Bootstrapping Semantics in an Autonomic Computing System

Sam S. Adams, Nancy Alvarado, Steve Burbeck, Craig Latta
IBM, Thomas J. Watson Research Center
{ssadams, nanalv, sburbeck, crl}@us.ibm.com

Abstract

Autonomic Computing Systems (ACS) are envisioned to function with the robustness of self-regulating biological systems. They are self organizing, self configuring, self optimizing, self healing, self protecting, aware of both themselves and their environment, and capable of adapting to circumstances unknown to their designers while they continue to deliver their specified behavior. These capabilities fundamentally require an ACS to develop and maintain some level of understanding of both itself and its environment over time. In this paper, we focus on the bootstrapping of semantics in such systems by comparing this process to the bootstrapping of the conscious mind in the human infant. We highlight the need for the pervasive integration of emotional control and present an overview of the Joshua Blue architecture for semiosis in autonomic computing systems.

Introduction

The complexity of large networked computing systems is rapidly heading toward a point of diminishing returns, where the costs of both IT personnel and system downtime will threaten to eclipse the incremental business value promised by the application designers. A straightforward approach to this problem has been proposed by Paul Horn, Senior Vice President of IBM Research: design systems that can take care of themselves by emulating the adaptive and intelligent behavior of complex biological systems. [Horn, 2000] These so-called Autonomic Computing Systems (ACS) are thus expected to demonstrate a wide range of capabilities previously only observed in higher life forms such as humans and other mammals. Achieving this vision will require many innovations in both technology and development methodologies, but one requirement stands out as particularly challenging. These systems must not only deliver their expected behavior within the range of operating conditions anticipated by their designers, but they must also successfully adapt to unanticipated and extraordinary changes in their environment. Regardless of the amount of explicit knowledge preprogrammed into these systems, they will need to develop their own understanding of both the changes in their operating environment and the effect of their own actions on that environment. This is essential if they are to make more than blind choices as they adapt their internal behavior. An ACS capable of this level of adaptivity must achieve semiosis with its environment. In fact, we believe that it must also achieve semiosis between its embodiment and itself.

We take the position that since human beings are the most adaptive and the most studied complex biological systems known, they are most likely to provide the best models for emulation in the design of ACS. While it is not the goal of Autonomic Computing to achieve adult levels of human intelligence and cognition within computing systems, we believe that the cognitive behavior of human toddlers around three years of age provides a useful target for this effort. In addition, we have found the developmental process that results in the bootstrapping of semantics in the human infant to be particularly rich in both design metaphors and algorithms as we attempt to develop adaptive semiotic control systems for Autonomic Computing. Examples include the dramatic changes that occur in neocortex synaptic density in infants around 2-3 months of age [Huttenlocher, 1994], the staged onset of competency in sensory/motor skills [Turkewitz and Kenny, 1982], the development of attention [Posner and Rothbart, 1980], and the acquisition of cause/effect knowledge [Chaput and Cohen, 2001].

We believe that key aspects of the semantic bootstrapping process that occurs in human infants can be emulated in computer systems to provide similar semantic capability. In this paper, we describe the bootstrapping of flexible, self-defined semantics in the human mind as the result of numerous emergent behavioral metalevels beginning at conception. We describe what we term the initial conscious level where such semantic flexibility is first possible and explain how the capabilities of such a system can guide the design of ACS. We then discuss the necessity for emotional control in humans even at this primitive cognitive level and argue its necessity in ACS as well. Finally, we briefly describe the Joshua Blue project, its architecture for semiosis, and suggest an approach for testing the semantic capabilities of ACS.

Bootstrapping Semantics via Experience

At conception, our lifetime of experience begins at a primitive biochemical level. As cell growth and differentiation ensues, our experience becomes more varied as newly specialized cell types emerge with different biochemical behavior and sensitivities. After only a few days of development, ensembles of similar cells generate emergent behaviors that start the blastosphere on the road toward tissue and organ development. These tissues interact at a supercellular level, generating yet another metalevel of behavior and sensitivities. Feedback loops abound at each level, from intracellular chemistry to systems of organs, with numerous interactions and interdependencies between the elements of each level, as well as interactions across these emergent levels of behavior. The dynamical behavior at each level bootstraps the emergent capabilities for the next level. The goal of individual cellular homeostasis becomes dominated by tissue homeostasis, which gives way to organ homeostasis that eventually yields to the overall goal of organism homeostasis. Individual elements at any one level may be suboptimized or even sacrificed for the optimization or survival of a higher level.

At some unknown point during this process, a level of behavior and sensitivity emerges that permits a system to become reflectively aware, or conscious, of its own state. When this happens, the system is no longer simply reactive to the changes in its environment. The initially conscious level enables it to be proactive, to intentionally manipulate its environment to improve its performance in meeting its needs. The system now begins the ongoing process of learning to understand itself and its changing environment well enough to continually improve on its ability to meet its goals.

Emergent Behavior and the *Tabula Rasa*

Given the bootstrapping process above, what can we say about the character of the behaviors and the knowledge structures they generate at each emergent metalevel? At all levels, each new layer of emergent behavior innately provides a metaontology for the knowledge that can be generated and utilized at that level. This metaontology is directly encoded in the behaviors themselves and knowledge generated at this level is restricted by the execution of those behaviors. The knowledge structures, both genotypes (classes) and phenotypes (instances), generated at each level are themselves compliant phenotypes of the metaontology encoded in the emergent behavior of the previous level. This implies that while there is lower level encoded knowledge in existence at the birth of a new emergent level, there is no *a priori* knowledge at the new level until it is generated via experience. So, while it is true that there is no *tabula rasa* when looking across metalevels, it is also true that each new level of emergent behavior begins its operational life with a clean slate, at least at its own level.

Conscious Behavior: Escaping the Shackles of Genetics

What makes consciousness so different from the lower levels of behavior that support it? A computer analogy provides useful insights. The behavior of any computer is constrained by a combination of its hardware, software, and the information available to be processed. One difference between hardware behavior and software behavior is that hardware behavior is grounded in the physics of the electronics, while software behavior is grounded in the "physics" or behavior of the hardware. This analogy can be extended to the information processed by the software as well, since the resulting information output from the system is constrained or "grounded" in the "physics" of the software.

This situation is analogous to the emergent behavioral levels discussed above in that each prior level constrains the behavior of the next because it provides the implementation of the primitive functions of that behavior [Joslyn, 2001]. From this limited perspective, the main difference between the computer system and the biological mind is that the behaviors and knowledge at each level of the computer example are directly encoded by their designer. They are not emergent. Of course, there are software techniques for generating data and software, and even techniques for generating new hardware designs, but with very few exceptions, there are no new levels of behavior and knowledge generated, only modifications to the existing levels.

While software does indeed implement behavior and orchestrate the hardware functions to support a specific information processing application, software itself *is* information. This allows applications to be developed whose ongoing behavior is the result of self-modifying software, where the information produced by a software program becomes part of the program itself and changes its future behavior. The computer science concepts of "interpreters," "virtual machines" and “genetic algorithms” are examples that exploit this property of software. What does this have to do with consciousness? We believe that the major difference between the first conscious level and the previous levels on which it rests is that this is the point where the system's behavior escapes its "genetics.” For the first time, awareness of and reflection upon its internal states produces intentional changes in its own "programming" to improve its ability to meet its needs. This new capability produces a major qualitative difference in behavior at the initially conscious level, and provides the fundamental architecture for the layers of behavior and knowledge that emerge throughout its experience. We believe these new emergent layers constitute what we know as human memory, learning, and attended cognition. We believe that the layer directly below the initial conscious level represents the sensor and effector interface between the "mind" and the body and its environment. The task of anchoring cognition to environmental interactions falls directly on this initial conscious level.

Semantics and the Emergence of Consciousness

The arrival of flexible, self-defined semantics marks the transition from the lower, preconscious metalevels of behavior to this first metalevel of initial consciousness. Prior to this level, the behavior at each lower level was predetermined by the chemical and physical reactions between the structures that composed it and its supporting behavioral metalevels. While one can describe these reactions and systems of reactions in semiotic terms, their ontologies and dynamics are strictly limited to the set of possible reactions and stable chemical states that are allowed by the laws of physics and chemistry. A system at this level has no choice in how to conceptualize its experience or how to interpret the molecular “signs.” Such a system cannot develop alternative models of events or generalize across similar experiences. It cannot anticipate future needs, form expectations, generate goals and select appropriate plans of actions. It has no way of evaluating whether some current state or event has a positive or negative effect on its existence. A system at this level simply reacts to its environment, and has no choice in the matter.

When the initial conscious level of behavior emerges, the system has an information processing capability that is no longer limited by a fixed, predetermined ontology. New semantics are generated to represent different aspects of experience, hypothetical cause and effect relationships between events, proposed generalizations for reoccurring phenomena (both internal and external), and potentially effective plans of action. Since these representations are themselves information generated by the emergent dynamics of the initially conscious level, the number and kinds of semantic structures generated by the system are only bound by its storage capacity. In addition, the potential number of different ontologies that the system can now support is also bound only by its storage capacity for information. A system at this level can now overcome the semantic shortcomings of its previous supporting behavioral levels. It can capture representations of its subjective experience over time. It can invent alternative models of that experience, even by trial and error, that prove more accurately predictive of its future states. Given a means of subjective evaluation such as emotional response, the system can select which of these alternative models should most influence its behavior in the future as it strives to reach both its innate goals and acquired subgoals.

Semantics in Autonomic Computing Systems

Computer systems are typically not the products of many levels of emergent behavior, although anyone who has observed the sometimes-chaotic process of large-scale system integration projects may argue this point. Traditional computer systems are explicitly designed and developed to provide specific information processing and control behavior within a predetermined operational context. And yet, as Freeman [Freeman, 2000] states, “Truly flexible and adaptive intelligence cannot operate in the real world without the construction of meaning.” As mentioned earlier, an ACS must be able to adapt to previously unanticipated changes in its environment. How can we apply what we have learned from the semantic bootstrapping process observed in humans to the design of ACS? First, we can provide these systems with sufficient monitoring and feedback capabilities to provide the minimum connectivity requirements for semiosis, an analog of the sensory/motor connection human beings have between their bodies and their environments. Second, to ensure that ACS can understand the effects of their actions in novel environments, a feedback loop must be included to notify the system of its own actions and internal states, analogous to proprioception in biological systems. Third, the ACS must continuously be able to generate, test, reinforce and discard conceptual models of its experiences. Finally, the control logic for the ACS must be designed to both evaluate these models and to choose the most promising model to guide it as it interacts with its environment. These last two steps require some form of innate self evaluation, and that is where we believe something akin to an emotional control system will be crucial for fully Autonomic Computing.

The Need for an Emotional Control System

In humans and other living organisms, emotion provides a means of structuring knowledge and a control system for emergent behavior. As long as an organism responds in a fixed manner to its environment, with an invariant relationship between elicitors of behavior and behavioral responses, there is no need for consciousness or for emotion because there is no flexibility in what may occur. The value of consciousness is that it permits greater behavioral choice in order to maximize chances for survival and optimize quality of life for an organism. However, the organism must be guided in that choice. There must be some link between expected positive outcomes and the many options that confront an organism when behavior is uncoupled from environmental elicitors. Prior experience, with a memory for outcomes of previous action, provides one source of guidance for future behavioral choices. Salience of goals and association of goals with cues signaling presence of opportunities for meeting them (sometimes called affordances) provides another source of guidance. In each of these cases, emotion structures memory and guides attention by determining salience of environmental cues. It structures choice by identifying the significance of information with respect to survival goals [Clore, 1992] and it structures memory and thereby knowledge by organizing representations based upon their meaning to the organism.

When an organism operates with multiple levels of behavioral response, it requires multiple levels of control. Humans are capable of responding automatically, as more primitive creatures do, and this characterizes much of human behavior. Examples include responses to pheromones, reflexive response to sources of pain, orienting the head and eyes in response to motion, and so on. Humans are also capable of making choices outside of conscious awareness. This lower level of emergent behavior involves choice guided by affect (emotional influence), not by consciousness. Examples include the biasing of memory, attention, and decision-making by mood, preferences, and familiarity. On an exception basis, unconscious emotional control of ongoing choice behaviors becomes the focus of conscious awareness. At this point, automatic governance by affect is interrupted and reason or other choice factors may be applied.

In human beings, affect has a biologically fixed relationship to physiology but a variable relationship to both environmental events and cognitions, including internal knowledge representations. Emotion arises along with psychological drive states in response to dysregulations of metabolism (hunger, thirst, injury, illness, fatigue). When these states arise due to environmental circumstances and behavioral choices, emotion becomes associated with those causative and predictive events through the mechanism of classical conditioning. With experience, cognition becomes an additional elicitor of affect. Thus the emotions that arise in different circumstances are a result of previous experiences and their consequences for survival. In humans, the same affects that are attached to actual experiences become attached to the mental representations of those experiences and form the basis for interpretations of their significance. Thus affect is intertwined with expectations, beliefs, attitudes, stereotypes and schemas, and the full range of mental representations (including sensory and nonverbal representations).

A consequence of this ability of affect to structure choice and representation is that it may be the mechanism for the "first distinction," the formation of the first emergent meaning that is represented and that guides behavior independent of a fixed response to an elicitor. In humans this presupposes that an infant responds to bodily sensations, has an innate preference for pleasure and a motive to avoid pain, and a functioning aversive/appetitive system at the neural level [Panksepp, 1998]. We believe that an ACS will need similar capabilities, albeit adapted to the domain of computer systems. In humans, the formation of meaning arises from the association of second-order significance with the events or actions that accompany or result in the sensations an infant can experience. The first meaning answers the self-query "what is the significance of this event to me?" and the answer consists of pain or pleasure associated with the experience. The cognitive representation and abstraction of pain or pleasure becomes an evaluation that something is positive or negative, good or bad in terms of one's own experience of it. Once the association is formed, the infant need not experience the actual pain- or pleasure-causing event, but will respond to the cue that signals it as if the event itself had occurred. From there it is a short step to engaging in preventative action to avoid occurrence of an aversive event (if only by crying in anticipation of pain).

With the acquisition of voluntary control over muscles, the infant has both a guidance system for behavioral choice and the means to carry out its choices. As Sutton [Sutton, 2002] states: "Any organism with the potential to selectively move about its environment must be able to organize mental, physiological, and environmental resources in order to turn incentives into rewards and keep threats from becoming punishments." We believe that the first meaning arises in service to that goal, as an expression of the behavioral flexibility made possible by the architecture of the human mind. We also believe that for ACS to achieve this level of open-ended semiosis, they will need to pervasively integrate similar emotional behavior.

Emotion, the Self, and the Emergence of Morality

Development of a sense of self permits a child to more richly conceptualize consequences to that self. Ego threats take on importance equivalent to physical threats and the affect arising from damage to the ego (e.g., shame or embarrassment, insult, loss of self-esteem) is described using pain as its dominant metaphor. A sense of self permits distinction between self and others. It becomes a heuristic for anticipating the actions of others and for explaining their behavior. In that process, it forms the basis for empathy. As one takes on the experience of another person via imagination, the same affects arise as would be felt during that experience. These affects motivate the same kinds of corrective or preventative action and become the basis for empathy-guided helping behavior. Affiliating the self with others permits association of emotional responses with the welfare of groups, including in humans the family, friends, ethnicities and nationalities and other social entities such as sports teams. These in turn motivate behaviors with consequences beyond the self and form the basis for morality. ACS that participate in multiple groups will need to manage multiple, possibly conflicting moralities that govern its interaction within each group. We speculate that this juggling act could even lead to a kind of situational ethics for ACS.

Moral behavior in autonomous systems functioning under changing circumstances more likely relies upon endowing a system with a sense of self, than it does on prescriptions such as Asimov's Three Laws of Robotics. Not only must a human, robot, or autonomous system know what constitutes morally correct behavior in complex situations, but it must be motivated to engage in it, even when there may be a negative consequence to the self. That requires an affective attachment to an identification that extends beyond the self.

Joshua Blue: An Architecture for Semiosis

The Joshua Blue project has been focused on developing a control system, patterned after the human mind, capable of autonomously learning to function in any number of embodiments and environments. We have based this work on research findings in developmental psychology and developmental neuroscience, focusing on the bootstrapping of the human mind within the first twelve months after conception and its development up to the onset of sufficient language acquisition to allow for complex verbal and written instruction at approximately 3 years of age.

The architecture for the Joshua Blue system is shown in Figure 1. It includes sensors and effectors (actuators) for interacting with the environment, an embodiment that provides an appropriate connection with both internal state variables and control parameters of an application, and the module performing the control functions, the system's "mind." The mind monitors and interprets sensory information provided by the body and executes appropriate actions via the body's actuators and other functional modules. All cognition takes place in the context of this sensory information, ensuring that the system remains continually grounded [Harnard, 1990].

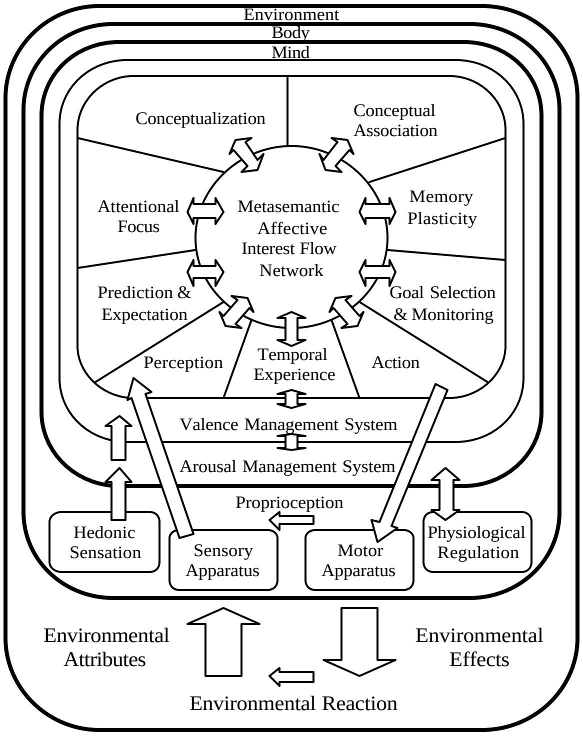

Figure 1. Joshua Blue Architecture for Semiosis

Affect is involved in two feedback loops. The first regulates internal functioning of the mind. The second regulates interactions with the environment and thus behavioral choice. The use of the same affective state variables in both feedback loops provides a connection between the mind, the body, and the world. The two regulatory state variables are "arousal," which indicates the general activity level of the system, and "valence," which represents the system's self-evaluation of its current state and performance. Changes in these state variables motivate changes in behavior.

The backbone of the Joshua Blue architecture is the Metasemantic Affective Interest-flow Network (MAIN). A series of subsystems providing "innate" cognitive behaviors interact with the MAIN. The MAIN is a highly specialized semantic network that supports the open-ended creation and association of concepts at any number of conceptual metalevels. Continual spreading activation based on an interest-oriented model of attention provides the emergent dynamics of the cognitive system. These dynamics are then modulated and steered by the simulated emotion system that performs the functions of prioritizing goals, structuring experience and guiding behavior, as emotion guides living organisms [Alvarado, Adams, Burbeck and Latta, 2002]. The resulting dynamic semantic network supports an ongoing process of entailment [Heylighen, 1997] that both bootstraps and continually adapts the system's conceptual model of both itself and its environment.

This architecture provides for semiosis with the environment in a manner similar to Meystel's "Six-box diagram of learning" [Meystel and Albus, 2002], but differs markedly in its internal design due to the MAIN's unified representation. The pervasive integration of emotion and emotional control are also distinctive features, as is the proprioceptive feedback loop that enables semiosis between the system's "mind" and its "body", a capability we believe essential to the emergent sense of self.

Testing for Acquired Semantics

In humans, the ultimate test for acquired semantics is success in adaptation and survival. Given the vast differences in the embodiments and environments between humans and ACS, specialized forms of testing will be required to determine progress and success in acquiring semantics and achieving semiosis. The literature of developmental psychology provides numerous well-developed experiments for measuring the cognitive capabilities of both humans and animals. We believe that adaptations of these experiments such as classical conditioning, operant conditioning, and social learning should be used as the standard for measuring both the intelligence and cognitive capabilities of computer systems [Alvarado, Adams, Burbeck and Latta, 2002]. To determine whether a system has acquired appropriate semantics, tests of acquired regulatory control are essential. These experiments must be tests of the acquisition of meaning via experience. While the values of key system variables can be inspected, observing the effects of learning exemplified in behavior change will provide the most reliable approach. The system is tested by first exposing it to a specific learning environment, then placing it in a carefully constructed context designed to demonstrate whether meaning has been acquired and whether it effectively governs the system's behavior. Among the important behaviors that must be demonstrated are: (1) ability to recognize and take corrective action based upon presence or absence of environmental signs; (2) ability to recognize and remember the effects of its own behavior; (3) ability to form goals and expectations and intervene only when change will be effective; (4) ability to define itself and others using salient characteristics or behavior; (5) ability to act in a self-preserving manner in contexts that threaten system performance or integrity.

Conclusion

For computer systems to reach the goal of Autonomic Computing, they will need to develop and maintain semiotic relationships with both their operational environments as well as their own "embodied implementation." We believe this will require either the direct design or emergence of a level of behavior that exhibits a primitive consciousness capable of self-evaluation and informed choice of action. This initially conscious level must be capable of open-ended generation of candidate ontologies with which to model its experiences, yet remain thoroughly grounded to the realities of its environment as it tries to reach its goals. Testing such systems will likely require the development of specialized versions of classic experiments from developmental psychology

References

Alvarado, N., Adams, S., Burbeck S., Latta C. (2002) "Beyond the Turing Test: Performance Metrics for Evaluating a Computer Simulation of the Human Mind", The 2nd International Conference on Development and Learning 2002 (ICDL'02), in submission.

Alvarado, N., Adams, S., Burbeck S., Latta C. (2002) "Integrating Emotion and Motivation into Intelligent Systems", IBM Systems Journal, in submission.

Asimov, I. (1950). I, Robot. New York: Gnome Press.

Chaput, H. H. & Cohen, L. B (2001). A Model of Infant Causal Perception and its Development. *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (pp. 182-187). Mahwah, NJ: Lawrence Erlbaum.

Clore, G. (1992) Cognitive phenomenology: Feelings and the construction of judgment. In L. Martin and A. Tesser (Eds.), The construction of social judgments (pp. 133-163). New York, NY: Lawrence Erlbaum Associates.

Freeman, W.J. (2000) A neurobiological interpretation of semiotics: meaning, representation, and information. Information Sciences 124(1-4): 93-102

Harnad, S. (1990) The Symbol Grounding Problem. Physica D 42: 335-34

Heylighen, Francis (1997) Bootstrapping knowledge representations: from entailment meshes via semantic nets to learning webs. Kybernetes 30 (5/6)[in press]

Horn, P. (2001). “The Autonomic Computing Manifesto”, http://www.research.ibm.com/autonomic/manifesto/autonomic_computing.pdf

Huttenlocher, P. R. (1994) Synaptogenesis, synapse elimination, and neural plasticity in human cerebral cortex. In C. A. Nelson (Ed.), Threats to optimal development: The Minnesota Symposium on Child Psychology, Vol. 27 (pp. 35-54). Hillsdale, NJ: Erlbaum.

Joslyn, C. (2001) "The Semiotics of Control and Modeling Relations in Complex Systems", Biosystems, v. 60:1-3, pp. 131-48, 2001

Meystel, A., Albus, J. (2002) Intelligent systems : architecture, design, and control (p. 93).New York : Wiley.

Panksepp, J. (1998) Affective neuroscience: The foundations of human and animal emotions. New York: Oxford University Press.

Posner, M. I., & Rothbart, M. K. (1980). The development of attentional mechanisms. Nebraska Symposium on Motivation, pp. 1-52.

Sutton, S. (2002). Incentive and threat reactivity: Relations with anterior cortical activity. In D. Cervone & W. Mischel (Eds). Advances in Personality Science (pp. 127-150). New York: Guilford Press.

Turkewitz, G. Kenny, P. A. (1982). Limitations on input as a basis for neural organization and perceptual development: A preliminary theoretical statement. Developmental Psychobiology, 15, 357--368.