresponding to a pure meaning, each module containing multiple neurons, and there are static borders that separate these modules) — words, any other linguistic units, and meanings expressible using a natural language, and actions? In other words, is any natural language truly symbolic in the brain’s internal representation, having clear-cut boundaries between symbols?

Dialog Initiation

4. **"How does physics give rise to meaning?"** I rewrite Müller's question more clearly: How interactions by a grounded brain with the physical word give rise to meanings inside the brain, if the brain does not have rigid symbolic modules in it? Can we address these interesting questions, at least Sloman's? If we truly cannot, let us shut off our computers and search for something non-computational, as Roger Penrose (Penrose, 1994; 1989) proposed.

It is known that cortical regions are typically inter-connected in both directions (Moran & Desimone, 1985; Fellman & Van Essen, 1991; Callaway, 1998). Based on this and well documented neuro-anatomic principles, I have proposed the first, as far as I know at this point, brain-mind network model (Weng, 2010) that has the five "chunks" — development (how the brain-mind emerges), architecture (how areas connect), area (how each building block represents), space (how the brain deals with spatial information) and time (how the brain deals with temporal information). In particular, the model has been preliminarily experimented for visual information processing with top-down attention (Luciw & Weng, 2010), early language acquisition and generalization (Miyan & Weng, 2010) and processing of natural language (from Wall Street Journal) (Weng et al. 2009). I think the model has provided my personal positions to the above four (4) questions. Of course, this brain-mind model has yet to be further biologically verified and enriched (e.g., neuro-modulation).

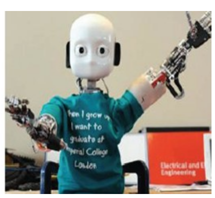
I invite colleagues to provide their perspectives and comments to the above questions, especially those that are different from mine.

References:

- Penrose, R. Shadows of the Mind: A Search for the Missing Science of Consciousness. *Oxford University Press*, Oxford, 1994.
- Penrose, R. The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. *Oxford University Press*, Oxford, 1989.
- Moran, J. and Desimone, R.. Selective attention gates visual processing in the extrastrate cortex. *Science*, 229(4715), pp. 782–784, 1985.
- Felleman, D. J. and Van Essen, D. C.. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, pp. 1–47, 1991.
- Callaway, E. M.Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience*, 21, pp. 47–74, 1998.
- Weng, J. A 5-chunk developmental brain-mind network model for multiple events in complex backgrounds. In Proc. *Int'l Joint Conf. Neural Networks*, pp. 1–8, Barcelona, Spain, July 18-23 2010.
- Luciw, M. and Weng, J. Where What Network 3: Developmental top-down attention with multiple meaningful foregrounds. In Proc. *IEEE International Joint Conference on Neural Networks*, pp. 4233–4240, Barcelona, Spain, July 18-23 2010.
- Miyan, K. and Weng, J. WWN-Text: Cortex-like language acquisition with What and Where. In Proc. *IEEE 9th International Conference on Development and Learning*, pp. 280–285, Ann Arbor, August 18-21 2010.
- Weng, J., Shen, Y., Chi, M., and X. Xue. Temporal context as cortical spatial codes. In Proc. *International Joint Conference on Neural Networks*, Atlanta, Georgia, June 14-19 2009.

Report of the Ninth IEEE International Conference on Development and Learning (IEEE ICDL 2010)





Thomas R. Shultz, McGill University, Canada
Benjamin Kuipers, University of Michigan, USA

This August, 116 researchers from 16 countries gathered at the University of Michigan in Ann Arbor to share their scientific interest in the development of knowledge in both humans and robots. ICDL-2010 was co-sponsored by the IEEE Computational Intelligence Society and by the Cognitive Science Society. The conference also received generous gifts from the Toyota Research Center in Ann Arbor and from Microsoft Research in Redmond, Washington.

ICDL is still small enough to be a single-track meeting. There were four keynote speakers, 23 accepted talks (28% of those submitted), 28 featured posters (35%), and 16 poster abstracts (89% of submitted). The consensus among the organizers was that the quality of papers was notably higher than in early years of the conference.

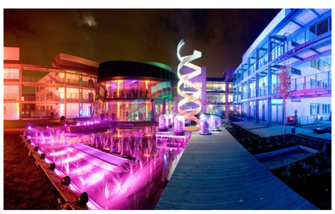
David Vernon of the Italian Institute of Technology gave a keynote talk on the iCub project, a largely European effort that has received more than $11 million in EU funding over the last few years. The central thrust of this project is to create robots with sufficient capabilities to replicate the developmental and learning processes that children go through, in areas as diverse as movement and sensory integration and eventually object recognition, language, and cooperation with humans.

Another keynote address was delivered by Rod Grupen of the University of Massachusetts, who also works on issues of locomotion and grasping in robots designed to learn like infants through reinforcement and intrinsic motivation, implementing Piagetian-style abilities to accommodate and assimilate. Visual tracking leads to touching and ultimately to grasping of objects. In his keynote talk, Felix Warneken of Harvard University spoke about cooperation in young children and chimpanzees. Remarkably, both chimpanzees and 18-month-old children spontaneously help a conspecific in need, even under cost and without being rewarded. In a final keynote address, Susan Gelman of the University of Michigan reported her work on when and how generalization occurs in children's language acquisition. She finds that a notion of "kind" is fundamental to children's fast mapping of words to meanings.

Many interesting talks and posters over the three days of ICDL 2010 made a convincing case for the idea that testing hypotheses in bodies interacting in real time in dynamic environments is indeed a worthy challenge. Best-paper awards and the conference program can be found at http://www.eecs.umich.edu/icdl-2010/home.html.

General Co-Chairs for the meeting were Ben Kuipers and Tom Shultz. Alex Stoytchev and Chen Yu designed the conference program, competently aided by 38 Program Committee members and 93 reviewers. Publicity Co-Chairs Ian Fasel (North America), Jun Tani (Asia), and Jochen Triesch (Europe) helped publicize the conference. Publication Chair Nicholas Butko assembled the program and the Proceedings DVD. Webmasters Vindhya Baddela, Seth Levine, and Jingen Liu created the website. In 2011, ICDL will be jointly organized with Epigenetic Robotics (EpiRob) in Frankfurt August 24-27, with Jochen Triesch and Angelo Cangelosi as general co-chairs, and Ian Fasel and Katharina Rohlfing as program chairs. Keynote speakers will be Andy Barto, Jean Mandler, Erin Schuman, and Michael Tomasello.

Call for Participation IEEE ICDL/Epirob 2011



Dates: August 24-27 2011

Location: Frankfurt, Germany

General chairs: Jochen Triesch, Angelo Cangelosi

Program chairs: Katharina Rohlfing, Ian Fasel

The past decade has seen the emergence of a new scientific field that studies how intelligent biological and artificial systems develop sensorimotor, cognitive and social abilities, over extended periods of time, through dynamic interactions of their brain and body with their physical and social environments. This field lies at the intersection of a number of scientific and engineering disciplines including Neuroscience, Developmental Psychology, Developmental Linguistics, Cognitive Science, Computational Neuroscience, Artificial Intelligence, Machine Learning, Robotics, and Philosophy. Various terms have been associated with this new field such as Autonomous Mental Development, Epigenetic Robotics, Developmental Robotics, etc., and several scientific meetings have been established. The two most prominent conference series of this field, the International Conference on Development and Learning (ICDL) and the International Conference on Epigenetic Robotics (EpiRob), are now joining forces and invite submissions for a joint meeting in 2011, to explore and extend the interdisciplinary boundaries of this field.

Keynote speakers:

Andy Barto, University of Massachussets Amherst, Amherst, USA.

Jean Mandler, University of California, San Diego , USA.

Erin Schuman, Howard Hugues Medical Institute, California Institute of Technology, USA.

Michael Tomasello, Max Planck Institute for Evolutionary Anthropology, Germany.

IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT

Volume 2, Issue 2, June 2010 **Link:** http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=5497035

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

Lopes, M.; Oudeyer, P.-Y. Page(s): 65-69 (**pdf**)

Abstract: Learning techniques are increasingly being used in todays' complex robotic system. Robots are expected to deal with a large variety of tasks, using their high-dimensional and complex bodies, to interact with objects and humans in an intuitive and friendly way. In this new setting, not all relevant information is available at design time, thus self-experimentation and learning by interacting with the physical and social world is very important to acquire knowledge and competences. Yet, the volume and dimensionality of sensorimotor spaces are often so large that active organized exploration strategies are necessary. We provide an overview of the articles published in this special issue, at the cross-roads of active learning and intrinsically motivated exploration. We also outline several important open challenges which remain to be addressed about exploration strategies for robot learning and development

Motivated Reinforcement Learning: An Evolutionary Perspective

Singh, S.; Lewis, R.L.; Barto, A.G.; Sorg, J. Kawamura, K.; Wilkes, D. M. Page(s): 70-82 (**pdf**)

Abstract: There is great interest in building intrinsic motivation into artificial systems using the reinforcement learning framework. Yet, what intrinsic motivation may mean computationally, and how it may differ from extrinsic motivation, remains a murky and controversial subject. In this paper, we adopt an evolutionary perspective and define a new optimal reward framework that captures the pressure to design good primary reward functions that lead to evolutionary success across environments. The results of two computational experiments show that optimal primary reward signals may yield both emergent intrinsic and extrinsic motivation. The evolutionary perspective and the associated optimal reward framework thus lead to the conclusion that there are no hard and fast features distinguishing intrinsic and extrinsic reward computationally. Rather, the directness of the relationship between rewarding behavior and evolutionary success varies along a continuum.

Genetic Programming for Reward Function Search

Niekum, S.; Barto, A.G.; Spector, L. Page(s): 83-90 (**pdf**)

Abstract: Reward functions in reinforcement learning have largely been assumed given as part of the problem being solved by the agent. However, the psychological notion of intrinsic motivation has recently inspired inquiry into whether there exist alternate reward functions that enable an agent to learn a task more easily than the natural task-based reward function allows. This paper presents a genetic programming algorithm to search for alternate reward functions that improve agent learning performance. We present experiments that show the superiority of these reward functions, demonstrate the possible scalability of our method, and define three classes of problems where reward function search might be particularly useful: distributions of environments, nonstationary environments, and problems with short agent lifetimes.

Correction to: What Is Needed for a Robot to Acquire Grammar? Some Underlying Primitive Mechanisms for the Synthesis of Linguistic Ability [Oct 09 187-195]

Lyon, C.; Sato, Y.; Saunders, J.; Nehaniv, C.L. Page(s): 144-144 (**pdf**)

Abstract: In the above titled paper (ibid., vol. 1, no. 3, pp. 187-195, Oct. 09), the acknowledgement to financial support was incompletely displayed. The correct acknowledgement is presented here.

Infomax Control of Eye Movements

Butko, N.J.; Movellan, J.R. Page(s): 91-107 (pdf)

Abstract: Recently, infomax methods of optimal control have begun to reshape how we think about active information gathering. We show how such methods can be used to formulate the problem of choosing where to look. We show how an optimal eye movement controller can be learned from subjective experiences of information gathering, and we explore in simulation properties of the optimal controller. This controller outperforms other eye movement strategies proposed in the literature. The learned eye movement strategies are tailored to the specific visual system of the learner-we show that agents with different kinds of eyes should follow different eye movement strategies. Then we use these insights to build an autonomous computer program that follows this approach and learns to search for faces in images faster than current state-of-the-art techniques. The context of these results is search in static scenes, but the approach extends easily, and gives further efficiency gains, to dynamic tracking tasks. A limitation of infomax methods is that they require probabilistic models of uncertainty of the sensory system, the motor system, and the external world. In the final section of this paper, we propose future avenues of research by which autonomous physical agents may use developmental experience to subjectively characterize the uncertainties they face.

Designing Interactions for Robot Active Learners

Cakmak, M.; Chao, C.; Thomaz, A.L. Page(s): 108-118 (pdf)

Abstract: This paper addresses some of the problems that arise when applying active learning to the context of human-robot interaction (HRI). Active learning is an attractive strategy for robot learners because it has the potential to improve the accuracy and the speed of learning, but it can cause issues from an interaction perspective. Here we present three interaction modes that enable a robot to use active learning queries. The three modes differ in when they make queries: the first makes a query every turn, the second makes a query only under certain conditions, and the third makes a query only when explicitly requested by the teacher. We conduct an experiment in which 24 human subjects teach concepts to our upper-torso humanoid robot, Simon, in each interaction mode, and we compare these modes against a baseline mode using only passive supervised learning. We report results from both a learning and an interaction perspective. The data show that the three modes using active learning are preferable to the mode using passive supervised learning both in terms of performance and human subject preference, but each mode has advantages and disadvantages. Based on our results, we lay out several guidelines that can inform the design of future robotic systems that use active learning in an HRI setting.

A Comparative Study of Value Systems for Self-Motivated Exploration and Learning by Robots

Merrick, K.E. Page(s): 119-131 (pdf)

Abstract: A range of different value systems have been proposed for self-motivated agents, including biologically and cognitively inspired approaches. Likewise, these value systems have been integrated with different behavioral systems including reflexive architectures, reward-based learning and supervised learning. However, there is little literature comparing the performance of different value systems for motivating exploration and learning by robots. This paper proposes a neural network architecture for integrating different value systems with reinforcement learning. It then presents an empirical evaluation and comparison of four value systems for motivating exploration by a Lego Mindstorms NXT robot. Results reveal the different exploratory properties of novelty-seeking motivation, interest and competence-seeking motivation.

Intrinsically Motivated Hierarchical Skill Learning in Structured Environments

Vigorito, C.M.; Barto, A.G. Page(s): 132-143 (pdf)

Abstract: We present a framework for intrinsically motivated developmental learning of abstract skill hierarchies by reinforcement learning agents in structured environments. Long-term learning of skill hierarchies can drastically improve an agent's efficiency in solving ensembles of related tasks in a complex domain. In structured domains composed of many features, understanding the causal relationships between actions and their effects on different features of the environment can greatly facilitate skill learning. Using Bayesian network structure (learning techniques and structured dynamic programming algorithms), we show that reinforcement learning agents can learn incrementally and autonomously both the causal structure of their environment and a hierarchy of skills that exploit this structure. Furthermore, we present a novel active learning scheme that employs intrinsic motivation to maximize the efficiency with which this structure is learned. As new structure is acquired using an agent's current set of skills, more complex skills are learned, which in turn allow the agent to discover more structure, and so on. This bootstrapping property makes our approach a developmental learning process that results in steadily increasing domain knowledge and behavioral complexity as an agent continues to explore its environment.

Volume 2, Issue 3, June 2010 **Link:** http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=5568771

Spatio-Temporal Multimodal Developmental Learning

Zhang, Y.; Weng, J. Page(s): 149-166 (**pdf**)

Abstract: It is elusive how the skull-enclosed brain enables spatio-temporal multimodal developmental learning. By multi-modal, we mean that the system has at least two sensory modalities, e.g., visual and auditory in our experiments. By spatio-temporal, we mean that the behavior from the system depends not only on the spatial pattern in the current sensory inputs, but also those of the recent past. Traditional machine learning requires humans to train every module using hand-transcribed data, using handcrafted symbols among modules, and hand-link modules internally. Such a system is limited by a static set of symbols and static module performance. A key characteristic of developmental learning is that the "brain" is "skull-closed" after birth-not directly manipulatable by the system designer-so that the system can continue to learn incrementally without the need for reprogramming. In this paper, we propose an architecture for multimodal developmental learning-parallel modality pathways all situate between a sensory end and the motor end. Motor signals are not only used as output behaviors, but also as part of input to all the related pathways. For example, the proposed developmental learning does not use silence as cut points for speech processing or motion static points as key frames for visual processing.

Integration of Action and Language Knowledge: A Roadmap for Developmental Robotics

Cangelosi, A.; Metta, G.; Sagerer, G.; Nolfi, S.; Nehaniv, C.; Fischer, K.; Tani, J.; Belpaeme, T.; Sandini, G.; Nori, F.; Fadiga, L.; Wrede, B.; Rohlfing, K.; Tuci, E.; Dautenhahn, K.; Saunders, J.; Zeschel, A. Page(s): 167-195 (**pdf**)

Abstract: This position paper proposes that the study of embodied cognitive agents, such as humanoid robots, can advance our understanding of the cognitive development of complex sensorimotor, linguistic, and social learning skills. This in turn will benefit the design of cognitive robots capable of learning to handle and manipulate objects and tools autonomously, to cooperate and communicate with other robots and humans, and to adapt their abilities to changing internal, environmental, and social conditions. Four key areas of research challenges are discussed, specifically for the issues related to the understanding of: 1) how agents learn and represent compositional actions; 2) how agents learn and represent compositional lexica; 3) the dynamics of social interaction and learning; and 4) how compositional action and language representations are integrated to bootstrap the cognitive system. The review of specific issues and progress in these areas is then translated into a practical roadmap based on a series of milestones. These milestones provide a possible set of cognitive robotics goals and test scenarios, thus acting as a research roadmap for future work on cognitive developmental robotics.

Top-Down Gaze Movement Control in Target Search Using Population Cell Coding of Visual Context

Miao, J.; Qing, L.; Zou, B.; Duan, L.; Gao, W. Page(s): 196-215 (**pdf**)

Abstract: Visual context plays an important role in humans' top-down gaze movement control for target searching. Exploring the mental development mechanism in terms of incremental visual context encoding by population cells is an interesting issue. This paper presents a biologically inspired computational model. The visual contextual cues were used in this model for top-down eye-motion control on searching targets in images. We proposed a population cell coding mechanism for visual context encoding and decoding. The model was implemented in a neural network system. A developmental learning mechanism was simulated in this system by dynamically generating new coding neurons to incrementally encode visual context during training. The encoded context was decoded with population neurons in a top-down mode. This allowed the model to control the gaze motion to the centers of the targets. The model was developed with pursuing low encoding quantity and high target locating accuracy. Its performance has been evaluated by a set of experiments to search different facial objects in a human face image set. Theoretical analysis and experimental results show that the proposed visual context encoding algorithm without weight updating is fast, efficient and stable, and the population-cell coding generally performs better than single-cell coding and k-nearest-neighbor (k-NN)-based coding.

Goal Babbling Permits Direct Learning of Inverse Kinematics

Rolf, M.; Steil, J.J.; Gienger, M. Page(s): 216-229 (**pdf**)

Abstract: We present an approach to learn inverse kinematics of redundant systems without prior- or expert-knowledge. The method allows for an iterative bootstrapping and refinement of the inverse kinematics estimate. The essential novelty lies in a path-based sampling approach: we generate training data along paths, which result from execution of the currently learned estimate along a desired path towards a goal. The information structure thereby induced enables an efficient detection and resolution of inconsistent samples solely from directly observable data. We derive and illustrate the exploration and learning process with a low-dimensional kinematic example that provides direct insight into the bootstrapping process. We further show that the method scales for high dimensional problems, such as the Honda humanoid robot or hyperredundant planar arms with up to 50 degrees of freedom.

Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010)

Schmidhuber, J. Page(s): 230-247 (**pdf**)

Abstract: The simple, but general formal theory of fun and intrinsic motivation and creativity (1990–2010) is based on the concept of maximizing intrinsic reward for the active creation or discovery of novel, surprising patterns allowing for improved prediction or data compression. It generalizes the traditional field of active learning, and is related to old, but less formal ideas in aesthetics theory and developmental psychology. It has been argued that the theory explains many essential aspects of intelligence including autonomous development, science, art, music, and humor. This overview first describes theoretically optimal (but not necessarily practical) ways of implementing the basic computational principles on exploratory, intrinsically motivated agents or robots, encouraging them to provoke event sequences exhibiting previously unknown, but learnable algorithmic regularities. Emphasis is put on the importance of limited computational resources for online prediction and compression. Discrete and continuous time formulations are given. Previous practical, but nonoptimal implementations (1991, 1995, and 1997–2002) are reviewed, as well as several recent variants by others (2005–2010). A simplified typology addresses current confusion concerning the precise nature of intrinsic motivation.

Top-Down Connections in Self-Organizing Hebbian Networks: Topographic Class Grouping

Luciw, M.; Weng, J. Page(s): 248-261 (**pdf**)

Abstract: We investigate the effects of top–down input connections from a later layer to an earlier layer in a biologically inspired network. The incremental learning method combines optimal Hebbian learning for stable feature extraction, competitive lateral inhibition for sparse coding, and neighborhood-based self-organization for topographic map generation. The computational studies reported indicate top–down connections encourage features that reduce uncertainty at the lower layer with respect to the features in the higher layer, enable relevant information to be uncovered at the lower layer so that irrelevant information can preferentially be discarded [a necessary property for autonomous mental development (AMD)], and cause topographic class grouping. Class groups have been observed in cortex, e.g., in the fusiform face area and parahippocampal place area. This paper presents the first computational account, as far as we know, explaining these three phenomena by a single biologically inspired network. Visual recognition experiments show that top–down-enabled networks reduce error rates for limited network sizes, show class grouping, and can refine lower layer representation after new conceptual information is learned. These findings may shed light on how the brain self-organizes cortical areas, and may contribute to computational understanding of how autonomous agents might build and maintain an organized internal representation over its lifetime of experiences.

Editor : Pierre-Yves Oudeyer Editorial Assistants : Adrien Baranes and Paul Spears
The AMD Newsletter is financially supported by the IEEE CIS and INRIA