IEEE CIS

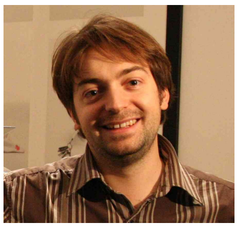
AMD Newsletter

Robotics
Psychology
Neuroscience
Machine Intelligence

The Newsletter of the Autonomous Mental Development Technical Committee

VOLUME 8, NUMBER 1 ISSN 1550-1914 April 2011

Editorial



For the first edition of the joint IEEE ICDL-EpiRob conference, we will have keynote presentations by four eminent scientists from the hard, the wet and the human sciences: Andrew Barto (reinforcement learning and cognitive robotics), Erin Schuman (neuroscience), Jean Mandler and Michael Tomasello (developmental psychology). This interdisciplinarity, focused around a shared set of core questions about development, is an important strength of our field: building bridges that allow for mixing various scientific cultures fosters creativity and insight. I encourage all of you to come and participate at ICDL-EpiRob which will happen in Frankfurt, 24th -27th August in Frankfurt, Germany.

This interdisciplinarity is also strongly reflected in the dialog of this month, initiated by John Weng and continuing the dialog on symbol grounding held in this newsletter last year. John Weng asked “Are natural languages symbolic in the brain?”. This question 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 modern conceptual invention of modern human culture? The answers, written by Stevan Harnad, Jürgen Schmidhüber, Aaron Sloman, Angelo Cangelosi, and Yuuya Sugita and Martin Butz, interestingly mix philosophical and mathematical arguments, showing how recent technical advances can illuminate old questions and vice versa, how philosophical theories can either question or support the assumptions and concepts of modern technical approaches.

Then, a new dialog is proposed by Yaochu Jin and Yan Meng about the interaction between phylogeny, ontogeny and epigenesis, as well as about the interaction between morphological and mental development, urging us to explore new areas in developmental robotics: “Evolutionary Developmental Robotics – The Next Step to Go?”. Those of you interested to react are welcome to submit a response (contact pierre-yves.oudeyer@inria.fr) by September 1st, 2011. The length of each response must be between 300 and 500 words (including references).

— Pierre-Yves Oudeyer, INRIA, Editor

Message from the Chair of IEEE AMD Technical Committee

I would first like to thank all the enthusiastic and continuous contributors to the AMD TC so far. I was appointed as AMD TC chair for the year 2011 (the second year) by IEEE CIS president. The new vice-chairs of our TC are Matthew Schlesinger (US) and Jochen Triesch (Germany). Jochen is one of the general co-chairs of the first ICDL-EpiRob conference 2011 to be held during August 24 -27 in Frankfurt, Germany (www.icdl-epirob.org), and Matt is expected to work for the same conference for the next year. A renewal of TC members has been achieved, and as a result we now have a good geographical balance: 26 from America, 20 from Europe/Africa, and 19 from Asia/Oceania though still very few members are female and from Africa or Oceania. Please encourage the researchers of this category to enter into our TC. I really appreciate your efforts to activate our community even more in the future.

By the way, Japan had a tragic disaster in the form of an earthquake and a tsunami. Nuclear plants were so seriously damaged and are still unstable. It is a huge challenge for current science and technology to confront this difficulty, and we, as researchers, should consider acting based on the right information. Japanese robotics researchers, including myself, opened a web page with their statement of the situation (http://roboticstaskforce.wordpress.com/english/). Any proposals and comments are welcome. I strongly thank our members around the world in advance for their support.

— Minoru Asada, New Chair of the AMD TC

Dialog Column

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 de- pends 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 Column

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).

References:

Penrose, R. (1994) Shadows of the Mind: A Search for the Missing Science of Consciousness. *Oxford University Press*, Oxford.

Penrose, R. (1989) The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. *Oxford University Press*, Oxford.

Moran, J. and Desimone, R. (1985) Selective attention gates visual processing in the extrastrate cortex. *Science*, 229(4715), pp. 782–784.

Felleman, D. J. and Van Essen, D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, pp. 1–47.

Callaway, E. M. (1998) Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience*, 21, pp. 47–74.

Weng, J. (2010) 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.

Luciw, M. and Weng, J. (2010) 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.

Miyan, K. and Weng, J. (2010) 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.

Weng, J., Shen, Y., Chi, M., and X. Xue. (2009) Temporal context as cortical spatial codes. In Proc. *International Joint Conference on Neural Networks*, Atlanta, Georgia, June 14-19.

The Language of Thought Is Symbolic, Grounded and Felt

Stevan Harnad
Chaire de recherche du Canada en sciences cognitives, Université du Québec à Montréal, Canada
School of Electronics and Computer Science, University of Southampton, UK

Weng asks "are the outputs from the brain symbolic?… inputs into the brain seem not really computer symbols either…does a normal… brain have symbols in it?"

I can't answer for everything in the brain, but words certainly are symbols: mental symbols. They refer to things, and kinds of things; and they can be combined into propositions that define and describe more kinds of things. But unlike the symbols in an ungrounded formal system like arithmetic or a computer program, which are manipulated solely on the basis of syntactic rules that operate on the symbols' (arbitrary) shapes and not their meanings, words in the mind are manipulated also on the basis of their meanings. Grounding the words in a symbolizer's sensorimotor (robotic) capacity to recognize, categorize and describe the words' referents in the world is more than ungrounded syntax, but it is not yet meaning. For words in the mind to have meaning, it also has to feel like something to say and mean them. That's the hard problem. Sensorimotor grounding is the "easy" problem.

Weng: "[does] the brain have internal atomic symbols… is any natural language truly symbolic?"

There are words that the brain can connect to referents through sensorimotor experience and interaction. I am not sure whether they should be called "atomic," but there are words that we learn directly from sensorimotor experience, and others that we learn from combining those words into propositions that define still further words, and describe further referents.

Dialog Column

To be "truly symbolic," language would have to be just syntactic: The semantic interpretation would be elsewhere, just as it is not in the pages of a book but in the head of a reader. But the words in the brain are the words in the mind of the reader; it would be homuncular to think of the mind as "reading" them. Rather, grafted onto the words is the sensorimotor grounding plus the feeling. The former will be "easy" to explain, eventually; the latter will not. So I would say that the language of thought is truly symbolic, but not only symbolic.

References:

Harnad, S. (2010) From Sensorimotor Categories and Pantomime to Grounded Symbols and Propositions. In: *Handbook of Language Evolution*, Oxford University Press. http://eprints.ecs.soton.ac.uk/21439/

Symbols as Natural By-Products of History Compression

Juergen Schmidhuber, IDSIA, Galleria 2, 6928 Manno-Lugano
University of Lugano & SUPSI
Switzerland

Isn't it time to end this ancient philosophical symbol grounding debate? Information theory (Shannon, 1948) and algorithmic information theory (Kolmogorov, 1965; Solomonoff, 1964; Solomonoff, 1978; Li, 1997) already provides in hindsight obvious technical explanations (Schmidhuber, 2009a; 2010) of where symbols and self-symbols come from: they may be viewed as simple by-products of data compression during problem solving.

As we interact with the world to achieve goals, we are experiencing a growing, eventually lifelong history of actions and perceptions, never quite the same, as pointed out by Weng, but nevertheless regular in many ways and thus at least partially predictable and therefore compressible—what you can predict you do not need to store.

By partially compressing this data history, we essentially construct an internal model of the world, that is, a comparatively short program able to encode or compute the raw data, in the spirit of the minimum description length principle (Solomonoff, 1964; Wallace, 1968; Solomonoff, 1978; Li, 1997). If the data-compressing program / predictor is a biological or artificial recurrent neural network (RNN) (Werbos, 1988; Williams, 1994; Hochreiter 1997a; 1997b; Graves, 2009), it will automatically create well-known feature hierarchies, lower level neurons corresponding to simple feature detectors similar to those found in human brains (Schmidhuber, 1996), higher layer neurons typically corresponding to more abstract, symbol-like features but fine-grained where necessary to achieve good compression of the whole (Schmidhuber, 1992). Like any reasonable adaptive compressor, the RNN will over time identify shared regularities among different already existing internal data structures, and represent frequently occurring observation sub-sequences by rather localized prototype encodings (*symbols*, if you will) across neuron populations to shrink the storage space (synapses) needed for the whole.

How does this apply to language and texts (Schmidhuber, 1996)? If we present to a *Long Short-Term Memory* RNN (Schmidhuber, 1996) examples of a simple context sensitive language such as $a^n b^n c^n$ (where short and long sequences may map onto the same meaning), it will learn to compactly encode those sequences of greatly varying size across a few symbol-like neurons, some of them used as counters or stacks (Gers, 2001; Schmidhuber, 2002), efficiently representing also many previously unseen instances of the language, since generalization is a natural by-product of compression (Li, 1997). Similarly, for words slowly or quickly spoken by various speaker, LSTM can learn to compactly encode them across relatively few neurons (Beringer, 2005), compressing the essential information into often quite localized, symbol-like representations. Even without a teacher, every regularity that helps to create a compact predictive code of the continually growing data stream will be able to claim some of the limited resources, which are neurons and synapses. In quasi-symbolic ways, reflecting the spatio-temporal statistics and regularities of the environment, some of the neurons will eventually encode individual speakers, or speaker types (e.g., male vs female); others will respond only to a certain frequently used phoneme or word (Beringer, 2005), etc.

Where does *goal-oriented meaning* of symbols come from? Since the brain is a reinforcement learning system (Kaelbling, 1996) embedded in the physical world, such meaning is defined by the reward signals obtained through the action sequences it computes in response to its sensory input sequences.

Dialog Column

For example, there may be *external* reward for eating when hungry, or *intrinsic* reward for improvements of the above-mentioned world model / history compressor (Schmidhuber, 2007; 2009a; 2010). The brain *expresses* such meanings by learning to create action sequences leading to more reward.

As compression by products, such neural symbols typically are not rigid or clear-cut (to answer Weng's question), but will often change during the learner's life time, for two reasons: (1) the data history is continually growing in a problem solving-oriented way that occasionally requires new or modified proto- type encodings (or *symbol changes*) for more efficient compression; (2) Learning may allow for encoding even older parts of the history more efficiently, identifying previously unknown regularities that allow for compression progress (Schmidhuber, 2007; 2009a; 2010).

Self-symbols may be viewed as a by-product of all of this, since there is one thing that is involved in all actions and sensory inputs of the learning agent, namely, the agent itself. To efficiently encode the entire data history, it will profit from creating some sort of internal prototype symbol or code (some compact neural activity pattern) representing itself (Schmidhuber, 2009a; 2010). Whenever this representation becomes activated above a certain threshold, say, by activating the corresponding neurons through new incoming sensory inputs or an internal "search light" or otherwise, the agent could be called self-aware. No need to see this as a mysterious process - it is just a natural by-product of partially compressing the observation history by efficiently encoding frequent observations.

Note that it is not at all ridiculous to assume that a human brain can in principle store a lifelong observation history (Schmidhuber, 2007). A human lifetime rarely lasts much longer than 3×10^9 seconds. The human brain has roughly 10^{10} neurons, each with 10^4 synapses on average. Assuming that only half of the brain's capacity is used for storing raw data, and that each synapse can store at most 6 bits, there is still enough capacity to encode the lifelong sensory input stream with a rate of at least 10^5 bits/s, comparable to the demands of a movie with reasonable resolution, but possibly at a much higher rate, assuming that human compressors are much smarter than those of cameras.

Note also that the recent mathematically optimal general problem solvers and universal AIs (Hutter, 2005; Schmidhuber 2009b) do *not at all* require something like an explicit concept of symbols or *consciousness* and the like — one more reason to consider symbols as a by-product of general intelligence, not a pre-condition.

References:

Beringer, N,. Graves, A., Schiel, F. and Schmidhuber, J. (2005) Classifying unprompted speech by retraining LSTM nets. In *Proc. Int. Joint Conf. on Neural Networks IJCNN 2005*.

Gers, F.A., and Schmidhuber, J. (2001) LSTM recurrent networks learn simple context free and context sensitive languages. *IEEE Transactions on Neural Networks*, 12(6):1333–1340.

Graves, A. and Schmidhuber, J. (2009) Offline handwriting recognition with multidimensional recurrent neural networks. In *Advances in Neural Information Processing Systems 21*. MIT Press, Cambridge, MA.

Hochreiter, H. and Schmidhuber, J. (1997a) Flat minima. *Neural Computation*, 9(1):1–42.

Hochreiter, H. and Schmidhuber, J. (1997b) Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Hutter, M. (2005). *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability*. Springer, Berlin.

Kaelbling, L.P., Littman, M.L. and Moore, A. W. (1996) Reinforcement learning: a survey. *Journal of AI research*, 4:237–285.

Kolmogorov, A.N. (1965) Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:1–11.

Li, M. and Vita Nyi, P.N.B. (1997) *An Introduction to Kolmogorov Complexity and its Applications (2nd edition)*. Springer.

Schmidhuber, J. (1992) Learning complex, extended sequences using the principle of history compression. *Neural Computation*, 4(2):234–242.

Schmidhuber, J. (2007) Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity. In *Proc. 10th Intl. Conf. on Discovery Science (DS 2007), LNAI 4755*, pages 26–38. Springer, 2007. Joint invited lecture for *ALT 2007 and DS 2007*, Sendai, Japan.

Schmidhuber, J. (2009a) Simple algorithmic theory of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. *SICE Journal of the Society of Instrument and Control Engineers*, 48(1):21–32.

Schmidhuber, J. (2009b) Ultimate Cognition à la Gödel. *Cognitive Computation*, 1(2):177–193.

Schmidhuber, J. (2010) Formal theory of creativity, fun, and intrinsic motivation (1990-2010). *IEEE Transactions on Autonomous Mental Development*, 2(3):230 –247.

Schmidhuber, J., Eldracher, M. and B. Foltin. (1996) Semilinear predictability minimization produces well-known feature detectors. *Neural Computation*, 8(4):773–786.

Schmidhuber, J., Gers, F. and Eck, D. (2002) Learning nonregular languages: A comparison of simple recurrent networks and LSTM. *Neural Computation*, 14(9):2039–2041.

Schmidhuber, J. and Heil, S. (1996) Sequential neural text compression. *IEEE Transactions on Neural Networks*, 7(1):142–146.

Dialog Column

Shannon, C. (1948) A mathematical theory of communication (parts I and II). *Bell System Technical Journal*, XXVII:379–423.
Solomonoff, R. J. A (1964) formal theory of inductive inference. Part I. *Information and Control*, 7:1–22.
Solomonoff, R.J. (1978) Complexity-based induction systems. *IEEE Transactions on Information Theory*, IT-24(5):422–432.
Wallace, C.S. and Boulton, D.M. (1968) An information theoretic measure for classification. *Computer Journal*, 11(2):185–194.
Werbos, P.J. (1988) Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, 1.
Williams, R.J. and Zipser, D. (1994) Gradient-based learning algorithms for recurrent networks and their computational complexity. In *Backpropagation: Theory, Architectures and Applications*. Hillsdale, NJ: Erlbaum.

Meaning-bearers in Computers, Brains, and Natural or Artificial Minds

Aaron Sloman
School of Computer Science, The University of Birmingham
Birmingham, B15 2TT
England, UK

Cognition includes symbol-use, both externally and internally, to express meaning or information (not in Shannon's sense). How that is possible is an old problem with several old unsatisfactory answers, such as: meaning is based on experience of things referred to, meaning depends on causal connections between symbol and referent, meanings are possible only because of social/cultural conventions, and expressing meaning requires a human language.

Reference cannot require causal connection, since we can refer to non-existent objects, e.g. "The elephant on the moon". There was a largest mammal in Africa 500 years ago, but our ability to refer to it does not require us to be causally linked to it: meaning, or reference, can use, but does not require, causal links. That's why we can ask questions without knowing the answers, select goals that we may not achieve, and have false beliefs. If meaning depended on causal links most human thinking would be impossible. Reference to genes, quarks and transfinite ordinals cannot depend on experience of referents. Percepts, intentions, learnt associations, and thoughts in pre-verbal children and non-human animals cannot depend on human conventions or language.

Symbols, i.e. discrete meaningful tokens, such as words, are not the only meaning-bearers. Spoken languages also use continuous variation, e.g. of pitch, or intensity; and human sign languages use continuous gestures. Maps, chemical formulae, equations, and semaphore signals, are among the external meaning-bearers we use. Internal meaning-bearers also exist but cannot be found in brains using physical sensing devices, any more than physical sensors can detect spelling correction, or a threatening chess move, in a multi-processing computer. Such events occur in virtual, not physical, machinery. But they exist, since they have causal consequences.

Most interesting contents of computers exist in “running” virtual machines (running VMs), which are implemented in physical machinery, using complex technology based on a tangled causal web of hardware, firmware and software, that alters virtual-physical mappings dynamically. Many tasks, including checking spelling, formatting documents, fetching web pages, and eliminating malware, require VMs. When a computer manipulates numbers the physical machine uses groups of switches implementing bit-patterns that, depending on context, represent numbers, instructions, pointers to complex structures, or other things. Whether a bit-pattern represents a number or something else depends on what procedures are active and what they are doing. When a chess program creates a threat, that is not a physical state. Likewise when I doubt whether someone is talking sense, the doubt is not a physical brain state. "Threat" and "doubt" cannot be defined in the language of physics nor instances detected by physical sensors -- though in simple cases physical footprints may be detectable under special conditions.

Over decades, human engineers found that complex control mechanisms need to operate on entities in virtual machinery that, unlike physical machinery, allows rapid construction and modification of complex structures, and rapid garbage collection after use.

Conjecture: long before that, evolution "discovered" the need for representation and control functions distinct from, but implemented in, physical mechanisms: so mental meaning-bearers exist in those biological VMs running on brains.

Dialog Column

Since perceptual and other contents must change faster than physical parts of brains can be rearranged (e.g. walking with eyes open in a busy city), biological minds need VMs. That can include symbols, for example if you solve equations in your head, rehearse a Shakespearian sonnet, or wonder how brains work. Brain-based VMs can also construct and manipulate diagrams, e.g. visualising the Chinese proof of Pythagoras' theorem, or designing a new information-processing architecture, or imagining the operation of a threaded bolt rotating as it goes into a nut. Virtual machinery includes, but is not restricted to, discrete, discontinuous, structures and processes. Interacting VMs on computers and attached devices run concurrently — their state being preserved in memory while CPUs switch tasks, relying on decades of complex design by hardware and software engineers, solving many different problems — including self-monitoring and control. Very few grasp the big picture combining their efforts.

Biological evolution did something similar, though far more complex and difficult to understand. Support for VMs used in human language, in construction of percepts, in formation of motives, in specifying actions, in generating, evaluating and executing plans, and learning, probably took thousands of intermediate design steps, not yet known to us. Clues exist in the competences of other animals and in pre-verbal children (Karmiloff-Smith, 1992). Exactly what the VMs are, how they evolved, how they are implemented in brains and what their functions are, are still unanswered questions. We cannot find answers simply by studying a narrow subset of products of evolution (e.g. humans) nor a narrow class of robots that mimic some tiny (often arbitrary) subset of animal competence.

Much thinking about language, mind and philosophy of science by roboticists, ignores most of what has already been written over hundreds of years, including work on semantics last century by philosophers of science. Previously, scholars could be familiar with all the important prior published work when investigating a problem, such as the problem of how meaning is possible. But that is no longer possible. I call this the Singularity of Cognitive Catchup (SOCC), see (Sloman, 2010a).

Does SOCC mean that on many important topics we are now doomed to arguing in circles, producing only minor variations on previous failed theories? Perhaps not, if we can find a new high level synthesis to reorganise our thinking. That may be possible if we replace pointless debates (e.g. about embodiment) with deep investigations of the evolutionary discontinuities in information processing requirements and mechanisms, not just in humans but in a wide range of organisms, including microbes, insects and other animals. That will help us focus on the real design issues and help us understand some of the solutions, as suggested in (Sloman, 2010b).

References:

Karmiloff-Smith, A. (1992) Beyond Modularity: A Developmental Perspective on Cognitive Science, *MIT Press*, Cambridge, MA

Sloman, A. (2010a) Yet Another Singularity of Intelligence, http://www.cs.bham.ac.uk/research/projects/cogaff/misc/another-singularity.html

Sloman, A. (2010b) Genomes for self-constructing, self-modifying information-processing architectures, SGAI 2010, http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#sgai10

Symbols in Input and Output, but Not in the Brain

Angelo Cangelosi
Adaptive Behaviour & Cognition Lab, Centre for Robotics and Neural Systems
University of Plymouth, UK

John Weng argues that we (our brains) receive no computer-like symbols in input, neither does the brain output discrete computer-like symbols or behaviors. As a consequence, he concludes that there is no need to assume symbol-like processing in the brain. My suggestion here is that our brain does indeed receive symbol-like entities in input (such as words and numbers), as our categorical perception ability allows us to convert continuous inputs (speech sound waves, retinal images) into discrete categorical and symbolic input for the brain (e.g. discrete representation for words and numbers). However, I argue that our brain does not process these discrete categorical inputs using computer-like symbolic representations and manipulations.

As the brain is based on a symbol-less parallel, distributed processing architecture, the brain is able to integrate embodied discrete representations through mechanisms like mental simulations (Barsalou, 1999). This can then lead to the production, in output, of symbol-like discrete behaviours and entities such as words, numbers, discrete behaviors, computer-like formal symbols etc.

Dialog Column

Our capability to handle input and output symbols is based on well known cognitive phenomena such as categorical perception (Harnad, 1987). Categorical perception is the experience of perceiving invariances in sensory phenomena that vary along a continuum. This is the case of colour perception, where we tend to split (warp) the continuum of light waves into discrete, symbol-like color categories. Categorical perception has been observed in animal and humans and for varying perceptual domains such as vision and speech. Therefore this provides us with the "cognitive groundwork" (Harnad, 1987) to perceive and handle discrete (symbol-like) representations out of inputs that vary over continua such as in speech, gestures and vision. As for the capability to produce symbol-like output entities, I believe that the fact that I am writing this commentary with a set of symbol-like words confirms that at some stage of cognitive processing I am able to produce computer-like (readable) symbols in outputs.

However, our brain is not a computer-like symbol manipulator, as it is not based on a symbolic von Neumann computer architecture, but rather on a parallel distributed processing (PDP) neuronal system. Through this symbol-less PDP architecture the brain is able to use symbol-like input representations (mediated by categorical perception) and produce output behaviors. In the computational modeling field, various neural networks models of categorical perception and symbol grounding exist that implement such a model of cognition (e.g. Damper & Harnad, 2000; Cangelosi, 2005). I believe that the model proposed by Luciw & Weng is consistent with this view of cognitive and symbol processing, with the advantage of proposing a richer, brain-plausible architecture combining top-down and bottom-up mechanisms.

References:

Barsalou, L.W. (1999) Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-609.

Damper, R.I. & Harnad, S. (2000) Neural Network Modeling of Categorical Perception. *Perception and Psychophysics* 62 (4): 843–867.

Cangelosi (2005) Approaches to Grounding Symbols in Perceptual and Sensorimotor Categories. In H. Cohen & C. Lefebvre (Eds), *Handbook of Categorization in Cognitive Science*, Elsevier, pp. 719-737

Harnad, S. (ed.) (1987) Categorical Perception: The Groundwork of Cognition. New York: *Cambridge University Press*. http://cogprints.org/1571/

Compositionality and Embodiment in Harmony


Yuuya Sugita and Martin V. Butz
Department of Psychology III (Cognitive Psychology)
University of Würzburg, Germany

Several aspects of the symbol grounding problem have been solved to significant degrees by now (Steels, 2008) – such as learning to associate pre-defined sensory features with symbols in agent communities. We propose, however, that the deeper problem of *situated linguistic meaning* lies in the challenge of learning to harmonize compositionality with the available embodiment.

In particular, based on Weng's arguments, we believe that the brain does neither establish clear-cut boundaries for symbolic concepts nor for representing relations between such concepts. If this is the case, however, then two deeper, complementary challenges need to be addressed:

1. **Acquiring situated domain-specific systems:** Non-symbolic processes need to learn to produce and thus represent flexible, systematic, and situated interactions, at least for bodily and verbal domains.
2. **Building inter-domain mappings:** Bodily and verbal interaction structures need to be co-associated to situate linguistic meaning.

Although these questions have been investigated on semantic levels in the context of image schema theories (Lakoff, 1987; Johnson, 1987) for more than twenty years, we discuss a fundamental, unsolved difficulty, which is inherent in these theories, from a computational point of view.

According to Hampe (2005), an image schema is an embodied pre-conceptual *gestalt*, which arises from – or is grounded in – human recurrent bodily movements. The most distinctive character of an image schema is that it has a dual nature: it has properties of continuous, analog patterns but it is a composite of identifiable parts. While the analog nature provides a good affinity with its referents in the real world, the compositional structure explains its flexible cognitive performance, representing novel events by adapted recombinations of past experiences.

Dialog Column

However, the dual nature prevents the direct revelation of the non-symbolic computational processes that can learn to represent the *internal* structure of compositional (pre-)concepts, including image schemas (cf. Clausner,2005).

We have proposed that emergent, self-organized geometric representations (Sugita and Tani, 2005; Sugita et al., 2011) may embed analog sensory and motor codes for their recombination. We have shown that geometric representations can be discovered by a connectionist learning approach based on analog, sensory and motor codes. In these emergent geometric, spatial representations, the functionally compositional structure is realized by focusing on the *external*, relational structure among concepts. In particular, each concept is embedded in a self-organized manifold based on the sensory-motor similarities amongst the concepts in a given *n*-dimensional vector space. The relationships among concepts are reflected by the self-organized geometric arrangements. For example, six actions specified by every possible combination of one of three target objects and one of two types of interactions may form a systematic geometric arrangement, similar to the triangular prism shown in Fig. 1. In our robotic experiments, comparable arrangements self-organized through the learning of subsets of body-environment interactions without using any explicit compositional cues. These arrangements generalized to conceptually related, but untrained interactions.

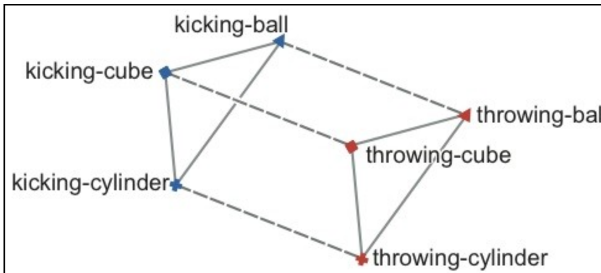


Figure 1. A geometrically arranged concept manifold, which embeds particular object-respective interactions

The inter-domain mapping problem proposed above has also been investigated based on self-organized, geometric representations. Sugita and Tani (2005) discussed how the associations between sentences and their referents may be generalized with- out decomposing the sentences into words explicitly. Sentences, represented as verbal sequences, were embedded in a representational manifold, similarly to the bodily interactions. The generalization of the associations was then illustrated by confirming structural alignment between the two representational manifolds without employing atomic symbols.

These robotic experiments have so far been conducted in rather small-scale environments with restricted sets of behavioral interactions and a supervised, connectionist learning approach. Nonetheless, we believe that there is lots of room for adding (1) modularity and (2) anticipatory, self-motivated drives – two ingredients that the brain appears to use excessively (Butz, 2008). Seeing that the brain solves the situated linguistic meaning problem seemingly at ease, the mixture of computational techniques, modular arrangements, and module interactions employed by it are clearly still to be unraveled. Robotic experiments with progressively more sophisticated neural, sub-symbolic learning architectures will help to solve this challenging puzzle.

References:

Butz, M. V. (2008) How and why the brain lays the foundations for a conscious self. Constructivist Foundations, 4(1):1–42.

Clausner, T. C. (2005) Image schema paradoxes: Implications for cognitive semantics. In Hampe, B., editor, From Perception to Meaning: Image Schemas in Cognitive Linguistics, pages 93–110. Mouton de Gruyter, Berlin.

Hampe, B. (2005) Image schemas in cognitive linguistics: Introduction. In Hampe, B., editor, From Perception to Meaning: Image Schemas in Cognitive Linguistics, pages 1–12. Mouton de Gruyter, Berlin.

Johnson, M. (1987) The Body in the Mind. The Bodily Basis of Meaning. *Chicago University Press*, Chicago.

Lakoff, G. (1987) Women, Fire and Dangerous Things. What Categories reveal about the Mind. *The University of Chicago Press*, Chicago.

Steels, L. (2008) The symbol grounding problem has been solved. So what's next? In de Vega, M., Glenberg, A. M., and Graesser, A. C., editors, Symbols and Embodiment: Debates on meaning and cognition, pages 223–244. Academic Press, New Haven.

Sugita, Y. and Tani, J. (2005) Learning Semantic Combinatoriality from the Interaction between Linguistic and Behavioral Processes. *Adaptive Behavior*, 13(1):33–52.

Sugita, Y., Tani, J., and Butz, M. V. (2011) Simultaneously Emerging Braitenberg Codes and Semantic Compositionality - *submitted to Adaptive Behavior*. The draft is available at http://www.coboslab.psychologie.uni-wuerzburg.de/people/yuuya sugita/

Dialog Column

Reply and Summary:
Through the Symbol-Grounding Problem See the Two Largest Hindrances of Science

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

While approaching a scientific subject, it seems ineffective to adopt an approach that many scientists like to take — restrict the subject to a narrowly defined domain. For example, the symbol-grounding problem that this dialog addresses seems to be related to a super science, called developmental science — the science about the development of multi-agent systems, natural and artificial. Why? Just through this short reply I am not able to make the reader to convincingly understand a brain-mind solution to this problem. But I hope that the reader could proceed in the way I suggest here to reach a conclusion for himself. Furthermore, I wish that we humans collectively can be sensitized to many fundamental problems in our existing infrastructures. I call them two largest hindrances of science.

First, our editor Pierre-Yves Oudeyer received five responses from North America, Europe, and Japan. This small sample size does not have any justifiable statistical significance. However, this fact appears to be consistent with what we have seen: these geographic regions have the environment necessary for researchers to study all important problems of AMD. Why other countries have less active research on this important subject? I think that this is due to, in other geographic regions, a lack of the necessary developmental program, which includes the constitution, laws and bylaws. What I like to say is that the largest hindrance of all science is the lack of the developmental program. A government might emphasize science, but the government fails to provide necessary developmental program that are necessary for researchers in its country to take the lead in solving major open science problems, such as the brain-mind problem. China and Russia are two examples. I have been personally involved with collaborative academic research in China and know the research environment there.
If I were in China it would have been impossible for me to conduct original research of this kind. All developed countries seem to have one thing in common — their constitutions and laws are superior in terms of check-and-balance of power (Kernell et al., 2008). This is like a superior "genome" program for a nation. Check-and-balance of power seems the most basic principle of such a genome program for a nation, not superficial guarantee of basic food and basic housing. I expressed to some Chinese administrators that China needs to seriously learn check-and-balance of power, extensively put this principle into its constitution, laws and regulations, and effectively execute them. Before the effect has taken place for many years, China could not enter the category of developed countries. I am fortunate to do research in US, although my experience of doing cross-disciplinary research in US has been harsh which I will allude to while I discuss the second hindrance below.

The second largest hindrance is as follows. The human race, collectively, may have sufficient knowledge to solve the brain-mind problem. However, such vast knowledge is piece-meals in the minds of many domain researchers. Understanding the solution of the brain-mind problem by each individual mind requires a superset (the set of union) of such vast knowledge in the mind. The existing infrastructure in the developed countries severely hinders the publication, awareness, proliferation, and technology transfer of the first brain-mind solution (in the 5+1 chunks to be outlined below) that has been discovered in my group. Let us use the symbol-grounding problem as a sub-problem of the brain-mind problem for us to discuss the second hindrance. I will discuss it in terms of following three intertwined aspects.

Existing university curricula. Existing university curricula continued their reliance on the works of Aristotle and how he defined science and the arts. Although various multi-disciplinary and newer academic programs have emerged, such as biophysics, cognitive science, neuroscience, and computer science, the scope of any single academic curriculum falls far short of the necessary background needed to solve or understand the brain-mind problem. The initial brain-mind solution that my coworkers and I have arrived at (Weng, 2010; 2011; Daly et al., 2011) seems the first that spans the 5+1 chunks: development, architecture, area, space, time, plus neuromodulation. I humbly express my view here: how the brain-mind works is not a myth any more. However, according to this 5+1 chunk solution, to understand how the brain-mind works, especially for convinced appreciation, one needs knowledge in at least 6 disciplines: biology, neuroscience, psychology, computer science, electrical engineering, and mathematics. Let us take a look at the symbol-grounding problem.

Dialog Column

I am happy to see that the commentaries in this Dialog are from some of the most active researchers on this broad and challenging subject. Harnad is correct in his comments, “words in the mind are manipulated also on the basis of their meanings”. He correctly insisted: “For words in the mind to have meaning, it also has to feel like something to say and mean them. That’s the hard problem.” However, he did not explain how his claimed “mental symbols” are represented in the brain. Our brain-anatomy inspired WWN (Luciw, 2010; Miyan, 2010) explains computationally that almost every neuron inside the brain skull is a mix of sensory and motor inputs, where motor is the meanings. That is how “words in the mind are manipulated also on the basis of their meanings”. WWN demonstrated that the meanings of words emerge in motor areas (i.e., in terms of muscles and glands) as their meanings but overt actions do not have to be always displayed. That is how it “feels like something to say and mean them.” Sugita and Butz proposed a geometrically-arranged concept manifold, in that the object and the associated action are represented as a node. This belongs to what is called object-action framework that has been popular in psychology and cognitive vision. However, their representation is symbolic which requires the human designer to handpick the contents for each symbol or node. Their work did not explain how the internal representations inside the brain skull emerge without requiring the genome (like a homunculus parent inside the brain) to completely specifying the contents of objects and their actions. The current curricula in psychology do not seem to provide a sufficient infrastructure for psychologists to learn brain science. Psychologists follow their tradition of observing the external behaviors from human subjects, but external behaviors explain little about the inner working of the brain.

Sloman is in the right direction when he eschews “causal connections”, a habit of many computer science researchers who use “fact-based” symbolic representations. However, he did not state what his “internal meaning-bearers” mean in the brain. Our brain-mind model (Weng, 2010; 2011, Daly, 2011) proposes that in general, every neuron in the brain is not meaning-pure since it is a mix of sensory input and motor input and it is shared by other meanings. Our brain-mind model learns immediately, error free, like a perfect logic machine but it does not use any symbolic representations. Cangelosi correctly emphasized that the brain is a “parallel distributed processing (PDP) neuronal system.” Nevertheless, he did not provide brain evidence for his assumption that the brain “convert continuous inputs (speech sound waves, retinal images) into discrete categorical and symbolic input for the brain (e.g. discrete representation for words and numbers).” Our brain-mind model (Miyan, 2010) clarifies that words are not represented as discrete representations for sensory objects exclusively, but as border-less representation mixed with temporal context, motor, and sensory information. Schmidhuber’s LDTM was an advance from some prior recurrent networks in that they defined some internally controlled nodes that serve as temporal “gates” for triggering some temporal events (called short-term memory by Schmidhuber). Our brain-mind model (Weng, 2011; Miyan & Weng, 2010) explains that the brain does not have any “gates” dedicated exclusively to time. Internal representation inside the closed skull autonomously emerge without separating space and time at all, so that it can effectively and efficiently deal with arbitrary spatial and temporal information and their arbitrary equivalence learned from the environment. This way, WWN does not treat any short-term memory as a discrete temporal block. But rather, any spatiotemporal chunk can be “warped”. Any combinations of “chunks” and parts of “chunks” can be learned by the model as equivalent.

These commentaries also manifest casualties from the second hindrance. Although exceptions do exist, the following situation is currently prevailing: computer scientists model logic reasoning but they use symbolic representations and, thus, they are not keen on emergent representations that the brain and neural networks use. They do not know, or are not interested in, that emergent representations enable the brain to dynamically learn new concepts directly from the real physical world, without requiring the human programmer to understand the tasks and concepts that the machine ends up learning. Electrical engineers use neural networks but they do not have sufficient exposure to computer science, e.g., what it takes for computers to conduct symbolic reasoning. Computer scientists and electrical engineers do read literature in psychology, but they do not find the neuroscience literature to be particularly helpful. This is because neuroscientists are not in a position to explain how the brain works. Symbolic AI methods and engineering-style neural networks become increasingly familiar to psychologists but their limitations are largely not yet well known by psychologists. Therefore, psychologists borrow Good Old-Fashioned Artificial Intelligence (GOFAI as Rodney Brooks put it) and superficially treat it as the mind. See, e.g., the symbolic methods recently reviewed in (Tenenbaum et al., 2010).

Statistics-based ranking and performance metrics. Although institution ranking by itself is not necessarily meant to be stereotypical, the data from various types of institution ranking tend to be used improperly. If racial profiling is considered not a fair practice by a government, how fair is institution ranking?

Dialog Column

Is it fair to judge a member based on a ranking of the institution he is affiliated with? Is it fair to bias the judgment on a person based on the mean or other statistical measures of whatever the group under which he can be easily identified? For example, should the NRC Assessment of Research Doctorate Programs be used as an NRC suggested metric by a graduate applicant to decide which PhD program he should apply and enroll? Should he look instead into the research activities of individual faculty members instead?

How much has the impact factors of a journal stereotypically biased our perception of papers in the journal? How much has the institution ranking of an author stereotypically biased the reviewers of a journal? How much has the superficial rejection rate biased the perception of a journal or a conference? Randomness is the human laziness in exploring the true causality. There is no true randomness in the physical world other than technical challenges to measure (the uncertainty principle in physics included).

Anthony van Raan said "You should never used the journal impact factor to evaluate research performance for an article or for an individual — that is a mortal sin.", "All these people should know better than to think that there is a single measure you can use", David Pendlebury stated (Van Noorden, 2010). Various rankings seem to have a negative influence when researchers consider exploring newer, less established, less mature research subjects, since such subjects do not have well established publication venues, and reviewers are not familiar with the new subjects, and the citation numbers will not be high since there are relatively few people working on such new subjects. How can we promote better understanding of scientific research so that the highly complex nature of research is not degenerated into a single number?

The current peer review system. The peer review system is like multiple flocks of sheep, each with a close leader so that others in the same flock can follow closely. Outliers that travel among the flocks fall prey to natural predators. This example of the symbol-grounding problem provides a picture how a proposal or research article for this cross-disciplinary problem would have been reviewed by the peer review system. A paper about a solution to the brain-mind problem has a more miserable fate, as it is about many "sheep flocks".

Here is an example of review cases repeated many times in the history of science (Kuhn, 1970). We have had a brain-like solution to the symbol-grounding problem. The Theorem 1 in (Weng, 2011) establishes that the symbolic representation of any finite automaton (FA) is a special case of the emergent representation of the Developmental Network (DN) in the following sense. An FA models only the external behaviors of an agent (e.g., brain) assuming it works in a disembodied symbolic world. The FA is a model of the collective knowledge of the human society as well as the physical world causality. Given any FA, the general purpose developmental program (DP) of DN fully autonomously self-organizes inside the network (e.g., inside the brain skull) when the DN incrementally learns the FA. Unlike the FA, the DN is grounded in the physical world since it directly takes input vectors as its sensory inputs and input vectors from its motor areas as direct and indirect effects from the physical world. It applies its responses in the motor area directly to the physical world through the muscles and the glands. We have proved that the learning of DN from FA is immediate and error-free for each state transition (i.e., step) of the FA. This theorem unifies the symbolic school and the connectionist school which have had over 30 years of heated debate and divide. This author submitted the Theorem 1 and its full mathematical proof to a long-established journal in computer science. However, the journal refused to review the paper without a clearly stated reason while indirectly stating we accept "mature work". After contacting a few well known colleagues, this author resubmitted the paper to the same journal, but the journal rejected the paper again stating "without the possibility of further resubmission". The second rejection email has a few lines of informal comments, without a formal review by at least two reviewers as for other normal submissions. In fact, not only the connectionist school, the symbolic school can also greatly benefit from the Theorem 1 and its proof, since every past symbolic work has now a connectionist counterpart. From it, we can see how the brain can incrementally learn the skills of every symbolic model published in this journal and elsewhere.

Currently, although the symbol-grounding problem has had a theoretical, brain-inspired solution — Theorem 1 — the proof does not seem possible to appear in this journal. This line of highly multidisciplinary work requires a sufficient background in computer science, electrical engineering, mathematics, biology, neuroscience and psychology. Reviewers with such diverse background are extremely rare, if existed. The same is true for the editorial boards of highly multi-disciplinary journals and the review boards of US federal funding agencies. The limited scope of existing university curricula is one background reason, while the flocks-of-sheep phenomenon is a psychological reason.

Dialog Column

Another critical issue I like to raise here for us to consider is: What is the purpose of cross-disciplinary review which I have also promoted? There are at least two alternatives: (1) Publishing only a common-denominator like paper that every reviewer, regardless of his discipline, can comfortably understand. (2) Encourage each reviewer to only check for his own home discipline so that a cross-disciplinary journal can publish work that spans the super set of several disciplines. If (1) is enforced as some reviewers and journals insist, what is in the intersection of all the 6 discipline? Is (1) appropriate for encouraging investigation on cross-disciplinary brain-mind related problems such as the symbol-grounding problem?

To summarize, from the hindsight, I think that a solution to this symbol-grounding problem is unlikely correct without a clear computational understanding about how the brain-mind develops and works. Fortunately, according to the 5+1 chunks solution from my students and I, it seems that all the knowledge necessary to reach a correct answer to the brain-mind problem has already been published in the literature across the 6 disciplines. April 15, 2011, MSU has had the first planning meeting to establish the Brain-Mind Institute (BMI) to host summer institute for the knowledge in these 6 disciplines from 2012. The BMI web site is at http://www.cse.msu.edu/bmi/. With a clear theoretical framework for the brain-mind and its mathematical properties as the guide, I predict that the rise of the brain-mind technology will soon arrive.

Reference:

Daly, J., Brown, J. and Weng, J. (2011) Neuromorphic motivated systems. In *Proc. Int. Joint Conference on Neural Networks*, pages +1–8, San Jose, CA, July 31 - August 5.

Kernell, S., Jacobson, G.C. and Kousser, T. (2008) The Logic of American Politics. *Congressional Quarterly*, Washington DC, 4th edition.

Kuhn, T.S. (1970) The Structure of Scientific Revolutions. *University of Chicago Press*, Chicago, IL, second edition.

Luciw, M. and Weng, J. (2010) Where What Network 3: Developmental top-down attention with multiple meaningful foregrounds. In *Proc. IEEE Int. Joint Conference on Neural Networks*, pages 4233–4240, Barcelona, Spain, July 18-23.

Miyan, K. and Weng, J. (2010) WWN-Text: Cortex-like language acquisition with What and Where. In *Proc. IEEE 9th Int. Conference on Development and Learning*, pages 280–285, Ann Arbor, August 18-21.

Van Noorden, R. (2010) A profusion of measures. *Nature*, 15:864–866.

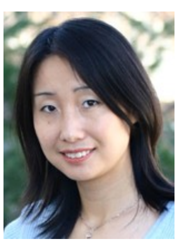
Tenenbaum, J.B., Kemp, C., Griffiths, T.L. and Goodman, N.D. (2011) How to grow a mind: Statistics, structure, and abstraction. *Science*, 331:1279–1285.

Weng, J. (2010) A 5-chunk developmental brain-mind network model for multiple events in complex backgrounds. In *Proc. Int. Joint Conf. Neural Networks*, pages 1–8, Barcelona, Spain, July 18-23.

Weng, J. (2011) Three theorems: Brain-like networks logically reason and optimally generalize. In *Proc. Int. Joint Conference on Neural Networks*, pages 1–8, San Jose, CA, July 31 - August 5.

Dialog Initiation

Evolutionary Developmental Robotics – The Next Step to Go?

Yaochu Jin, Department of Computing, University of Surrey, UK
Yan Meng, Department of Electrical and Computer Engineering
Stevens Institute of Technology, USA

Developmental robotics, commonly known as *epigenetic robotics* (Metta & Berthouze, 2006) is mainly concerned with modeling the postnatal development of cognitive behaviors in living systems, such as language, emotion, curiosity, anticipation, and social skills. Over the past decade, epigenetic robotics has enjoyed great success and achieved significant progress in understanding mental development.

Despite its significant success, the current approaches to developmental robotics can be challenged both from ontogenetic and phylogenetic point of view. First, development of biological organisms consists of both physical and mental development. Thus, developmental robotics that concentrates on mental development is incomplete. Ontogenetically, mental development is based on and closely coupled with physical development of an organism, including development of both the body plan and the nervous systems. For example, new findings in neuroscience reveal that early development of nervous systems are also considerably driven by neural activity (Spitzer, 2006), suggesting that activity-dependent and activity-independent development of neural networks cannot be separated. *Morphogenetic robotics* (Jin & Meng 2011), inspired from biological morphogenesis, has recently been proposed as a new research area that deals with the physical development of robotic systems, thus complementing epigenetic robotics to provide a full picture of developmental robotics.

Second, biological evidence suggests that autonomous mental development is driven by intrinsic motivational systems (Deci and Ryan, 1985), among others. Such findings have triggered one key research area in epigenetic robotics that studies the role of intrinsic motivation systems in mental development (Oudeyer et al., 2007). Typically, an intrinsic motivation system is pre-defined in the robotic system. However, autonomous mental development in living system was not *hard-wired* from scratch, rather has gradually shaped by a brain-body co-evolution embedded in a changing environment.

The introduction of morphogenetic robotics addresses the first challenge in developmental robotics to a certain extent by integrating mental and physical development. Nevertheless, the evolutionary origin that accounts for both physical and mental development is still missing. In addition, in morphogenetic robotics, the gene regulatory network that governs the neural and morphological development cannot be designed manually. Thus, the role of *evolution* becomes increasing important in studying physical and mental development in robotics systems.

Evolutionary robotics (Nolfi & Floreano, 2006) applies evolutionary algorithms to the automatic design of neural controllers for autonomous robots. Unfortunately, the role of development has largely been neglected. We argue that integrating research on developmental (including epigenetic and morphogenetic) robotics and evolutionary robotics is indispensable. Developmental plasticity can not only bias evolution, but also enhance evolvability by maintaining genetic diversity in changing environments and resolving robustness-variability trade-off. For example, it has been shown most recently that an evolutionary perspective on autonomous learning might bring about new insight into understanding the driving force behind mental development (Singh et al, 2010).

The past decade has witnessed rapid technical and theoretical advances in evolutionary developmental biology (Müller, 2007) (often known as evo-devo) understanding molecular and cellular mechanisms that control the biological morphogenesis. Evolutionary developmental biology has also helped us gain deep insight into human cognitive development, resulting in a new discipline known as evolutionary developmental psychology (Griffiths, 2007). As the role of genes is too important to be neglected in cognitive development (Ramus, 2006), we believe that it is high time to bring together evolutionary robotics and developmental robotics to form a new discipline *evolutionary developmental robotics* (evo-devo-robo).

Going from epigenetic robotics to evolutionary developmental robotics is not straightforward. A few important questions remain to be answered. For example, how to realize physical development in robotic systems? Must we use hardware (real robotic system) for research in evo-devo-robo? Many researchers in evolutionary robotics are convinced that a real robotic system must be used, which, unfortunately might have been the biggest hurdle that has prevented this area from further progressing. Another question is, given two different time scales in evolution and development, how much resources should be devoted to evolution and how much to development? How to provide an environment, real or simulated, that is complex enough to enable the system to evolve behaviors of cognitive significance? Thus, our proposal to move from epigenetic robotics towards evo-devo-robo may have raised more questions than answers in study autonomous mental development.

References:

Metta, G. and Berthouze, L. (2006) Epigenetic robotics: Modeling cognitive development in robotic systems. *Interaction Studies,*7:129–134

Spitzer, N.(2006) Electrical activity in early neuronal development. *Nature*, 444:707–712.

Jin, Y. and Meng, Y. (2011) Morphogenetic Robotics: An Emerging New Field in Developmental Robotics. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 41:145-160.

Deci, E. L. and Ryan, R. M. (1985) Intrinsic motivation and self-determination in human behavior. New York: Plenum Press

Oudeyer, P. and Kaplan, F. and Hafner, V.V. (2007) Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, 11:265-286.

Nolfi, S. and Floreano, D. (2004) Evolutionary Robotics. The MIT Press.

Singh, S., Lewis, R.L., Barto, A.G. and Sorg J. (2010) Intrinsically Motivated Reinforcement Learning: An Evolutionary Perspective. *IEEE Transactions on Autonomous Mental Development*, 2:70-82.

Müller, G. (2007) Evo-devo: extending the evolutionary synthesis. *Nature Review Genetics*, 8:943–949.

Griffiths, P.E. (2007) "Evo-devo meets the mind: Towards a developmental evolutionary psychology." in Integrating Development and Evolution, R. Sanson and R. Brandon, Eds. Cambridge University Press.

Ramus, F. (2006) Genes, brain, and cognition: A roadmap for the cognitive scientist. Cognition, 101:247-269

Call for Participation: IEEE ICDL-EpiRob 2011

http://www.icdl-epirob.org/

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 Massachusetts 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.

The IEEE CIS will offer travel grants for students and researchers from developing countries presenting a paper at IEEE ICDL-EpiRob 2011. Students and researchers must have both IEEE and IEEE CIS memberships (CIS is listed as an optional society membership under IEEE) and must have joined the society before applying for a travel grant. In addition students / researchers must have papers accepted at the conference. Further information about how travel grants are awarded can be found here http://ieee-cis.org/members/travel/. The deadline for travel grant submissions is June 20. See Conference website for additional information.

IEEE TRANSACTIONS ON AUTONOMOUS MENTAL DEVELOPMENT

Volume 2, Issue 4, December 2010 **Link:** http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=5645770

Guest Editorial Representations and Architectures for Cognitive Systems

Metta, G. Cheng, G. Asfour, T. Caputo, B. Tsotsos, J.K. Page(s): 265-266 (pdf)

Abstract: The seven papers in this special issue focus on the development of artificial cognitive systems.

A Probabilistic Appearance Representation and Its Application to Surprise Detection in Cognitive Robots

Maier, W. Steinbach, E. Page(s): 267-281 (pdf)

Abstract: In this work, we present a novel probabilistic appearance representation and describe its application to surprise detection in the context of cognitive mobile robots. The luminance and chrominance of the environment are modeled by Gaussian distributions which are determined from the robot's observations using Bayesian inference. The parameters of the prior distributions over the mean and the precision of the Gaussian models are stored at a dense series of viewpoints along the robot's trajectory. Our probabilistic representation provides us with the expected appearance of the environment and enables the robot to reason about the uncertainty of the perceived luminance and chrominance.

Hence, our representation provides a framework for the detection of surprising events, which facilitates attentional selection. In our experiments, we compare the proposed approach with surprise detection based on image differencing. We show that our surprise measure is a superior detector for novelty estimation compared to the measure provided by image differencing.

Self-Understanding and Self-Extension: A Systems and Representational Approach

Wyatt, J.L. Aydemir, A. Brenner, M. Hanheide, M. Hawes, N. Jensfelt, P. Kristan, M. Kruijff, G.M. Lison, P. Pronobis, A. Sjoo, K. Vrecko, A. Zender, H. Zillich, M. Skocaj, D. Page(s): 282-303 (**pdf**)

Abstract: There are many different approaches to building a system that can engage in autonomous mental development. In this paper, we present an approach based on what we term self-understanding, by which we mean the explicit representation of and reasoning about what a system does and does not know, and how that knowledge changes under action. We present an architecture and a set of representations used in two robot systems that exhibit a limited degree of autonomous mental development, which we term self-extension. The contributions include: representations of gaps and uncertainty for specific kinds of knowledge, and a goal management and planning system for setting and achieving learning goals.

Body Schema in Robotics: A Review

Hoffmann, M. Marques, H. Arieta, A. Sumioka, H. Lungarella, M. Pfeifer, R. Page(s): 304-324 (**pdf**)

Abstract: How is our body imprinted in our brain? This seemingly simple question is a subject of investigations of diverse disciplines, psychology, and philosophy originally complemented by neurosciences more recently. Despite substantial efforts, the mysteries of body representations are far from uncovered. The most widely used notions-body image and body schema-are still waiting to be clearly defined. The mechanisms that underlie body representations are coresponsible for the admiring capabilities that humans or many mammals can display: combining information from multiple sensory modalities, controlling their complex bodies, adapting to growth, failures, or using tools. These features are also desirable in robots. This paper surveys the body representations in biology from a functional or computational perspective to set ground for a review of the concept of body schema in robotics. First, we examine application-oriented research: how a robot can improve its capabilities by being able to automatically synthesize, extend, or adapt a model of its body. Second, we summarize the research area in which robots are used as tools to verify hypotheses on the mechanisms underlying biological body representations. We identify trends in these research areas and propose future research directions.

Epigenetic Robotics Architecture (ERA)

Morse, A.F. de Greeff, J. Belpeame, T. Cangelosi, A. Page(s): 324-339 (**pdf**)

Abstract: In this paper, we discuss the requirements of cognitive architectures for epigenetic robotics, and highlight the wider role that they can play in the development of the cognitive sciences. We discuss the ambitious goals of ongoing development, scalability, concept use and transparency, and introduce the epigenetic robotics architecture (ERA) as a framework guiding modeling efforts. A formal implementation is provided, demonstrated, and discussed in terms of meeting these goals. Extensions of the architecture are also introduced and we show how the dynamics of resulting models can transparently account for a wide range of psychological phenomena, without task dependant tuning, thereby making progress in all of the goal areas we highlight.

Multilevel Darwinist Brain (MDB): Artificial Evolution in a Cognitive Architecture for Real Robots

Bellas, F. Duro, R.J. Faina, A. Souto, D. Page(s): 340-354 (**pdf**)

Abstract: The multilevel Darwinist brain (MDB) is a cognitive architecture that follows an evolutionary approach to provide autonomous robots with lifelong adaptation. It has been tested in real robot on-line learning scenarios obtaining successful results that reinforce the evolutionary principles that constitute the main original contribution of the MDB. This preliminary work has lead to a series of improvements in the computational implementation of the architecture so as to achieve realistic operation in real time, which was the biggest problem of the approach due to the high computational cost induced by the evolutionary algorithms that make up the MDB core. The current implementation of the architecture is able to provide an autonomous robot with real time learning capabilities and the capability for continuously adapting to changing circumstances in its world, both internal and external, with minimal intervention of the designer. This paper aims at providing an overview or the architecture and its operation and defining what is required in the path towards a real cognitive robot following a developmental strategy. The design, implementation and basic operation of the MDB cognitive architecture are presented through some successful real robot learning examples to illustrate the validity of this evolutionary approach.

Integration of Active Vision and Reaching From a Developmental Robotics Perspective

Hülse, M. McBride, S. Law, J. Lee, M. Page(s): 355-367 (**pdf**)

Abstract: Inspired by child development and brain research, we introduce a computational framework which integrates robotic active vision and reaching. Essential elements of this framework are sensorimotor mappings that link three different computational domains relating to visual data, gaze control, and reaching. The domain of gaze control is the central computational substrate that provides, first, a systematic visual search and, second, the transformation of visual data into coordinates for potential reach actions. In this respect, the representation of object locations emerges from the combination of sensorimotor mappings. The framework is tested in the form of two different architectures that perform visually guided reaching. Systematic experiments demonstrate how visual search influences reaching accuracy. The results of these experiments are discussed with respect to providing a reference architecture for developmental learning in humanoid robot systems.

Development of Object and Grasping Knowledge by Robot Exploration

Kraft, D. Detry, R. Pugeault, N. Bas¸eski, E. Guerin, F. Piater, J.H. Krüger, N. Page(s): 368-383 (**pdf**)

Abstract: We describe a bootstrapping cognitive robot system that-mainly based on pure exploration-acquires rich object representations and associated object-specific grasp affordances. Such bootstrapping becomes possible by combining innate competences and behaviors by which the system gradually enriches its internal representations, and thereby develops an increasingly mature interpretation of the world and its ability to act within it. We compare the system's prior competences and developmental progress with human innate competences and developmental stages of infants.

Volume 3, Issue 1, March 2011 **Link:** http://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=5729985

Editorial: Healthy and Prosperous Development

Zhang, Z. Page(s): 1 - 2 (**pdf**)

Editorial Board Update

Zhang, Z. Page: 3 (**pdf**)

Understanding Psychological Development in Biological and Artificial Agents: Report on the International Conference on Development and Learning (ICDL 2010)

Shultz, R. Kuipers, B. Page(s): 4 - 5 (**pdf**)

The Impact of Participants' Beliefs on Motor Interference and Motor Coordination in Human–Humanoid Interactions

Qiming Shen Kose-Bagci, H. Saunders, J. Dautenhahn, K. Page(s): 6 - 16 (**pdf**)

Abstract: This study compared the responses of human participants studying motor interference and motor coordination when they were interacting with three different types of visual stimuli: a humanoid robot, a pendulum, and a virtual moving dot. Participants' responses indicated that participants' beliefs about the engagement of the robot affected the elicitation of the motor interference effects. Together with research supporting the importance of other elements of robot appearance and behavior, such as bottom-up effects and biological motion profile, we hypothesize that it may be the overall perception (in this study, by the term "overall perception," we mean the human observer's overall perception of the robot in terms of appearance, motion, and observer's beliefs) of a robot as a "social entity" instead of any individual appearance or motion feature that is critical to elicit the interference effect in human-humanoid interaction. Moreover, motor coordination responses indicated that the participants tended to synchronize with agents with better overall perception, which were generally in-line with the above hypothesis. Based on all the results from this experimental study, the authors suggest that a humanoid robot with good overall perception as a "social entity" may facilitate "engaging" interactions with a human.

Integration of Speech and Action in Humanoid Robots: iCub Simulation Experiments

Tikhanoff, V. Cangelosi, A. Metta, G. Page(s): 17 - 29 (**pdf**)

Abstract: Building intelligent systems with human level competence is the ultimate grand challenge for science and technology in general, and especially for cognitive developmental robotics. This paper proposes a new approach to the design of cognitive skills in a robot able to interact with, and communicate about, the surrounding physical world and manipulate objects in an adaptive manner. The work is based on robotic simulation experiments showing that a humanoid robot (iCub platform) is able to acquire behavioral, cognitive, and linguistic skills through individual and social learning. The robot is able to learn to handle and manipulate objects autonomously, to understand basic instructions, and to adapt its abilities to changes in internal and environmental conditions.

Using the Rhythm of Nonverbal Human–Robot Interaction as a Signal for Learning

Andry, P. Blanchard, A. Gaussier, P. Page(s): 30 - 42 (**pdf**)

Abstract: Human-robot interaction is a key issue in order to build robots for everyone. The difficulty for people to understand how robots work and how they must be controlled will be one of the mains limit for broad robotics. In this paper, we study a new way of interacting with robots without needing to understand how robots work or to give them explicit instructions. This work is based on psychological data showing that synchronization and rhythm are very important features for pleasant interaction. We propose a biologically inspired architecture using rhythm detection to build an internal reward for learning. After showing the results of keyboard interactions, we present and discuss the results of real human-robots (Aibo and Nao) interactions. We show that our minimalist control architecture allows the discovery and learning of arbitrary sensorimotor associations games with expert users. With nonexpert users, we show that using only the rhythm information is not sufficient for learning all the associations due to the different strategies used by the human. Nevertheless, this last experiment shows that the rhythm is still allowing the discovery of subsets of associations, being one of the promising signal of tomorrow social applications.

Implicit Sensorimotor Mapping of the Peripersonal Space by Gazing and Reaching

Chinellato, E. Antonelli, M. Grzyb, B.J. del Pobil, A.P. Page(s): 43 - 53 (**pdf**)

Abstract: Primates often perform coordinated eye and arm movements, contextually fixating and reaching towards nearby objects. This combination of looking and reaching to the same target is used by infants to establish an implicit visuomotor representation of the peripersonal space, useful for both oculomotor and arm motor control. In this work, taking inspiration from such behavior and from primate visuomotor mechanisms, a shared sensorimotor map of the environment, built on a radial basis function framework, is configured and trained by the coordinated control of eye and arm movements. Computational results confirm that the approach seems especially suitable for the problem at hand, and for its implementation on a real humanoid robot. By exploratory gazing and reaching actions, either free or goal-based, the artificial agent learns to perform direct and inverse transformations between stereo vision, oculomotor, and joint-space representations. The integrated sensorimotor map that allows to contextually represent the peripersonal space through different vision and motor parameters is never made explicit, but rather emerges thanks to the interaction of the agent with the environment

Towards an Understanding of Hierarchical Architectures

Goerick, C. Page(s): 54 - 63 (**pdf**)

Abstract: Cognitive systems research aims to understand how cognitive abilities can be created in artificial systems. One key issue is the architecture of the system. It organizes the interplay between the different system elements and thus, determines the principle limits for the performance of the system. In this contribution, we focus on important properties of hierarchical cognitive systems. Therefore, we first present a framework for modeling hierarchical systems. Based on this framework, we formulate and discuss some crucial issues that should be treated explicitly in the design of a system. On this basis, we analyze and compare several well-established cognitive architectures with respect to their internal structure.

Cognitive Development in Partner Robots for Information Support to Elderly People

Yorita, A. Kubota, N. Page(s): 64 - 73 (pdf)

Abstract: This paper discusses an utterance system based on the associative memory of partner robots developed through interaction with people. Human interaction based on gestures is quite important to the expression of natural communication, and the meaning of gestures can be understood through intentional interactions with a human. We therefore propose a method for associative learning based on intentional interaction and conversation that can realize such natural communication. Steady-state genetic algorithms (SSGA) are applied in order to detect the human face and objects via image processing. Spiking neural networks are applied in order to memorize the spatio-temporal patterns of human hand motions and various relationships among the perceptual information that is conveyed. The experimental results show that the proposed method can refine the relationships among this varied perceptual information that can then inform an updated relationship to natural communication with a human. We also present methods of assisting memory and assessing a human's state.

Dynamic Neural Fields as Building Blocks of a Cortex-Inspired Architecture for Robotic Scene Representation

Zibner, S.K.U. Faubel, C. Iossifidis, I. Schoner, G. Page(s): 74 - 91 (pdf)

Abstract: Based on the concepts of dynamic field theory (DFT), we present an architecture that autonomously generates scene representations by controlling gaze and attention, creating visual objects in the foreground, tracking objects, reading them into working memory, and taking into account their visibility. At the core of this architecture are three-dimensional dynamic neural fields (DNFs) that link feature to spatial information. These three-dimensional fields couple into lower dimensional fields, which provide the links to the sensory surface and to the motor systems. We discuss how DNFs can be used as building blocks for cognitive architectures, characterize the critical bifurcations in DNFs, as well as the possible coupling structures among DNFs. In a series of robotic experiments, we demonstrate how the DNF architecture provides the core functionalities of a scene representation.

Visual Attention for Robotic Cognition: A Survey

Begum, M. Karray, F. Page(s): 92 - 105 (pdf)

Abstract: The goal of the cognitive robotics research is to design robots with human-like cognition (albeit reduced complexity) in perception, reasoning, action planning, and decision making. Such a venture of cognitive robotics has developed robots with redundant number of sensors and actuators in order to perceive the world and act up on it in a human-like fashion. A major challenge to deal with these robots is managing the enormous amount of information continuously arriving through multiple sensors. The primates master this information management skill through their custom-built attention mechanism. Mimicking the attention behavior of the primates, therefore, has gained tremendous popularity in robotic research in the recent years (Bar-Cohen , Biologically Inspired Intelligent Robots, 2003, and B. Webb , Biorobotics, 2003). The difficulties of redundant information management, however, is the most severe in case of visual perception of the robots. Even a moderate size image of the natural scene generally contains enough visual information to easily overload the on-line decision making process of an autonomous robot. Modeling primates-like visual attention mechanism for the robot, therefore, is becoming more popular among the robotic researchers. A visual attention model enables the robot to selectively (and autonomously) choose a “behaviorally relevant” segment of visual information for further processing while relative exclusion of the others. This paper sheds light on the ongoing journey of robotics research to achieve a visual attention model which will serve as a component of cognition of the modern-day robots.

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