

# Preface: Toward a Growing Computational Replica of the Human Mind

**Alexei V. Samsonovich**

George Mason University

**Shane T. Mueller**

Klein Associates Division, ARA Inc.

The fields of artificial intelligence, machine learning, machine vision, and robotic control have all produced successes that demonstrate superiority of machines over the human for narrow problems, and have even sometimes performed tasks that are impossible for humans. Despite these successes, a true human-level machine intelligence remains a far-off dream. Nevertheless, people of the 21st century are looking forward to new intelligent artifacts: partners rather than tools, machines that may become useful members of the human society or replace humans at the frontiers. Are we ready to take this challenge seriously?

A contraargument goes like this. Producing a replica of a Picasso painting does not make the painter a Picasso. Similarly, a capable but brittle machine can be engineered to take on some cognitive tasks of a human, but it may fail to work in an unexpected domain. Yet the current approach to engineering new technology is focused on incremental expansion and integration of machine replicas of human intelligent capabilities — e.g., replacing traffic cops and housemaids with intelligent traffic lights and Roombas. Extrapolating on this trajectory, it seems that future artifacts are bound to continue automating more and more cognitive tasks that were previously considered as exclusively human expertise. This process results in high-cost state-of-the-art solutions for narrow problems, while even the most intelligent of these systems will never be on an equal scale with humans. This approach is unlikely to produce a chain-reaction factor that could change our civilization in several years.

Yet, there is an alternative to the approach of highly-engineered incremental development and integration. A way to understand the flavor of this alternative is by looking at what we may expect from it, and how it is different from the above. Our main expectation for the new technology is its low cost: it should make most human-level cognitive abilities available at a much lower price than they are available today (in their natural or machine replica form).

How? By designing the systems so that they can learn on their own, set their own learning goals and develop their own skills in new and unanticipated domains. As a result, artifacts will be able to grow up, taking care of themselves in their own design, education, development and deployment, with minimal and gradually vanishing over time cost of required human interventions during this process! In other words, artifacts will be able to decide what abilities to develop and how, and achieve their goals autonomously, mostly using existing hardware and information resources. The bottom line is that society can and should afford development of useful intelligent artifacts, if those artifacts will be capable of self-sustainable cognitive growth up to the human level.

Understood in this sense, the challenge becomes very attractive. The challenge is to identify and to create the *critical mass* of artificial intelligence that will be able to start the process of self-sustainable cognitive growth, plus to design the *ladder* (the sequence of goals, or the curriculum) that artifacts will follow in their human-assisted, human-like development. The success will be measured by the law of decay of the cost of newly developed human-level cognitive capabilities.

The only existing cognitive system that is known to possess the above desired properties is the human brain. Therefore, our hope is that an approach based on biological inspirations will be successful. Accordingly, the objective of this symposium is to explore how the development of biologically-inspired cognitive architectures (BICA) may allow for a deeper understanding of the biological underpinnings of human cognitive and learning skills, and therefore move the field of artificial intelligence beyond capable but brittle systems, to a level that is robust, flexible, and useful in unanticipated domains.

The general focus of the symposium is on the challenge of creating a computational equivalent of the human mind taken in its highest cognitive abilities: specifically, the

ability to grow cognitively, virtually in any domain of human expertise, including unrestricted social interactions with humans. This fundamental scientific problem calls for the design and experimental study of BICA. Despite a substantial number of recent conferences and workshops focused on computational theories of biological information processing, this specific topic of research remains underrepresented. It is therefore our view that a forum, under the auspices of the AAAI 2008 Fall Symposia, dedicated to the topic of BICA is both timely and beneficial in stimulating further research in this field.

BICA include intelligent systems that use neurally-plausible information processing mechanisms, engage in ubiquitous learning and autonomous exploration, are embodied in a virtual or physical and social world, can accomplish real tasks through instruction or observation, integrate functions across multiple complex perceptual and motor systems, etc. As such, these are intelligent systems that operate like the human brain-mind. Efforts to implement, integrate, and measure the effectiveness of such biologically-inspired cognitive systems are at the primary focus of the symposium.

Studies of implemented embodied BICA are represented by several papers in this report (Barone et al., Noelle, Lebiere et al., and Goertzel & de Garis). The same topic, with more focus on industrial and military applications, is covered by related papers (Khosla & Huber, Srinivasa & Grossberg, and Bhattacharyya et al.). At the same time, theoretical discussion of the underlying mechanisms receives equal attention (Kandefor & Shapiro, Grossberg, Perlis, Epstein, Wyatt & Hawes, Hammerstrom & Zaveri, Thomsen, Hall, and Sanz et al.).

From the point of view of building human-compatible BICA possessing human-level social intelligence and capable of cognitive growth in a human society, natural language acquisition and processing capabilities are vital. Therefore, a large subset of works included in this symposium are focused on language, where using BICA can provide new solutions to currently burning problems in natural language processing and analysis (Kandefor & Shapiro, Boschee et al., Sierra & Santibáñez, Malaia & Wilbur, Kogut et al., Madani, Rohrer, Chen & Fahlman, Sikos et al., and Nirenburg et al.).

The term “biologically inspired” can be reinterpreted as “biologically constrained”. Indeed, biological constraints on human cognition informed by neuroscience become most valuable, when it comes to a computational description of higher cognition. Understanding vital biological constraints and their leverage in information processing is a key to success in the so-called “bottom-up”, or “reverse-engineering” approach in BICA research. A number of works included in this symposium address biological constraints, together covering a hierarchy of levels: from individual neuronal information processing principles to large-scale connectivity and functional organization of the human brain (Albus, Schneider et al., Anderson, Lörincz, Achler & Amir, and Ventura).

This enterprise of bottom-up exploration and smart exploitation of vital biological constraints extends further into the field of cognitive psychology, addressing questions directly related to the challenge. Here, the topic of human-like memory and learning and their computational models is intended as the specific focus of the symposium. It is covered by several interrelated works: in experimental cognitive psychology (Gardner et al.), in cognitive modeling of human-like memory and learning (Brom & Lukavský, Chella et al.), and in theoretical and computational modeling of knowledge representation in the human brain, human reasoning and expertise (Gómez et al., Staszewski, Eilbert, Gero, and Gros & Kaczor; see also Freedman, Sanz, and Huyck).

One particular topic in studies of humanlike learning deserves special attention in the context of this symposium. It is represented by works that use models of self-regulated learning (SRL) for applications in intelligent tutoring systems (Azevedo & Witherspoon, Pirrone et al., and Samsonovich et al.). One problem with bringing ideas and results from SRL studies to artificial intelligence is that theoretical research in educational science is done at a relatively abstract, human-oriented, functionalist level. The ambitious goal is to translate educational facts and SRL models to a precise computational format. This goal will be achieved with the implementation of the proposed new intelligent tutoring systems.

Emotions and human values play a key role in SRL and will play a key role in human-level intelligent systems. Several works included in this symposium are devoted to this topic, including theoretical and BICA-based cognitive modeling studies of emotions (Hudlicka), experimental studies of emotions (Rohlfshagen & Bryson) and theoretical analysis of the emergence of ethical values in a society of agents (Waser).

Finally, we come to the central question: how to put all efforts together, and how to validate the results? This question is addressed essentially by three contributions to the symposium: one describing an integration framework for testing BICA (Pope & Langley), one addressing adaptation of traditional intelligence tests for artifacts (Mueller & Minnery), and one presenting a feasible roadmap to solving the challenge (Oblinger).

With all these necessary components covered by papers and associated with them talks, we hope that the main value of the symposium will be determined by the discussion panels, where ideas of new approaches and new research projects may be born. The symposium program includes three discussion panels intended as brainstorming. We would like to end this introduction with the list of topics – open questions assigned to the panels.

## **Discussion Panel 1**

Embodiment, language and the critical mass of intelligence. (a) How to define a threshold level of cognitive abilities (the critical mass: architecture and interface) that will enable autonomous cognitive growth of the agent? (b) How to design a successful cognitive growth scenario (the ladder),

from simple egg hunt to a general human level of intelligence?

### **Discussion Panel 2**

“B” in BICA and the natural cognitive growth of BICA. How do biological constraints help us to build flexible, more robust, human-compatible cognitive systems?

### **Discussion Panel 3**

The challenge, metrics, and the roadmap for developing an autonomously growing human-level intelligence. What can we conclude from this symposium? What should be our strategy and research plan for the future?

We believe that this symposium will help all of us to ignite, coordinate and support efforts aimed at creating human-compatible, autonomously growing intelligent cognitive systems. We would like to thank all contributors for their hard work and their support of the idea of this symposium. We welcome all committed participants to this highly stimulating and exciting event: the first AAAI Fall Symposium on Biologically Inspired Cognitive Architectures.

– *Alexei V. Samsonovich*  
– *Shane T. Mueller*