IEEE CIS

CDS NEWSLETTER

The Newsletter of the Technical Committee on Cognitive and Developmental Systems

Developmental Robotics
Machine Intelligence
Neuroscience
Psychology

Volume 14, number 2
Fall 2017

Editorial

The Development of Human General Intelligence and Extreme Specialization

Pierre-Yves Oudeyer

Inria and Ensta ParisTech, France
Editor

pierre-yves.oudeyer@inria.fr

Humans have a unique capability to achieve and learn a wide diversity of skills of all kinds, from low-level sensorimotor skills to very abstract linguistic or mathematical skills. At the same time, they also show a striking capability to extreme specialization: particular individuals can become exceptional experts in particular skills. Is it possible to develop theories of how general cognitive architectures can display such a general flexibility for skill learning? And which aspects of cognitive mechanisms could explain extreme specialization?

The first question is addressed in this issue of the newsletter, through the dialog initiated by Matthias Rolf, Lorijn Zaadnoordijk and Johan Kwisthout, entitled "One developmental architecture to rule them all?". Responses from Niels Taatgen, John Spencer, Gary Jones, Gerard Wolff, Clément Moulin-Frier and Paul Verschure collectively discuss whether and how it would be useful both epistemologically and in practice to aim towards the development of a "standard integrated cognitive architecture", akin to "standard models" in physics, or whether focusing on simple and partial models should be a better approach. In particular, this question is discussed in the context of understanding development in infants, and of building developmental architectures, thus addressing the issue of architectures that not only learn, but that are adaptive themselves.

Then, a new dialog initiation is proposed by Celeste Kidd, exploring the question of why and how humans can be driven to extremely specialize. In particular, she proposes the hypothesis that curiosity may play a fundamental role in this process, and highlights many important open questions about how this could happen, and what are the actual mechanisms of curiosity-driven exploration and learning. Those of you interested in reacting to this dialog initiation are welcome to submit a response by May 30th, 2018. The length of each response must be between 600 and 800 words including references (contact pierre-yves.oudeyer@inria.fr).

Message From the CDS TC Chair

Kathryn Kasmarik

School of Engineering and Information Technology, University of New South Wales, Canberra, Australia

Chair of the Technical Committee on Cognitive and Developmental Systems

k.merrick@adfa.edu.au

I am honoured to be appointed to the role of Chair of the IEEE CIS Technical committee on Cognitive and Developmental Systems again for 2018. I look forward to continuing to work with you this year.

In 2018 the IEEE Computational Intelligence Society aims to celebrate and grow the diversity of its membership and their representation in scientific leadership. The strength of the scientific community lies in its recognition of creativity and excellent science, without ethnic, geographic, gender or other biases. As a technical committee we are asked to review our leadership to ensure a fair rotation of roles and unbiased distribution of leadership tasks among our members. This will be one of jobs I am committed to fulfilling this year.

From a research perspective, 2018 promises to further themes such as the study of ethics and social implications of computational intelligence, and it is timely to consider these in the context of developmental learning and cognitive systems. The implications for human machine interaction and application areas such as smart cities are wide ranging areas that can be influenced by our own research.

A number of upcoming events are planned for this year. The World Congress on Computational Intelligence (http://www.ecomp.poli.br/~wcci2018/) will take place in Rio de Janeiro, Brazil in July. In the second half of the year, we look forward to ICDL-Epirob (http://www.icdl-epirob.org/) in Tokyo, Japan in September. Members of the technical committee and task forces also have plans for workshops and special sessions at these and other international events. A Frontiers Research Topic organized by Vieri Giuliano Santucci, Pierre-Yves Oudeyer, Andrew Barto, Gianluca Baldassarre on Intrinsically Motivated Open-Ended Learning in Autonomous Robots is also due for publication later this year.

It is exciting to see the ongoing efforts of members of our community and 2018 promises to be a busy year for the technical committee.

Previous open-access editions of the newsletter can be found at: http://icdl-epirob.org/cdsnl
Web site of the IEEE TC on Cognitive and Developmental Systems: http://icdl-epirob.org/cdstc
IEEE ICDL-Epirob conference: http://www.icdl-epirob.org

Table of Contents

Dialogue

One Developmental Cognitive Architecture to Rule Them All?

Matthias Rolf

Dept. of Computing and Communication Technologies
Oxford Brookes University, UK
mrolf@brookes.ac.uk

Lorijn Zaadnoordijk

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, Netherlands
l.zaadnoordijk@donders.ru.nl

Johan Kwisthout

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, Netherlands
j.kwisthout@donders.ru.nl

A (cognitive) architecture describes the structure of an intelligent agent's mind, which may include emergent or even purely reactive approaches. Classical cognitive architectures typically describe grown behavioral or reasoning skills and are typically not embodied and structurally static, which makes their transfer to developmental problems problematic (Vernon et al., 2007). There have been some efforts to dedicatedly create architecture of learning and development, e.g. (Morse et al., 2010, Bellas et al., 2010). Many studies in developmental science describe or investigate the interplay of action and perception, motivation, and other aspects in closed and often embodied loops, thereby inevitably describing architectural aspects, even though not comprehensive ones that can span entire skill sets.

There clearly is not any cognitive architecture or general structural description that could "rule" developmental science (psychology/robotics), yet. The real questions of this dialogue initiation are therefore what purpose a single standard model and architecture could serve, and in how far the process of searching for one could be useful along the way.

What is the purpose of a developmental cognitive architecture?
The answer likely depends on whether one specifically looks at the scientific understanding of (human) intelligence, or at the engineering capability to build intelligence (that is, besides generally providing a potentially common language for researchers). Architectures potentially do more for science than "this is how it could work" descriptions. Unlike purely behavioral or descriptive models (e.g. sheer statistics of behavior), architectures describe hidden structure that is meant to explain the "how" and that might be experimentally verified. At the engineering end we might, in fact, find ourselves developing toolkit like solutions that also practically aid the creation of a developing intelligence.

Within either science or engineering, where would we find benefits from striving for unifying architectures?

Complexity monster or shackle?
Architectures naturally aim to address more than a single skill or a single scenario. If any single skill is investigated at a time (which is the practical norm), using a whole architecture involves complexity that is not strictly necessary for the task at hand, potentially violating Occam's razor. On the other hand, it has been argued that architectures actually constrain (instead of inducing unnecessary complexity) by confining models to a fixed formal language (Jones et al., 2000).

What areas or research could currently benefit from architectural efforts without being over-constraint by such a fixed language? How can practically good scientific experiments be conducted with such architectures?

Acknowledgements
We would like to thank all participants of the Lorentz-NIAS workshop "Perspectives on Developmental Robotics" (May 2017), whose discussions shaped this dialogue initiation.

Morse, A. F., De Greeff, J., Belpeame, T., & Cangelosi, A. (2010). Epigenetic robotics architecture (ERA). IEEE Transactions on Autonomous Mental Development, 2(4), 325-339
Vernon, D., Metta, G., & Sandini, G. (2007). A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. IEEE transactions on evolutionary computation, 11(2), 151-180.
Jones, G., Ritter, F. E., & Wood, D. J. (2000). Using a cognitive architecture to examine what develops. Psychological Science, 11(2), 93-100.
Bellas, F., Faiña, A., Varela, G., & Duro, R. J. (2010). A cognitive developmental robotics architecture for lifelong learning by evolution in real robots. In Neural Networks (IJCNN), The 2010 International Joint Conference on (pp. 1-8). IEEE.

Cognitive Architectures Should Not Be Computer Architectures

Niels Taatgen

Artificial Intelligence, University of Groningen, Netherlands

n.a.taatgen@rug.nl

The goal of cognitive architectures, as defined by Newell (1990) and Anderson (1983), is to provide a single theory that can explain and predict all aspects of the human mind. An architecture is not just a theory, it is also a simulation platform to build models of human task performance. A simulation platform means that certain representations have to be chosen and primitives have to be defined that serve as the building blocks of human cognition. Therefore, current cognitive architectures operate at a certain level of abstraction: Newell's Soar has chosen a purely symbolic level of abstraction, whereas Anderson's ACT-R has symbolic representations augmented with subsymbolic parameters. This choice of representational level is what Newell calls "carving nature at its joints". Despite the poetic nature of this claim, I would like to put forward the idea that this is a mistake.

Cognitive architectures provide the innate capabilities of the mind, which means that anything that is learned is not part of the architecture, even though the learning mechanisms themselves are. This means, assuming cognitive architectures operate on a certain level of abstraction, that anything below the level of abstraction of the architecture is implementation, and anything above that level has to be explained by knowledge that is accumulated through the learning mechanisms of the architecture. This reflects how architectures are designed in computer science, where each level of abstraction (e.g., logical circuits, microprogramming, machine language, higher level programming) is self-contained and virtually independent of constraints of the lower level.

Although this approach has been very successful in modeling many aspects of cognition, it fails if the phenomena that it wants to model are too far removed from the symbolic (or subsymbolic) level. For example, mechanisms for perceptual and motor learning are considered to be part of the implementation, and are therefore not covered. On the other hand, understanding natural language instruction requires so much knowledge and skills (all learned) that modeling that process in a cognitive architecture becomes programming in an awkward programming language. Instead, we have to acknowledge that learning and processing happens at many different levels of abstraction, and that we need a cognitive architecture with multiple levels of abstraction to be able to "rule them all": more similar to levels of abstraction in physics, chemistry and biology.

A multi-level cognitive architecture (Taatgen, 2017) provides representations and learning mechanisms for different levels of abstraction, from the neural level to higher-level reasoning. Each level has its own representations, which are composed of units from the level below. The composition or learning processes differs by level. At the lower levels of abstraction, where processing is typically fast, learning is often slow, for example attenuation of cells in the visual cortex to particular line orientations in the visual field. Learning at that level is a form of unsupervised learning. At the highest level of abstraction people can interpret natural language instructions for a new task, translate these into an instantiation of the necessary skills, and carry out that new task right away. Therefore, learning is fast ("one-shot") but processing is slow relative to processing at the lowest level.

In between we probably need several levels of abstraction: one in which we learn new skills to carry out tasks, such as the ability to count, or interpret language. Below that is traditional level of cognitive architectures, where units of representation are single "thinking steps" of in the order of 100ms each. Although there are many proposals for many levels of abstraction, tying them all together into a stable system will still be a big puzzle, which will hopefully not require a dark overlord.

Anderson, J. R. (1983). The architecture of cognition. book, Cambridge, MA: Harvard university press.
Newell, A. (1990). Unified theories of cognition. book, Cambridge, MA: Harvard university press.
Taatgen, N. A. (2017). Cognitive architectures: innate or learned? In The 2017 AAAI Fall Symposium Series (pp. 476–480). Palo Alto, CA: AAAI Press.

One (Dynamic Field) Theory to Rule Them All

John P. Spencer

School of Psychology, University of East Anglia, Norwich, UK

j.spencer@uea.ac.uk

Background

In a 2015 book, my colleagues and I proposed a large-scale neural process model of executive function (EF) and word learning (Schöner, Spencer, & The DFT Research Group, 2015). The model was constructed by integrating smaller models of attention and visual exploration (Perone & Spencer, 2012; Ross-Sheehy, Schneegans, & Spencer, 2015), visual working memory (Johnson, Spencer, Luck, & Schöner, 2009; Simmering, 2016), object representation and binding (Johnson, Spencer, & Schöner, 2008), and a variety of word learning phenomena (Samuelson, Smith, Perry, & Spencer, 2011). Thus, this line of work demonstrated the scalability of dynamic field theory—to move from smaller scale models and integrate them into a larger scale neural architecture that explains key aspects of higher-level cognition. And in each case—at small and large scales—the theory has led to testable predictions about how infants and children develop (Buss & Spencer, 2014; Perone, Molitor, Buss, Spencer, & Samuelson, 2015).

Here, I reflect on where this work might go: are we systematically moving toward one (dynamic field) theory to rule them all? If that is the aim, then what challenges do I see at present and what benefits might there be for pursuing this future? I discuss each in turn.

Challenges of large-scale modeling

• The integrated word learning/EF model is really complicated: The large scale architecture we are currently working with has 24 coupled cortical fields representing the neural activation of roughly 24,000 neurons (i.e., about 1000 neurons per population). How do we communicate the details of a model like this to a non-expert? Our first presentation of this model was in a book. That was reasonable because we could build the model across chapters. In a journal format, it is not possible to review all of the 'smaller' models that were integrated into the larger scale architecture in detail. This presents a unique communication challenge.

• Fitting the data must be done by hand. We have tried multiple model-fitting approaches. The computational load is too great—years of computation would be required to search the parameter space, even using the most advanced approaches. Indeed, we are not yet confident that Markov Chain Monte Carlo (MCMC) methods would even converge with our smaller scale models—we are exploring that now. This means that there is an entire literature in mathematical psychology dealing with model-fitting and 'free' parameters that we have to step outside of. It is not that we disagree with the arguments there—they just are not applicable to large scale neural models.

Benefits of large-scale modeling

• The integrated model should have massive generalizability. Our word learning model is the same as our executive function model—this is striking. Few models of EF even mention word learning; the fact that we have developed a large-scale model in two literatures in parallel suggests a deep theoretical link between these domains. Moreover, because our large-scale model has been built on the back of smaller architectures, this means we have integrated a large set of phenomena ranging from early attention to dimensional category labelling.

• The integrated model gives us a glimpse of neural reality. We have developed ways to generate hemodynamic predictions from dynamic field models (Buss, Wifall, Hazeltine, & Spencer, 2014). Here, the neural complexity of the model can be an advantage. We often simplify neural functions in smaller scale models (often to create fewer 'free' parameters). When we fit a brain model, however, those details are useful—each piece of the architecture hopefully leaves a neural signature that can be detected in the brain. Interestingly, when we use a general linear modelling approach (Wijeakumar, Ambrose, Spencer, & Curtu, 2017), we don't have to create separate regressors for each function or event in the task. Instead, the model specifies the entire neural pattern through time as the task unfolds. This approach could be very powerful.

• Large scale neural models might help us understanding developmental cascades. Embedded in our large scale model is a potential cascade of developmental changes moving from early perceptual and motor systems to attention and working memory to higher-level word learning and EF. Embedding these different functionalities into a single system allows us to look for signatures of cascade effects—as one system wires itself up, how does that affect the other systems to which it is coupled? This could shed new light on how development constructs itself step-by-step.

Conclusions: Are we hobbits or wizards?

• In my view, cognitive and developmental science have a Tolkien problem. As in Tolkien's trilogy, there is an intrinsic fear about 'one theory to rule them all'. Often, the idea is dismissed out of hand—the idea that someone might propose a large-scale model of the brain that is accurate is deemed to be a pipe dream. But if you dig deeper, I think there is something threatening about the idea. Let's say our model of word learning and executive function does a good job of explaining behavioural and neural data with kids and adults. This would be threatening in that the theory does something no other theory currently does—and if we are good scientists, that should matter! Now, everyone working in these domains should be required to learn about that theory. This places constraints on future work in this area. That's often uncomfortable for scientists—it is much easier to ignore the theory; to live the life of a hobbit.

• Other people might want to rise above—to become wizards and master this new theory. That takes time and energy, and should be encouraged. Physics has chosen this route. Does that mean there can only be a small handful of wizards? Here it depends on your definition of a wizard. There are certainly only a handful of theoreticians who understand Einstein, string theory, and the like. But there are wizards in experimental physics who are absolutely central to modern progress in physics. It is just that the wizarding world has bifurcated into theoretical physics and experimental physics. We think this is healthy.

• Thus, in the near future, we think cognitive and developmental science will have to decide whether to pursue the route of the hobbit or the wizard. It will soon become unreasonable to think that scientists could master empirical work with children, computational modelling, neuroscience, and robotics—something has to give. And we think that large scale models of brain function might be the tipping point that forces a sea change. So don't fear the ring. Ultimately, science is about discovering truth. What we do with that knowledge is a completely different dialog.

Buss, A. T., & Spencer, J. P. (2014). The emergent executive: a dynamic field theory of the development of executive function. Monographs of the Society for Research in Child Development. http://doi.org/10.1002/mono.12096

Buss, A. T., Wifall, T., Hazeltine, E., & Spencer, J. P. (2014). Integrating the behavioral and neural dynamics of reponse selection in a dual-task paradigm: A dynamic neural field model of Dux et al. (2009). Journal of Cognitive Neuroscience, 26(2), 334–351.

Johnson, J. S., Spencer, J. P., Luck, S. J., & Schöner, G. (2009). A dynamic neural field model of visual working memory and change detection. Psychological Science, 20(5), 568–577. http://doi.org/10.1111/j.1467-9280.2009.02329.x

Johnson, J. S., Spencer, J. P., & Schöner, G. (2008). Moving to a higher ground: the dynamic field theory and the dynamics of visual cognition. New Ideas in Psychology, 26, 227–251.

Perone, S., Molitor, S., Buss, A. T., Spencer, J. P., & Samuelson, L. K. (2015). Enhancing the executive functions of 3-year-olds in the dimensional change card sort task. Child Development, 86, 812–827.

Perone, S., & Spencer, J. P. (2012). Autonomy in action: Linking the act of looking to memory formation in infancy in infancy via dynamic neural fields. Cognitive Science, 1–59.

Ross-Sheehy, S., Schneegans, S., & Spencer, J. P. (2015). The Infant Orienting With Attention Task: Assessing the Neural Basis of Spatial Attention in Infancy. Infancy, 20(5), 467–506. http://doi.org/10.1111/infa.12087

Samuelson, L. K., Smith, L. B., Perry, L. K., & Spencer, J. P. (2011). Grounding word learning in space. PloS One, 6, E28095.

Schöner, G., Spencer, J. P., & The DFT Research Group. (2015). Dynamic Thinking: A Primer on Dynamic Field Theory. New York, NY: Oxford University Press.

Simmering, V. R. (2016). Working memory capacity in context: Modeling dynamic processes of behavior, memory, and development. Monographs of the Society for Research in Child Development, 81(3), 7–148.

Wijeakumar, S., Ambrose, J. P., Spencer, J. P., & Curtu, R. (2017). Model-based functional neuroimaging using dynamic neural fields: An integrative cognitive neuroscience approach. Journal of Mathematical Psychology, 76, 212–235. http://doi.org/10.1016/j.jmp.2016.11.002

Considering Simple Developmental Cognitive Architectures Before Complex Ones

Gary Jones

Department of Psychology, Nottingham Trent University, Nottingham, UK

gary.jones@ntu.ac.uk

What is the starting point of a developmental cognitive architecture? The minimum would presumably involve a mechanism that enabled one to learn things and another that constrained what could be learned. That way, we have something that is able to change its behaviour based on interacting with the world, while also "developing" i.e. not learning too much in a short space of time. Such a basic cognitive architecture is clearly not going to be able to explain the multi-faceted ways in which children develop so why not introduce complexity? At least two reasons spring to mind: first, how confident are we that the extra processes we include actually exist in the developing child? Second, hadn't we better find out how far we can get with the most basic architecture, because it might tell us which of the more complex processes aren't needed?

What can be taken as a given is that the child experiences the world and learns something from that experience. What does this experience involve? Unfortunately this is the magic question yet we are beginning to be able to answer it. Work by Linda Smith and colleagues (Jayaraman, Fausey & Smith, 2015; Yu & Smith, 2013) begins to illustrate infant experience of the visual world while linguistic experience in infancy can be estimated from large-scale transcripts and videos of mother-child interactions.

My own work involves the latter case of language where one estimates the child's experience based on the large-scale maternal speech that they hear. It would seem that gradual associative learning together with some constraint on the information processed from linguistic input can tell us a great deal about the language learning process and what may be involved. For example, measures of what can be referred to as verbal short-term memory appear to assess the child's current linguistic knowledge, be it sublexical information (where tests of nonword repetition can be simulated on the basis of the child's current linguistic knowledge, Jones, 2016) or lexical information (where tests of digit span can be explained from exposure to seemingly random digit sequences that appear in natural language, Jones & Macken, 2015). One even sees effects of syntax on the basis of associative learning (Kidd, 2012). Given that language is often perceived to be an indicator of intelligence, it would seem that one can get reasonably far in explaining linguistic phenomena purely from a straightforward associative learner and a good estimate of linguistic input.

Clearly, such a simplistic view is not going to account for all of child development, and that is where adding complexity comes in. But it is important to have a good idea of the child's experience of the particular tasks being examined because it seems that a lot of what appears as 'complexity' may reflect the environmental stimulus. It seems to me then that the role of the cognitive architecture is to encapsulate what the child learns from experience while also capturing higher-level cognition that can use learned knowledge to create new knowledge. In addition, this needs to go across numerous domains. That, I believe, is the challenge for a developmental cognitive architecture and whether we are yet at the stage where one can achieve this appears debatable.

Jayaraman, S., Fausey, C. M., & Smith, L. B. (2015). The faces in infant-perspective scenes change over the first year of life. PLoS ONE, 10: e0123780.
Jones, G. (2016). The influence of children's exposure to language from two to six years: The case of nonword repetition. Cognition, 153, 79-88.
Jones, G. & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. Cognition, 144, 1-13.
Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. Developmental Psychology, 48, 171.
Yu, C. & Smith, L. B. (2013). Joint attention without gaze following: Human infants and their parents coordinate visual attention to objects through eye- hand coordination. PLoS ONE, 8: e79659.

Yes, One Developmental Cognitive Architecture Is Necessary and Feasible

J. Gerard Wolff

CognitionResearch.org,
Menai Bridge, UK

jgw@cognitionresearch.org

In their dialog initiation, Matthias Rolf, Lorijn Zaadnoordijk, and Johan Kwisthout ask:

> '... whether and how it would be useful both epistemologically and in practice to aim towards the development of a "standard integrated cognitive architecture", akin to "standard models" in physics. In particular, [the authors] ask this question in the context of understanding development in infants, and of building developmental architectures, thus addressing the issue of architectures that not only learn, but that are adaptive themselves.'

In brief, my answer is "yes", it is essential to aim for such an architecture, and to abandon the quest only when there is overwhelming evidence that it cannot be done. The main reasons are: 1) That, while the gathering of empirical evidence is an important part of any science, it is at least as important to try to develop parsimonious theories to make sense of empirical evidence; 2) In a 20-year programme of research, derived from earlier research on language learning (Wolff, 1988), I have developed a cognitive architecture which already has a lot to say about the nature of cognition, including learning, adaptation, and cognitive development. This research demonstrates what can be achieved, suggesting that it will indeed be possible to develop a "standard integrated cognitive architecture".

The SP theory of intelligence and its realisation in the SP computer model is a unique attempt to simplify and integrate observations and concepts across artificial intelligence, mainstream computing, mathematics, and human learning, perception, and cognition, with information compression as a unifying theme (See Wolff, (2006, 2013, 2016) and other papers on www.cognitionresearch.org/sp.htm).

A central idea in the SP system is the powerful concept of SP-multiple-alignment, borrowed and adapted from bioinformatics. This yields:

• Versatility in aspects of intelligence. The SP system has strengths in several aspects of intelligence including: 'unsupervised' learning (which has the potential to be the foundation of other kinds of learning); the analysis and production of natural language; pattern recognition that is robust in the face of errors in data; pattern recognition at multiple levels of abstraction; computer vision; best-match and semantic kinds of information retrieval; planning; problem solving; and:

– Versatility in the representation of knowledge. The SP system has strengths in the representation of diverse kinds of knowledge including: the syntax of natural languages; class-inclusion hierarchies (with or without cross classification); part-whole hierarchies; discrimination networks and trees; if-then rules; entity-relationship structures; relational tuples; and concepts in mathematics, logic, and computing.

– Versatility in reasoning. Strengths of the SP system in reasoning include: one-step 'deductive' reasoning; chains of reasoning; abductive reasoning; reasoning with probabilistic networks and trees; reasoning with 'rules'; nonmonotonic reasoning and reasoning with default values; Bayesian reasoning with 'explaining away'; causal reasoning; reasoning that is not supported by evidence; the inheritance of attributes in class hierarchies; and inheritance of contexts in part-whole hierarchies. There is also potential for spatial reasoning, and for what-if reasoning.

• Seamless integration of diverse kinds of knowledge and diverse aspects of intelligence. Because diverse kinds of knowledge and diverse aspects of intelligence all flow from a single coherent and relatively simple source—the SP-multiple-alignment framework—there is clear potential for the SP system to provide seamless integration of diverse kinds of knowledge and diverse aspects of intelligence, in any combination. It appears that that kind of seamless integration is essential for human levels of fluidity, versatility and adaptability in intelligence.

There is more detail in Wolff (2017, Appendix B).

J. G. Wolff. Learning syntax and meanings through optimization and distributional analysis. In Y. Levy, I. M. Schlesinger, and M. D. S. Braine, editors, Categories and Processes in Language Acquisition, pages 179–215. Lawrence Erlbaum, Hillsdale, NJ, 1988.
J. G. Wolff. Unifying Computing and Cognition: the SP Theory and Its Applications. CognitionResearch.org, Menai Bridge, 2006. ISBNs: 0- 9550726-0-3 (ebook edition), 0-9550726-1-1 (print edition).
J. G. Wolff. The SP theory of intelligence: an overview. Information, 4(3):283–341, 2013. arXiv:1306.3888.
J. G. Wolff. The SP theory of intelligence: its distinctive features and advantages. IEEE Access, 4:216–246, 2016. arXiv:1508.04087.
J. G. Wolff. Software engineering and the SP theory of intelligence. Tech- nical report, CognitionResearch.org, 2017. Submitted for publication. arXiv:1708.06665.

Clément Moulin-Frier
Synthetic Perceptive, Emotive and Cognitive Systems (SPECS), Universitat Pompeu Fabra, Barcelona, Spain
clement.moulinfrier@gmail.com

Paul F.M.J. Verschure
Director of SPECS Lab (IBEC, BIST, ICREA), Barcelona, Spain
pverschure@ibecbarcelona.eu

Distributed Adaptive Control As an Integration Framework for Cognition

As pointed out by Matthias Rolf, Lorijin Zaadnoordijk and Johan Kwisthout, there is a gap between the description of the structure of an intelligent agent's mind, which may include emergent or even purely reactive approaches and classical cognitive architectures that describe advanced behavioral or reasoning skills. The latter are typically not embodied and structurally agnostic. This contrast reflects an old debate in Cognitive Science and Artificial Intelligence, where two opposing approaches have been advanced to explain how cognitive functions can arise. Top-down approaches rely on a priori symbolic representations of a task, which have to be recursively decomposed into simpler ones to be executed by the agent. These depend principally on methods from symbolic artificial intelligence, as, e.g., in Soar (Laird, Newell, & Rosenbloom, 1987). The alternative, bottom-up approaches instead implement behavior without relying on advanced knowledge representation and reasoning. This is typically the case in behavior-based robotics, where low-level sensory-motor control loops form the starting point of emergent behavioral complexity as Simon's "ant on the beach" example (Simon, 1969), Braitenberg's Vehicles (Braitenberg, 1986) and implemented in the Subsumption architecture (Brooks, 1986).

Interestingly, a machine-learning-oriented version of this old debate has emerged from recent advances in Artificial Intelligence. On the one hand, strong emphasis is placed on so-called Deep Learning frameworks, where large feed-forward networks are trained end to end with an extremely large amount of training data. On the other hand, a drastically different approach has also received considerable attention, arguing that Deep Learning is not able to solve key aspects of human cognition without having access to advanced prior knowledge (Lake, Ullman, Tenenbaum, & Gershman, 2017). This approach states that human cognition relies on causal models of the world built through combinatorial processes to rapidly acquire knowledge and generalize it to new tasks and situations. This solution, however, comes at a cost: the underlying algorithms require non-trivial a priori knowledge, and an assumption of such models is that learning should be grounded in intuitive theories of physics and psychology.

This illustrates that despite the recent advances, we still face the old debate between bottom-up and top-down models of cognition. It is, therefore, a major challenge to structure these heterogeneous aspects of cognition in one unified theory. The Distributed Adaptive Control (DAC) theory of the mind and brain (see Verschure, 2016, for a recent review) provides a principled framework for realizing this structuring and integration effort by grounding it into biology, neuroscience, and ecology. DAC proposes that cognition is based on the interaction of four interconnected control layers operating at different levels of abstraction (see Figure 1). The first level, the somatic layer, corresponds to the embodiment of the agent within its environment, with its sensors and actuators as well as the physiological needs (e.g. exploration or safety). Extending bottom-up approaches with drive reduction mechanisms, complex behavior is bootstrapped in DAC from the self-regulation of an agent's physiological needs when combined with reactive behaviors (the reactive layer). This reactive interaction with the environment drives the joint acquisition of both perceptual and action hierarchical representations modulated by value signals in the adaptive layer. These compressed (hierarchical) and informative (modulated by value) representations support the acquisition of causal models of the world for goal creation and planning at the fourth contextual layer, which comprises systems for episodic, procedural and working memory and an autobiographical memory supporting life-long learning. These high-level processes, in turn, modulate the activity at the lower levels via top-down pathways shaped by behavioral feedback, i.e. acting through the environment itself rather than through direct internal control signals. The control flow in DAC is therefore distributed, both from bottom-up and top-down interactions between layers, as well as from lateral information processing within each layer.

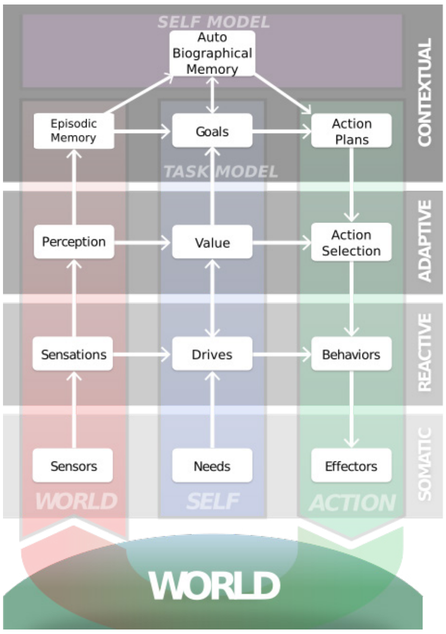

Figure 1: Abstract representation of The DAC architecture, see text and (Verschure, 2016) for details.

Braitenberg, V. (1986). Vehicles. Cambridge, MA, USA: MIT Press.
Brooks, R. (1986). A robust layered control system for a mobile robot. IEEE J. Robotics and Automation, RA-2, 14–23.
Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. Artificial Intelligence. https://doi.org/10.1016/0004-3702(87)90050-6
Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building Machines That Learn and Think Like People. Behavioral and Brain Sciences. https://doi.org/10.1017/S0140525X16001837
Moulin-Frier, C., Puigbò, J.-Y., Arsiwalla, X. D., Sanchez-Fibla, M., & Verschure, P. F. M. J. (2017). Embodied Artificial Intelligence through Distributed Adaptive Control: An Integrated Framework.
Simon, H. A. (1969). Sciences of the artificial. Cambridge: Cambridge, Ma.: MIT Press.
Verschure, P. F. M. J. (2016). Synthetic consciousness: the distributed adaptive control perspective. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 371(1701), 263–275. https://doi.org/10.1098/rstb.2015.0448
Verschure, P. F. M. J. P., & Althaus, P. (2003). A real-world rational agent: unifying old and new AI. Cognitive Science, 27(4), 561–590. Retrieved from http://dx.doi.org/10.1016/S0364-0213(03)00034-X

One Developmental Cognitive Architecture to Rule Them All? Responses to Commentaries

Matthias Rolf

Dept. of Computing and Communication Technologies
Oxford Brookes University, UK
mrolf@brookes.ac.uk

Lorijn Zaadnoordijk

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, Netherlands
l.zaadnoordijk@donders.ru.nl

Johan Kwisthout

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen, Netherlands
j.kwisthout@donders.ru.nl

During the NIAS-Lorentz workshop on 'perspectives on developmental robotics', that we organized in May 2017, the participants identified the need for a *developmental* cognitive architecture as a means for modeling development. In our dialogue initiation we focused on two questions: What is the research goal of a developmental cognitive architecture? And is the complexity inherent in a cognitive architecture a hindrance or a blessing?

Jones responds negatively to our question whether it is feasible and timely to focus on developmental cognitive architectures. He suggests that it might be too early to aim for such a general architecture, and that it is probably better to focus first on simple (associative) models before postulating more elaborate mechanisms and processes, as we yet do not really understand what actually constitutes the experiences that infants use in their learning. *Spencer*, in contrast, is more positive towards a general architecture; his approach is to focus on the low-level neural architecture (dynamic field theory) and scale up to simulations with thousands of neurons. Here, explaining and fitting the model is a challenge, but might hint at the neural reality of developmental cascades.

Moulin-Frier and Verschure reflect on the historical debate between symbolic and connectionist AI, that recently re-emerged in the form of structure-rich (Bayesian) causal models on the one hand and model-free (deep-learning) approaches on the other hand. Their DAC (Distributed Adaptive Control) approach aims to connect higher cognitive causal models with somative, reactive, and adaptive layers. *Wolff* proposes the SP theory of intelligence (and corresponding computer model) which is rooted in information compression. It still remains to be seen how either architectural approaches allow for studying development, i.e., the effect of physical change on cognition. *Taatgen*, in addition, points at the difficult problem of uniting in a single multi-level architecture both low-level learning (often slow and model free) and higher-level (instructed and/or planned) learning that is relative fast.

Our preliminary conclusion of this discussion is that there is (obviously) no 'simple solution' towards the developmental problem, where 'development' transcends 'learning' in the sense that the architectural cognitive features *themselves* are changing during development. Important questions are still open: For example whether we can define development in and through architectures, and whether the architectural level is necessary to define "development", as opposed to "learning". A dedicated workshop on these questions—for example at a future ICDL/EpiRob conference—might be timely to address these vital research questions.

New Dialogue Initiation

Curiosity as Driver of Extreme Specialization in Humans

Celeste Kidd

Assistant Professor of Brain and Cognitive Sciences, University of Rochester

celestekidd@gmail.com

The features that make us uniquely and distinctly human have been of interest to many people, from psychologists to philosophers to religious scholars, for centuries. Typical candidate traits include things like speech (Lieberman, 1991), upright posture (Clarke & Tobias, 1995), protracted childhoods (Jolly, 1972), helpless infants (Piantadosi & Kidd, 2016), sophisticated social cooperation (Melis & Semmann, 2010), and creativity (Carruthers, 2002).

There is, however, an essential human trait that has received far less recognition: the capacity for extreme specialization. Many humans spend a lifetime perfecting a single niche skill, such as a musical instrument, art medium, or style of dance. Others specialize in trades with economic roles (e.g., butchers, bakers, and candlestick makers). And while some other species exhibit certain forms of specialization—ants, for example, exhibit increased task specialization as the colony size increases (Amador-Vargas et al., 2015)—none approach the breadth and depth of specialization found in humans. In particular, specialization in species usually seems to hinge on abilities that are directly relevant to survival. Human specialization, in contrast, knows no limits or bounds and seems applicable to virtually any domain of existence. Here we will argue that this extreme specialization is enabled in large part due to key mechanisms within the human attentional system—specifically those mechanisms that bias learners towards material for which they already possess some background knowledge. More broadly, this extreme specialization is enabled by the driving pressures that underlie human curiosity.

Curiosity can be thought of as the force behind the acquisition of new knowledge (James, 1913; Pavlov, 1927; Skinner, 1938; Oudeyer & Kaplan, 2007; Gottlieb et al., 2013). It is a strong determinant of how we spend our days, and influences not just our intellectual interests, but also a myriad of recreational decisions, from who we speak to and what we discuss, to what we listen to and watch, to what we fixate on in a scene and what we learn about the world. It is a key driving force behind the grandest human innovations, yet less sophisticated, purpose-specific forms of curiosity for can be observed in more primitive intelligences (e.g., C. elegans). Curiosity, or intrinsic motivation, is likely a necessary feature of intelligent systems generally. Even robotic and artificial intelligence systems must possess a mechanism to seek out and learn material that is relevant to their present and future goals (Oudeyer & Kaplan, 2007).

Human curiosity is known to relate to our existing knowledge. For example, work from the infant attention literature suggests that infants prefer novel stimuli, defined as distinct from what the infant already knows (Sokolov, 1963) or partially encoded representations over either entirely known or entirely novel ones (Dember & Earl, 1957; Kinney & Kagan, 1976; Berlyne, 1978; Kidd et al., 2013). More contemporary theories observe that curiosity is triggered when a gap is detected between what a learner currently knows, and what they could know (Loewenstein, 1994). This suggests the involvement of metacognition, since a learner must first identify that there is a gap to be filled before curiosity should be piqued. Yet little work to date has explored the relationship between metacognitive processes and curiosity. Are people who possess more metacognitive abilities pertaining to their own knowledge more curious? Can you make someone more curious by calling attention to what they do not know?

While we know that there exists some relationship between existing knowledge about a stimulus and the learner's degree of interest in that stimulus, we still do not fully understand precisely how those two factors relate to each other, nor do we understand the cognitive or neural mechanisms underlying how and why the learner's curiosity is piqued (for a review of what we don't know, see Hayden & Kidd, 2015). For example, we do not understand how neural reward systems treat information and weigh it in decision-making, though it is clear that humans and monkeys are willing to sacrifice some reward to gain even useless information (Blanchard, Hayden, & Bromberg-Martin, 2015). Is there a common currency for reward and information, and how is the value of information determined, represented, and integrated neurally?

We have limited evidence to suggest that being in a curious state could facilitate learning (Gruber et al., 2014; Stahl & Feigenson, 2015); however, we also have evidence that learners are more curious when they possess information that is partially encoded, and thus on the verge of being learned (Kang et al., 2009). Thus, we must be sensitive to the fact that some of the apparent boosts to learning attributed to curiosity in the literature may have the direction of causality wrong—being on the verge-of-learning may induce greater curiosity, rather than curiosity inducing better learning. How do we understand curiosity and the biological mechanisms underlying it in a way that reasonably accounts for these two apparently opposing causal mechanisms?

What is the purpose of this curiosity system, and why does it yield the sort of specialization that we see in humans but not other species? How does it function, and what purposes does it serve? Why are there humans that become compelled to acquire information about fictitious worlds (e.g., Harry Potter, Star Wars)? What might be the connection between curiosity, creativity, and specialization?

Amador-Vargas, S., Gronenberg, W., Wcislo, W. T., & Mueller, U. (2015). Specialization and Group Size: Brain and behavioural correlates of colony size in ants lacking morphological castes. Proceedings of the Royal Society B, 282(1801), 20142502.
Berlyne, D. E. (1978). Curiosity and learning. Motivation and Emotion, 2(2), 97-175.
Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. Neuron, 85(3), 602-614.
Carruthers, P. (2002). The cognitive functions of language. Behavioral and Brain Sciences, 25(6), 657-674.
Clarke, R. J., & Tobias, P. V. (1995). Sterkfontein Member 2 foot bones of the oldest South African hominid. Science, 269(5223), 521.
Dember, W. N., & Earl, R. W. (1957). Analysis of exploratory, manipulatory, and curiosity behaviors. Psychological Review, 64(2), 91.
Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, Curiosity, and Attention: Computational and Neural Mechanisms. Trends in Cognitive Sciences, 17(11), 585-593.
Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. Neuron, 84(2), 486-496.
James, W. (1913). The Principles of Psychology, Volume II. New York, NY, United States: Henry Holt and Company.
Jolly, A. (1972). The Evolution of Primate Behavior. New York, NY, United States: MacMillan.
Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T. Y., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. Psychological Science, 20(8), 963-973.
Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. Neuron, 88(3), 449-460.
Kidd, C., Piantadosi, S.T., & Aslin, R.N. (2012.) The Goldilocks Effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. PLOS ONE, 7(5), e36399.
Kinney, D. K., & Kagan, J. (1976). Infant attention to auditory discrepancy. Child Development, 155-164.
Lieberman, P. (1991). Uniquely Human: The evolution of speech, thought, and selfless behavior. Cambridge, MA, United States: Harvard University Press.
Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. Psychological Bulletin, 116(1), 75.
Melis, A. P., & Semmann, D. (2010). How is human cooperation different?. Philosophical Transactions of the Royal Society of London B: Biological Sciences, 365(1553), 2663-2674.
Oudeyer, P. Y., & Kaplan, F. (2007). What is intrinsic motivation? A typology of computational approaches. Frontiers in Neurorobotics, 1, 6.
Pavlov, I.P. (1927). Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex. Oxford, England: Oxford University Press.
Piantadosi, S. T., & Kidd, C. (2016). Extraordinary intelligence and the care of infants. Proceedings of the National Academy of Sciences, 113(25), 6874-6879.
Stahl, A. E., & Feigenson, L. (2015). Observing the unexpected enhances infants' learning and exploration. Science, 348(6230), 91-94.
Skinner, B. F. (1938). The Behavior of Organisms: An Experimental Analysis. Oxford, England: Appleton–Century.
Sokolov, E. (1963). Perception and the Conditioned Reflex. Oxford, England: Pergamon Press.

IEEE TCDS Table of Contents

Volume 9, Issue 3, September 2017

SNet: Co-Developing Artificial Retinas and Predictive Internal Models for Real Robots

Ricardo Santos, Ricardo Ferreira, Ângelo Cardoso, Alexandre Bernardino

This paper focuses on a recently developed biologically inspired architecture, here denoted as sensorimotor network (SNet), able to co-develop sensorimotor structures directly from data acquired by a robot interacting with its environment. Such networks learn efficient internal models of the sensorimotor system, developing simultaneously sensor and motor representations as well as predictive models of the sensorimotor relationships adapted to their operating environment. Here, we describe our recent model of sensorimotor development and compare its performance with neural network models in predicting self-induced stimuli. In addition, we illustrate the influence of available resources and environment characteristics in the development of the SNet structures. Finally, an SNet is trained using real data recorded during a quadricopter drone flight.

Multilevel Behavioral Synchronization in a Joint Tower-Building Task

Moreno I. Coco, Leonardo Badino, Pietro Cipresso, Alice Chirico, Elisabetta Ferrari, Giuseppe Riva, Andrea Gaggioli, Alessandro D'Ausilio

Human to human sensorimotor interaction can only be fully understood by modeling the patterns of bodily synchronization and reconstructing the underlying mechanisms of optimal cooperation. We designed a tower-building task to address such a goal. We recorded upper body kinematics of dyads and focused on the velocity profiles of the head and wrist. We applied recurrence quantification analysis to examine the dynamics of synchronization within, and across the experimental trials, to compare the roles of leader and follower. Our results show that the leader was more auto-recurrent than the follower to make his/her behavior more predictable. When looking at the cross-recurrence of the dyad, we find different patterns of synchronization for head and wrist motion. On the wrist, dyads synchronized at short lags, and such a pattern was weakly modulated within trials, and invariant across them. Head motion, instead, synchronized at longer lags and increased both within and between trials: a phenomenon mostly driven by the leader. Our findings point at a multilevel nature of human to human sensorimotor synchronization, and may provide an experimentally solid benchmark to identify the basic primitives of motion, which maximize behavioral coupling between humans and artificial agents.

Dogs as Behavior Models for Companion Robots: How Can Human–Dog Interactions Assist Social Robotics?

Gabriella Lakatos

This position paper (re)presents a relatively new approach for the behavioral design of companion robots, the use of dogs' behavior as a model. This paper discusses the advantages of this approach compared to other prevalent approaches in the field of social robotics and analyzes its effectiveness through the review of three different experimental studies utilizing this concept.

Bio-Inspired Embedded Vision System for Autonomous Micro-Robots: The LGMD Case

Cheng Hu, Farshad Arvin, Caihua Xiong, Shigang Yue

In this paper, we present a new bio-inspired vision system embedded for micro-robots. The vision system takes inspiration from locusts in detecting fast approaching objects. Neurophysiological research suggested that locusts use a wide-field visual neuron called lobula giant movement detector (LGMD) to respond to imminent collisions. In this paper, we present the implementation of the selected neuron model by a low-cost ARM processor as part of a composite vision module. As the first embedded LGMD vision module fits to a micro-robot, the developed system performs all image acquisition and processing independently. The vision module is placed on top of a micro-robot to initiate obstacle avoidance behavior autonomously. Both simulation and real-world experiments were carried out to test the reliability and robustness of the vision system. The results of the experiments with different scenarios demonstrated the potential of the bio-inspired vision system as a low-cost embedded module for autonomous robots.

Online Algorithm for Robots to Learn Object Concepts and Language Model

Joe Nishihara ; Tomoaki Nakamura ; Takayuki Nagai

Humans form concept of objects by classifying them into categories, and acquire language by simultaneously interacting with others. Thus, the meaning of a word can be learned by connecting a recognized word to its corresponding concept. We consider this ability important for robots to flexibly develop knowledge of language and concepts. In this paper, we propose an online algorithm for robots to acquire knowledge of natural language and learn object concepts. A robot learns the language model from word sequences, which are obtained by the segmentation of phoneme sequences provided by a user, by using unsupervised word segmentation each time it is provided with a new object. Moreover, the robot acquires object concepts using these word sequences as well as multimodal information obtained by observing objects. The crucial aspect of our model is the interdependence of words and concepts: there is a high probability that the same words will be uttered to describe objects in the same category. By taking this relationship into account, our proposed method enables robots to acquire a more accurate language model and object concepts online. Experimental results verify this.

A Partial Contour Similarity-Based Approach to Visual Affordances in Habile Agents

Thomas E. Horton ; Robert St. Amant

In a typical tool use task, we can view both the relationship between the agent and the tool and the relationship between the tool and the target in terms of affordances. One set of affordances relates to the ability of the agent to manipulate the tool, while a second set of affordances relates to the ability of the agent to manipulate the target by means of the tool. In both cases, effective tool use is facilitated by the coupling of one object to another: agent-to-tool-to-target. In this paper, we focus on the visual identification of such affordances via contour similarity. Objects with complementary contour segments can fit together, which suggests possible opportunities for effective interactions. We present a system for the identification and evaluation of partial contour-based matches and analyze the system's behavior. We propose a set of sample tool-use scenarios as part of our analysis. We demonstrate the use of the system in providing guidance to an autonomous robotic agent performing tool selection tasks.

Multichannel EEG-Based Emotion Recognition via Group Sparse Canonical Correlation Analysis

Wenming Zhengi

In this paper, a novel group sparse canonical correlation analysis (GSCCA) method is proposed for simultaneous electroencephalogram (EEG) channel selection and emotion recognition. GSCCA is a group sparse extension of the conventional CCA method to model the linear correlationship between emotional EEG class label vectors and the corresponding EEG feature vectors. In contrast to conventional CCA method or previous GSCCA methods, a major advantage of our GSCCA method is the ability of handling the group feature selection problem from raw EEG features, which makes it very suitable for simultaneously coping with both EEG emotion recognition and automatic channel selection issues where each EEG channel is associated with a group of raw EEG features. To deal with EEG emotion recognition problem, we adopt the popularly used frequency feature to describe the EEG signal by dividing the full EEG frequency band into five parts, i.e., delta, theta, alpha, beta, and gamma frequency bands, and then extract the frequency band features from each band for GSCCA model learning and emotion recognition. Finally, we conduct extensive experiments on EEG-based emotion recognition based on the SJTU emotion EEG dataset and experimental results demonstrate that the proposed GSCCA method would outperform the state-of-the-art EEG-based emotion recognition approaches.

Volume 9, Issue 4, December 2017

New Algorithms for Encoding, Learning and Classification of fMRI Data in a Spiking Neural Network Architecture: A Case on Modeling and Understanding of Dynamic Cognitive Processes

Nikola Kasabov, Lei Zhou, Maryam Gholami Doborjeh, Zohreh Gholami Doborjeh, Jie Yang

This paper argues that, the third generation of neural networks—the spiking neural networks (SNNs), can be used to model dynamic, spatio-temporal, cognitive brain processes measured as functional magnetic resonance imaging (fMRI) data. This paper proposes a novel method based on the NeuCube SNN architecture for which the following new algorithms are introduced: fMRI data encoding into spike sequences; deep unsupervised learning of fMRI data in a 3-D SNN reservoir; classification of cognitive states; and connectivity visualization and analysis for the purpose of understanding cognitive dynamics. The method is illustrated on two case studies of cognitive data modeling from a benchmark fMRI data set of seeing a picture versus reading a sentence.

Place Classification With a Graph Regularized Deep Neural Network

Yiyi Liao, Sarath Kodagoda, Yue Wang, Lei Shi, Yong Liu

Place classification is a fundamental ability that a robot should possess to carry out effective human-robot interactions. In recent years, there is a high exploitation of artificial intelligence algorithms in robotics applications. Inspired by the recent successes of deep learning methods, we propose an end-to-end learning approach for the place classification problem. With deep architectures, this methodology automatically discovers features and contributes in general to higher classification accuracies. The pipeline of our approach is composed of three parts. First, we construct multiple layers of laser range data to represent the environment information in different levels of granularity. Second, each layer of data are fed into a deep neural network for classification, where a graph regularization is imposed to the deep architecture for keeping local consistency between adjacent samples. Finally, the predicted labels obtained from all layers are fused based on confidence trees to maximize the overall confidence. Experimental results validate the effectiveness of our end-to-end place classification framework in which both the multilayer structure and the graph regularization promote the classification performance. Furthermore, results show that the features automatically learned from the raw input range data can achieve competitive results to the features constructed based on statistical and geometrical information.

Online Multimodal Ensemble Learning Using Self-Learned Sensorimotor Representations

Martina Zambelli, Yiannis Demiris

Internal models play a key role in cognitive agents by providing on the one hand predictions of sensory consequences of motor commands (forward models), and on the other hand inverse mappings (inverse models) to realize tasks involving control loops, such as imitation tasks. The ability to predict and generate new actions in continuously evolving environments intrinsically requiring the use of different sensory modalities is particularly relevant for autonomous robots, which must also be able to adapt their models online. We present a learning architecture based on self-learned multimodal sensorimotor representations. To attain accurate forward models, we propose an online heterogeneous ensemble learning method that allows us to improve the prediction accuracy by leveraging differences of multiple diverse predictors. We further propose a method to learn inverse models on-the-fly to equip a robot with multimodal learning skills to perform imitation tasks using multiple sensory modalities. We have evaluated the proposed methods on an iCub humanoid robot. Since no assumptions are made on the robot kinematic/dynamic structure, the method can be applied to different robotic platforms.

Human Activity Recognition Based on Spatial Distribution of Gradients at Sublevels of Average Energy Silhouette Images

Dinesh Kumar Vishwakarma, Kuldeep Singh

The aim of this paper is to present a unified framework for human action and activity recognition by analysing the effect of computation of spatial distribution of gradients (SDGs) on average energy silhouette images (AESIs). Based on the analysis of SDGs computation at various decomposition levels, an effective approach to compute the SDGs is developed. The AESI is constructed for the representation of the shape of action and activity and these are the reflection of 3-D pose into 2-D pose. To describe the AESIs, the SDGs at various sublevels and sum of the directional pixels (SDPs) variations is computed. The temporal content of the activity is computed through R-transform (RT). Finally, the shape computed through SDGs and SDPs, and temporal evidences through RT of the human body is fused together at the recognition stage, which results in a new powerful unified feature map model. The performance of the proposed framework is evaluated on three different publicly available datasets, i.e., Weizmann, KTH, and Ballet and the recognition accuracy is computed using hybrid classifier. The highest recognition accuracy achieved on these datasets is compared with the similar state-of-the-art techniques and demonstrate the superior performance.

Emergent Structuring of Interdependent Affordance Learning Tasks Using Intrinsic Motivation and Empirical Feature Selection

Emre Ugur, Justus Piater

This paper studies mechanisms that produce hierarchical structuring of affordance learning tasks of different levels of complexity. Guided by intrinsic motivation, our system detects easy tasks first, and learns them in selected environments which are maximally different from the previously encountered ones. Easy tasks are learned from observed low-level attributes of the environment, and provide abstractions over these attributes. As learning progresses, the system shifts its focus and starts learning harder tasks not only from low-level attributes but also from previously-learned abstract concepts. Therefore, hard tasks are autonomously placed higher in the hierarchy if the easy task concepts are identified as distinctive input attributes of hard tasks. Use of abstract concepts allows hard tasks to be learned faster than learning them from scratch, i.e., from low-level perception only. We tested our system with the tasks of learning effect predictions for poke and stack actions using a dataset that includes 83 real-world objects. On the basis of a large number of runs of the method, our analysis shows that the hierarchical task structure emerged as expected, along with a consistent learning order. Furthermore, a significant bootstrapping effect in learning speed of the stack action was observed with the discovered hierarchy, albeit only when fully-learned poke actions were used from the beginning.

Design and Evaluation of a Unique Social Perception System for Human–Robot Interaction

Abolfazl Zaraki, Michael Pieroni, Danilo De Rossi, Daniele Mazzei, Roberto Garofalo, Lorenzo Cominelli, Maryam Banitalebi Dehkordi

Robot's perception is essential for performing high-level tasks such as understanding, learning, and in general, human–robot interaction (HRI). For this reason, different perception systems have been proposed for different robotic platforms in order to detect high-level features such as facial expressions and body gestures. However, due to the variety of robotics software architectures and hardware platforms, these highly customized solutions are hardly interchangeable and adaptable to different HRI contexts. In addition, most of the developed systems have one issue in common: they detect features without awareness of the real-world contexts (e.g., detection of environmental sound assuming that it belongs to a person who is speaking, or treating a face printed on a sheet of paper as belonging to a real subject). This paper presents a novel social perception system (SPS) that has been designed to address the previous issues. SPS is an out-of-the-box system that can be integrated into different robotic platforms irrespective of hardware and software specifications. SPS detects, tracks, and delivers in real-time to robots, a wide range of human- and environment- relevant features with the awareness of their real-world contexts. We tested SPS in a typical scenario of HRI for the following purposes: to demonstrate the system capability in detecting several high-level perceptual features as well as to test the system capability to be integrated into different robotics platforms. Results show the promising capability of the system in perceiving real world in different social robotics platforms, as tested in two humanoid robots, i.e., FACE and ZENO.

Deep Reinforcement Learning With Visual Attention for Vehicle Classification

Dongbin Zhao, Yaran Chen, Le Lv

Automatic vehicle classification is crucial to intelligent transportation system, especially for vehicle-tracking by police. Due to the complex lighting and image capture conditions, image-based vehicle classification in real-world environments is still a challenging task and the performance is far from being satisfactory. However, owing to the mechanism of visual attention, the human vision system shows remarkable capability compared with the computer vision system, especially in distinguishing nuances processing. Inspired by this mechanism, we propose a convolutional neural network (CNN) model of visual attention for image classification. A visual attention-based image processing module is used to highlight one part of an image and weaken the others, generating a focused image. Then the focused image is input into the CNN to be classified. According to the classification probability distribution, we compute the information entropy to guide a reinforcement learning agent to achieve a better policy for image classification to select the key parts of an image. Systematic experiments on a surveillance-nature dataset which contains images captured by surveillance cameras in the front view, demonstrate that the proposed model is more competitive than the large-scale CNN in vehicle classification tasks.

Editor Pierre-Yves Oudeyer, Inria and Ensta ParisTech, France, pierre-yves.oudeyer@inria.fr
Editorial assistant Fabien C. Y. Benureau, CNRU, Okinawa Institute of Science and Technology (OIST), Japan
ISSN 1550-1914