Journal of Artificial General Intelligence 4(2) 1-22, 2013
DOI: 10.2478/jagi-2013-0002

Submitted 2013-04-09
Accepted 2013-07-08

Conceptual Commitments of the LIDA Model of Cognition

Stan Franklin FRANKLIN@MEMPHIS.EDU
Steve Strain SFSTRAIN@MEMPHIS.EDU
Ryan McCall RMCCALL@MEMPHIS.EDU
Fedex Institute of Technology 301
The University of Memphis
Memphis, TN 38152, USA

Bernard Baars BAARSBJ@GMAIL.COM
6615 Fisher Ave.
Falls Church, VA 22046, USA

Editor: Haris Dindo, James Marshall, Giovanni Pezzulo

Abstract

Significant debate on fundamental issues remains in the subfields of cognitive science, including perception, memory, attention, action selection, learning, and others. Psychology, neuroscience, and artificial intelligence each contribute alternative and sometimes conflicting perspectives on the supervening problem of artificial general intelligence (AGI). Current efforts toward a broad-based, systems-level model of minds cannot await theoretical convergence in each of the relevant subfields. Such work therefore requires the formulation of tentative hypotheses, based on current knowledge, that serve to connect cognitive functions into a theoretical framework for the study of the mind. We term such hypotheses "conceptual commitments" and describe the hypotheses underlying one such model, the Learning Intelligent Distribution Agent (LIDA) Model. Our intention is to initiate a discussion among AGI researchers about which conceptual commitments are essential, or particularly useful, toward creating AGI agents.

Keywords: asynchrony, cognitive cycle, cognitive model, commitments, consciousness, embodied, Global Workspace Theory, learning, LIDA, memory, motivation, non-linear dynamics, theta-gamma coupling

1. Introduction

Not long after its inception, artificial intelligence abandoned its original aim of reproducing human-level intelligence in favor of developing highly practical systems that behave intelligently in narrow, however important, domains. After a half century, a movement in AI research toward that original quest has emerged under the rubric of *artificial general intelligence* (Goertzel & Pennachin, 2007; Wang, Goertzel, & Franklin, 2008). After an initial invitational workshop in 2006 (Goertzel & Wang, 2007), five successful AGI conferences have been held in several locations in Europe and the United States (de Garis & Goertzel, 2009a; de Garis & Goertzel, 2009b), A recent one hosted over 200 researchers on Google's main campus. The last two conference proceedings were published in Springer's *Lecture Notes in AI* book series (Bach, Goertzel, & Iklé, 2012; Schmidhuber, Thorisson, & Looks, 2011). Additionally, *The Journal of Artificial General Intelligence* has published several volumes.

A parallel movement flies under the rubric of BICA (Biologically Inspired Cognitive Architectures). First appearing as several AAAI symposia carrying that name (Samsonovich, 2008), the movement has produced successful conferences of its own (Samsonovich &

Johannsdottir, 2011; Samsonovich, Jóhannsdóttir, Chella, & Goertzel, 2010), and has started a journal, Biologically Inspired Cognitive Architectures, published by Elsevier. The aims and scope declaration of the journal begins with the sentence, "The focus of the journal is on the integration of many research efforts in addressing the challenge of creating a real-life computational equivalent of the human mind." Note that the "BICA challenge," as it has come to be called, is quite equivalent to the goal of AGI.

Another such parallel movement was ushered in by the First Annual Conference on Advances in Cognitive Systems, held in Palo Alto, California in December 2012. Its call for papers asserts "The purpose is to provide a venue for research on the initial goals of artificial intelligence and cognitive science, which aimed to explain the mind in computational terms and to reproduce the entire range of human cognitive abilities in computational artifacts." A new online journal, also entitled Advances in Cognitive Systems, has published its second volume.

Yet another movement in this same direction is arising in the form of an AAAI Spring symposium entitled "Designing Intelligent Robots: Reintegrating AI." The Overview of its second incarnation (Spring 2013) includes the following lines: "AI is fragmented field: well-developed and largely independent research communities exist for learning, planning, reasoning, language, perception and control. Since the challenges posted by each of these subfields are immense, most researchers have found it necessary to devote their careers to specializing in a single subfield. While immense progress has been made in each of these subfields in the last few decades, it remains unclear how they can be integrated to produce an intelligent robot. Unifying these disparate technologies will open up new avenues of research and create new application opportunities. Therefore, we believe that integration should be considered a valid research endeavor in its own right." (Designing Intelligent Robots, 2012)

It would seem that the quest for a generally applicable, integrated AI system capable of human-level intelligence is an idea whose time has come. The question is how to go about it. Necessary components include human-level perception, action selection, and everything in between. This calls for controlling our generally applicable, integrated AI agent with a systems-level cognitive architecture capable of human-level intelligence. Many such cognitive architectures exist, though none, as of yet, at human-level intelligence (Samsonovich, 2010). Some of these architectures, for instance SOAR, ACT-R, CLARION, are decades old, while others are relatively new. A few have been developed specifically in response to the stated goals of the AGI and BICA movements. Each of these various architectures is based on conceptual commitments, which may have profound implications for their implementation, evaluation, and efficacy in controlling an agent at human-level intelligence.

Here we argue for a particular set of conceptual commitments as being potentially useful in the development of such an AGI architecture. Our argument can be summarized as follows: 1) An AGI agent must be built upon a systems-level cognitive architecture. 2) Each AGI research group develops and uses its own such architecture. 3) It is at least difficult, and likely impossible, to compare the efficacy of such architectures without building working AGI agents which, currently, seems not possible. Under these conditions, how might productive discussions and collaborations occur among different AGI research groups? Perhaps by specifying conceptual commitments made by the various architectures, and trying to agree upon a common set of commitments that all the research groups think will have to be made in order to successfully create an AGI agent. As a beginning move in this direction, we lay out below a first draft of the conceptual commitments of our LIDA model.

2. Conceptual Commitments

Since the scientific study of mental phenomena began with the work of Helmholtz, Wundt and James in the 19th century, abundant evidence has accumulated from research efforts in the distinct fields of experimental psychology, cognitive neuroscience and artificial intelligence. Although instances of dialogue between these fields are numerous, relatively little progress has been made toward a broad-based understanding of mind that incorporates findings from these separate avenues of study into a testable model that can explain the range of observed cognitive phenomena in the context of a cohesive comparative and evolutionary biology of cognition. (The one exception is the work done through ACT-R, e.g., (Anderson, 2007)). Much remains to be explored and decided before the assembly of such a model can be accomplished with exactness, but we find it imperative to engage in such modeling efforts in order better to evaluate current knowledge, and to guide further research, particularly research into possible AGI agent architectures.

To this end, we have made efforts to integrate the currently available evidence into a single broad, systems-level model of the mind, the LIDA model. Given the breadth of this task, several conceptual commitments have been taken. These are the fundamental and influential working hypotheses within the model, some of which are speculative and still open to significant debate. However, such theoretical activity guides empirical research and provides a means for comparing and communicating about the various cognitive architectures.

Our attempt to integrate evidence across disciplines at times stretches, or even breaks, conventions, perspectives, and assumptions that hold within the confines of a given discipline. Within a discipline, we often find it necessary to take positions with respect to ongoing controversies that many empirical researchers may find premature. However, such conceptual commitments are practical rather than conclusive. For instance, the hierarchical self-organization of brain rhythms plays an important role in our model, since (as will be elaborated below) it provides an elegant basis for several other model elements, including the cognitive cycle, process formation and selection, and a consciousness mechanism based on Global Workspace Theory (Baars, 1988; Baars & Franklin, 2003). Moreover, it constitutes a plausible explanation for the connection between brain and mind.

Nonetheless, we stop short of asserting that the current evidence base is sufficient to make a conclusive claim regarding the functional role of brain rhythms. We think of our model as an early map of a world largely unexplored, much like the 16th century world map of Martin Waldseemüller.[1] While empirical researchers can and should focus on details, broad-based modelers must make educated guesses to integrate what is known into a global framework. To make an analogy with cartography, the empiricist might attempt “an accurate coastline of Florida” while a broad-based modeler will focus on a rough outline of the “continents” (Bach, 2008). If new details proved inconsistent with a global map that was correct in the broad-based sense, it likely would not invalidate the entire map; rather, the map would need updating to accommodate the new data.

In this paper we first briefly review the LIDA model and its cognitive cycle. The subsequent sections provide brief descriptions, with some justifications, of several conceptual commitments that have guided the growth of our LIDA model of cognition, both the conceptual model whose historical development is described in our numerous published papers (e.g., Faghihi & Franklin, 2012; Franklin, Strain, Snaider, McCall, & Faghihi, 2012), and its underlying computational model as embodied in the LIDA Computational Framework (Snaider, McCall, Franklin, 2011). The conceptual commitments described in the sections below are presented in decreasing order of abstractness. Many of them can be thought of as hypotheses based either on the interpretation of empirical data from artificial intelligence, cognitive neuroscience or cognitive psychology, or on the needs of the model’s current implementation. Please note that we explicitly make no claim that any of the conceptual commitments described below are unique to the LIDA model. Likely, each can also be found elsewhere.

A list of these conceptual commitments, numbered according to their order of appearance as subsections of Section 4, is presented as a preview for the reader: 1) Systems-level Modeling, 2)

[1] See http://www.loc.gov/rr/geogmap/waldexh.html for an example.

Biologically Inspired, 3) Embodied (Situated) Cognition, 4) Cognitive Cycles as Cognitive Atoms, 5) Global Workspace Theory, 6) Learning via Consciousness, 7) Comprehensive Decay of Representations and Memory, 8) Profligacy in Learning, 9) Feelings as Motivators and Modulators of Learning, 10) Asynchrony, 11) Transient Episodic Memory, 12) Consolidation, 13) Non-linear Dynamics Bridge to Neuroscience, 14) Theta Gamma Coupling from the Cognitive Cycle.

3. The LIDA Model and its Cognitive Cycle

The LIDA model is a systems-level, conceptual and computational model covering a large portion of human cognition [2]. Based primarily on Global Workspace theory, the model implements and fleshes out a number of psychological and neuropsychological theories (see the Biologically Inspired section below). The design of LIDA was also influenced by the Copycat model of Hofstadter and Mitchell (1995). The LIDA computational architecture is derived from the LIDA cognitive model. The LIDA model and its ensuing architecture are grounded in the LIDA cognitive cycle. Every autonomous agent, be it human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). More sophisticated agents, such as humans, process (make sense of) the input from such sampling in order to facilitate their decision making. The agent's "life" can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle constitutes a unit of sensing, attending and acting. A cognitive cycle can be thought of as a moment of cognition, a cognitive "moment."

We will now briefly describe what the LIDA model hypothesizes as the rich inner structure of the LIDA cognitive cycle. More detailed descriptions are available elsewhere (Baars & Franklin, 2003; Franklin, Baars, Ramamurthy, & Ventura, 2005). During each cognitive cycle the LIDA agent first makes sense of its current situation as best as it can by updating its representation of its current situation, both external and internal. By a competitive process, as specified by Global Workspace Theory, it then decides what portion of the represented situation is the most salient, the most in need of attention. Broadcasting this portion, the current contents of consciousness[3], enables the agent to chose an appropriate action and execute it, completing the cycle.

Thus, the LIDA cognitive cycle can be subdivided into three phases, the understanding phase, the attention (consciousness) phase, and the action selection phase. Figure 1 should help the reader follow the description. It starts in the upper left corner and proceeds roughly clockwise. Beginning the understanding phase, incoming stimuli activate low-level feature detectors in Sensory Memory. The output is sent to Perceptual Associative Memory where higher-level feature detectors feed in to more abstract entities such as objects, categories, actions, events, etc. The resulting percept moves to the Workspace where it cues both Transient Episodic Memory and Declarative Memory producing local associations. These local associations are combined with the percept to generate a Current Situational Model, the agent's understanding of what is going on right now.

Attention Codelets[4] begin the attention phase by forming coalitions of selected portions of the Current Situational Model and moving them to the Global Workspace. A competition in the Global Workspace then selects the most salient, e.g., the most relevant, important, urgent, novel, unexpected, loud, bright, etc. coalition, whose contents become the content of consciousness. These conscious contents are then broadcast globally, initiating the action selection phase. The action selection phase of LIDA's cognitive cycle is also a learning phase in which several

[2] "Cognition" is used here in a unusually broad sense, so as to include perception, feelings and emotions.

[3] Here "consciousness" refers to functional consciousness (Franklin 2003). We take no position on the need for, or possibility of, phenomenal consciousness.

[4] A codelet is a small piece of code that performs a specific task in an independent way. It could be interpreted as a small part of a bigger process, similar to an ant in an ant colony.

processes operate in parallel (see Figure 1). New entities and associations, and the reinforcement of old ones, occur as the conscious broadcast reaches Perceptual Associative Memory. Cognitive maps are created or updated in Spatial Memory. Events from the conscious broadcast are encoded as new memories in Transient Episodic Memory. Possible action schemes, each consisting of an action together with its context and expected result, and an activation measuring the likelihood of the result occurring if the action is taken in the context (Drescher, 1991), are learned into Procedural Memory from the conscious broadcast. Older schemes are reinforced. In parallel with all this learning, and using the conscious contents, possible action schemes are recruited from Procedural Memory. A copy of each such is instantiated with its variables bound, and sent to Action Selection (Maes, 1989), where it competes to be the behavior selected for this cognitive cycle. The selected behavior triggers Sensory-Motor Memory to produce a suitable algorithm for the execution of the behavior. Its execution completes the cognitive cycle.

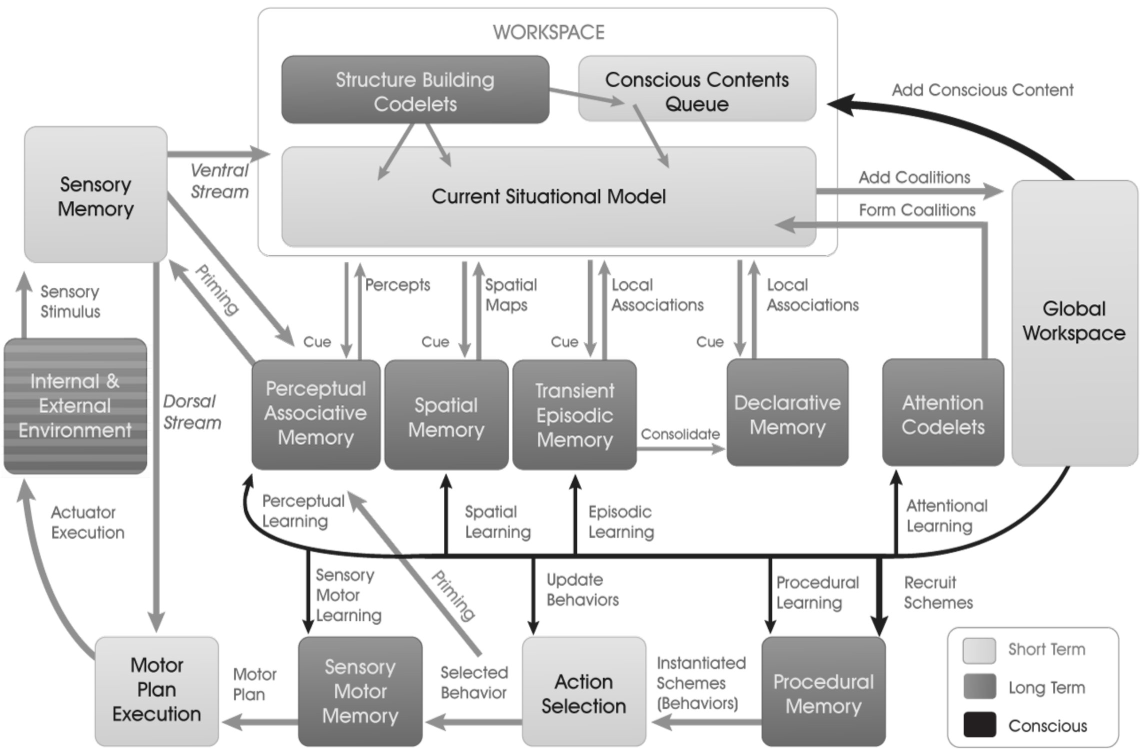

Figure 1. The LIDA Cognitive Cycle Diagram.

The Workspace requires further explanation. Its internal structure includes the Current Situational Model and the Conscious Contents Queue. The Current Situational Model is where the structures representing the actual current internal and external events are stored. Structure building codelets are responsible for the creation of these structures using elements from the various sub-modules of the Workspace. The conscious contents queue holds the contents of the last several broadcasts and permits LIDA, using codelets, to understand and operate upon time related concepts (Snaider, McCall, & Franklin, 2009).

4. Specific Conceptual Commitments of the LIDA Model

In this section we provide brief descriptions, with some justifications, of over a dozen conceptual commitments that have steered the development of our LIDA model of cognition, both the conceptual model, and its underlying computational architecture. Our intention is to offer them as tentative commitments to be considered for AGI architectures in general. The conceptual commitments described in the subsections below are presented in decreasing order of abstractness. Many of them can be thought of as hypotheses based either on the interpretation of empirical data from artificial intelligence, cognitive neuroscience or cognitive psychology, or on the needs of the model's current implementation. Please note once again that we explicitly make no claim that any of the conceptual commitments described below are unique to the LIDA model. Likely, each can also be found elsewhere. Finally, at the end of each commitment's description we will speculate on the possible usefulness or importance of that commitment to AGI architectures. For clarity's sake, we have assessed each commitment's importance to the LIDA model on one of four levels: essential, very significant, significant, and subsidiary.

4.1 Systems-level Modeling

Level of importance to the LIDA Model: Essential
Level of importance to AGI: Essential

Scientists use models, be they conceptual, mathematical, computational, to explain and to predict. Both explanation and prediction are more easily accomplished using specific models restricted to some limited function of cognition, for example, individual models of perception, memory, attention, learning, action selection, etc. Due to this facilitation almost all cognitive models have a restricted scope.

Still, questions arise as to how these various restricted functionalities interact, or relate to one another. To answer such questions requires a systems-level model that accommodates an incoming stimulus, an outgoing action, and everything in between, that is, the entire cognitive system. All of the various cognitive functionalities must be modeled individually to some level. In addition, the model must, in principle, account for all of their interactions, both pairwise and of higher order.

The need for using systems-level cognitive architectures has been championed in the past by several researchers. As social psychologist Kurt Lewin so succinctly pointed out, "There is nothing so practical as a good theory" (1951, p. 169). Artificial intelligence pioneer Allen Newell strongly supported the need for systems-level theories/architectures, asserting that "You can't play 20 questions with nature and win" (1973). Echoing Newell in decrying the reliance on modeling individual laboratory tasks, memory researcher Douglas Hintzman (2011) wrote, "Theories that parsimoniously explain data from single tasks will never generalize to memory as a whole…" Hintzman's arguments, which rest upon the need for systems-level cognitive architectures in memory research, carry over into the realm of intelligent agents, again calling for systems-level architectures. In their review article, Langley, Laird, and Rogers (2009) argue that "Instead of carrying out micro-studies that address only one issue at a time, we should attempt to unify many findings into a single theoretical framework, then proceed to test and refine that theory." They are all calling for the use of a broad-based, systems-level cognitive architecture.

We find much of value in the university metaphor used by Bullock in his cry for systems-level neuroscience (1993). Models of the university might be based on parameters such as the composition, structure, or mechanisms underlying components such as typewriters, telephones and people, or built around principles such as, "It works by shuffling a material called paper" and "It works by spatiotemporal configurations of units called committees" (Bullock, 1993, p. 1). While indubitably valuable, such approaches provide a picture of the modeled entity that is neither unified nor complete. Bullock continues:

The university is found to work by (i) interactions among partially equivalent but mostly nonredundant individuals, (ii) each with rich but fragmentary and filtered inputs, (iii) making decisions at widely different levels of consequence, which are (iv) based on those inputs but integrated with endogenous tendencies, (v) taking actions partly in concert, partly quite out of phase with others, (vi) every individual being unique but none indispensible. (vii) The system is adjustable due to a network of connectivity and shared competencies, (viii) though normally the individuals operate with distinct responsibilities. (Bullock, 1993, p. 2)

Bullock speculates, based on his then 53 years of empirical study, that this high-level description probably applies to the brain as well as universities. We find it extremely illuminating as well for the study of cognition.

Any such systems-level cognitive model must necessarily be quite complex, and so, time consuming to design and implement. Thus, at present, Samsonovich (2010) catalogs only about two-dozen such models, including our LIDA model.

It seems to us that any AGI agent may have to be controlled by a systems-level architecture, making this commitment one of wide generality for AGI research.

4.2 Biologically Inspired

Level of importance to the LIDA Model: Very significant
Level of importance to AGI: Undetermined

Many computational models, from artificial neural networks (ANNs) to the large-scale brain simulation of (Izhikevich & Edelman, 2008) can claim some degree of biological inspiration. Such models vary greatly not just in this degree, but also in purpose. ANNs have a wide range of applications, but stake no claim as a comprehensive model of brains or even of neurons. On the other hand, while large-scale brain simulations model a broad spectrum of neural phenomena, they do not seek to explain cognition. A commitment to biological inspiration in AGI entails the belief that the principles organizing biological minds can be useful to a complete understanding of cognition. However, such a commitment does not imply a particular degree of commitment to modeling specific biological implementations of the cognitive processes themselves.

Not all of the few dozen broad, systems-level cognitive models intend to be biologically inspired. For most of those that do, the claim goes unargued. For a few, such as ACT-R and CLARION, such a claim may be justified by the replication of data from experiments with humans and other animals (e.g., Gunzelmann, Gluck, Van Dongen, O'Connor, & Dinges, 2005; Sun & Naveh, 2004). Moreover, ACT-R bases its architecture of buffers and modules directly on the gross functional anatomy of the human brain, and has been successful in replicating data from fMRI studies (Anderson et al., 2004). With the development of the LIDA Computational Framework (Snaider, McCall, & Franklin, 2011), our LIDA model has begun to spawn such replications (Faghihi, McCall, & Franklin, 2012; Madl, Baars, & Franklin, 2011; Madl & Franklin, 2012). More are forthcoming. In addition, the model implements and fleshes out central ideas from a number of psychological and neuropsychological theories. These include Global Workspace Theory (Baars, 1988, 2002), situated (embodied) cognition (Glenberg & Robertson, 2000; Varela, Thompson, & Rosch, 1991), perceptual symbol systems (Barsalou, 1999a), working memory (Baddeley & Hitch, 1974), memory by affordances (Glenberg, 1997), long-term working memory (Ericsson & Kintsch, 1995), transient episodic memory (Conway, 2002), and Sloman's H-CogAff cognitive framework (1999).

Biological inspiration is not only espoused by LIDA to validate it as a model of biological minds. Biological minds represent the sole examples of the sort of robust, flexible, systems-level control architectures needed to achieve human-level intelligence. Often copying after a known biological example is a good strategy. As such it appears fruitful to study and incorporate the functionalities of biological minds that make them successful.

One argument against biological inspiration maintains that since biological flight did not prove to be a good model for mechanical flight in the genesis of modern aviation, biological inspiration might prove to be a waste of time for the design of other systems (Sukthankar, 2000), perhaps even AGI. However, biological and mechanical flight fulfill two highly divergent sets of modeling constraints, including very different speeds, weight loads, fuel ranges, and material strengths; an organic vs. a petroleum-based power source; the motivation for diversity vs. uniformity in production; and a biological vs. a commercial or military mandate. Certainly biological and computational cognition will also possess differing constraints. We strongly suggest that the study of the neurocognitive strategies fulfilling biological constraints will deepen our understanding of how cognition relates to its boundary conditions in general. Such an understanding should prove quite valuable to implementing cognition in a computational domain. Even if one finds this argument unconvincing given the current lack of theoretical consensus, and the need for further empirical validation of competing hypotheses, our approach will surely provide a valuable contribution toward achieving an artificial general intelligence.

It certainly seems possible to us that an AGI agent can be successfully developed using an architecture that is not biologically inspired. Thus, in our view, no claim can be made as yet as to the general utility or level of importance of this particular commitment to AGI research.

4.3 Embodied (Situated) Cognition

Level of importance to the LIDA Model: Significant
Level of importance to AGI: Undetermined

Embodied cognition argues that all aspects of cognition are shaped by the body (in particular the brain), and its interaction with the environment (situatedness) through incoming stimuli and outgoing motor actions (de Vega, Glenberg, & Graesser, 2008). We interpret the concept of embodiment broadly as a structural coupling between an autonomous agent and its environment (Franklin 1997), providing a criterion for embodiment in agents that interact with non-physical environments.

The LIDA model complies with this position by abjuring the use of amodal symbols, and depending on perceptual symbols instead (Barsalou, 1999). Though labels appear in the diagrams used to describe the conceptual LIDA model, they are only for the use of external human observers, and play no role in LIDA's internal dynamics. LIDA does not employ symbolic mechanisms, but, rather, abstract concepts would be learned from several examples (experiences) in a situated and embodied way. Though seemingly more difficult to implement, this approach is more consistent with our biologically motivated assumption (Barsalou, 2008). Older architectures, such as SOAR and ACT-R, typically employ symbolic mechanisms.

As with several of our commitments, there is still much controversy. Some might argue that we do not go far enough to claim a commitment to embodiment. For example, we do not insist on a robotic implementation (Franklin, 1997), while others certainly would (Pfeifer & Bongard, 2006). Still others are critical of the whole embodied cognition movement itself (Daum, Sommerville, & Prinz, 2009; Longo, 2009). In actuality, we do not reject the physical symbol system hypothesis (PSSH) of SOAR and ACT-R, feeling that PSSH is not truly in conflict with embodied cognition. In short, perceptual symbols a la Barsalou (1999b) work as well within a physical symbol system as do amodal symbols. Certainly, the symbols in SOAR and ACT-R implementations to date have been almost exclusively amodal, contributing to a misunderstanding of PSSH, but such a restriction has been pragmatic rather than axiomatic. Nonetheless, representations in LIDA will possess references to the sensory and perceptual primitives that invoked them, in keeping with the concept of embodiment as we interpret it.

Though firmly committed to embodiment for LIDA, we are hesitant to claim its necessity for all AGI architectures. It seems possible that successful AGI agents may be controlled by symbolic architectures.

4.4 Cognitive Cycles as Cognitive Atoms

Level of importance to the LIDA Model: Essential
Level of importance to AGI: Undetermined

Every autonomous agent, human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). Sophisticated agents, such as humans, process (make sense of) the input from such sampling in order to facilitate their action selection. The agent's "life" can be viewed as consisting of a continual cascading sequence of these cognitive cycles. The LIDA model suggests that basic human and animal cognition can be usefully viewed as functioning by means of cognitive cycles whose internal structure fleshes out the composition of the action-perception cycles of the psychologists (Neisser, 1976) and the neuroscientists (Cutsuridis, Hussain, & Taylor, 2011; Dijkstra, Schöner, & Gielen, 1994; Freeman, 2002; Fuster, 2002; 2004). Such cognitive cycles comprise continual interactions between conscious contents, the various memory systems, and action selection.

While these cycles overlap, their various modules asynchronously producing parallel effects, they must preserve the seriality of consciousness. Each LIDA agent's cycle consists of three phases, a perceiving and understanding phase, an attending and consciousness phase, and an action selection and learning phase. There is not a sharp boundary between the perceiving and understanding phase and the attending and consciousness phase. Coalitions can be forming during the understanding process. The conscious broadcast is, however, a relatively sharp boundary.

A cognitive cycle can be thought of as a cognitive "moment," lasting roughly 300-500ms in humans (Madl, et al., 2011). Higher-level cognitive processes are composed of many of these cognitive cycles, each a cognitive building block. Figure 1 depicts a flow diagram of much of the internal structure and processing of LIDA's cognitive cycle. Note that while memories appear as separate boxes, this should not be taken as suggesting that there are sharp boundaries between memory modules; in fact, memories may use or "point back to" other memories, e.g. an episodic memory could reference a Perceptual Associative Memory (Fuster, 2006; Fuster & Bressler, 2012). Additionally, while arrows are depicted passing information between modules, the LIDA model is not committed to copying representations in the computational sense. Some believe that pointers, which refer back to a source, are used in brains in the place of copying (Fuster, 2006; Fuster & Bressler, 2012). The lack of such a commitment may differentiate LIDA from most other such systems-level cognitive architectures.

Moreover, while many, and perhaps all cognitive architectures are based on a computational cycle of some kind, LIDA has no system clock to which its cycle is lock-stepped. Thus LIDA's cognitive cycle is not explicitly hard-wired but rather emerges from constraints on its many component processes, which operate in parallel (see Asynchrony, below).

While this emergent cognitive cycle commitment is central to the LIDA model, we hesitate to propose it as important for AGI research in general, since to our knowledge no other systems-level cognitive architecture makes such a commitment.

4.5 Global Workspace Theory

Level of importance to the LIDA Model: Essential
Level of importance to AGI: Very significant

Global Workspace Theory (GWT) views the nervous system as a distributed parallel system with many different specialized processes (Baars, 1988, 2002). Coalitions of these processes enable an agent to make sense of the sensory data coming from the current environmental situation. Other coalitions, incorporating the results of the processing of sensory data, compete for attention in what Baars calls the "Global Workspace." The contents of the winning coalition are broadcast to all other processes. The contents of this broadcast are proposed to be phenomenally conscious. This conscious broadcast serves to recruit other, unconscious processes to be used to select an action in response to the current situation, and to facilitate the various modes of learning (see the following section and the memory systems in Figure 1). GWT is therefore a theory of how consciousness functions within cognition (Figure 2) (Baars & Franklin, 2003; Franklin, et al., 2005). There is considerable empirical support for GWT from neuroscience studies (Baars, 2002), and other researchers have employed it in their architectures (Connor & Shanahan, 2010; Sergent & Dehaene, 2004; Shanahan, 2006; Wallace, 2005). It has also attracted much attention as the primary exemplar of what the philosophers call access consciousness (Baars, 2002; Block, 2007; Dennett, 2005; Dennis & Schutter, 2004). Of course, there are other theories of consciousness (e.g., Augustenborg, 2010; Edelman & Tononi, 2000; Sun & Franklin, 2007; Taylor, 2011; Tononi, 2008), but none to our knowledge, other than GWT, explicitly serve as the basis of a systems-level cognitive architecture.

While the commitment to GWT is certainly central to our LIDA model, we consider only the functional aspect, and not the specific implementation, to be crucial to AGI research. This functionality includes filtering the agent's current situation for the most salient portion to be used to select the next action (see Figure 2).

4.6 Learning via Consciousness

Level of importance to the LIDA Model: Essential
Level of importance to AGI: Significant

A second major function of consciousness in cognition, at least functional consciousness as implemented in the global broadcast of GWT[3], is the enabling of learning, the encoding of knowledge about the past for use in the present. The conscious broadcast selects the most salient portion of the current situation to be learned by the various memory systems. GWT supports the Conscious Learning Hypothesis: significant learning takes place via the interaction of consciousness with the various memory systems (Baars & Franklin, 2003; Franklin, et al., 2005). That is, all memory systems rely on conscious cognition for their updating, either in the course of a single cycle or over multiple cycles (see above for a discussion of cognitive cycles). Note that learning as used here does not refer to Hebbian learning or other forms of neural plasticity, rather it refers to the Conscious Learning Hypothesis, a theory of cognitive, not neural, learning. Nonetheless, this does not preclude that learning may be implemented via Hebbian mechanisms. Priming, of course, can occur unconsciously (Boltea & Goschke, 2008; Eimer & Schlagecken, 2003), but is of such limited scope and brief duration as not to be considered significant learning. Both implicit learning (Cleeremans, Destrebecqz, & Boyer, 1998; Jimenez, 2003) and latent learning (Campanella & Rovee- Collier, 2005; Chamizo & Mackintosh, 1989; Franks et al., 2007) require subjects to be awake and alert, and thus presumably conscious.

Data flow according to GWT can be visualized as having an hourglass shape with sensory data coming in the left and flowing through the leftmost cone (please see Figure 2). The bottleneck at the center represents the limited capacity global workspace acting as an attentional (relevance) filter before the broadcasting of conscious contents throughout the brain, represented by the rightmost cone. GWT is therefore a theory of how consciousness functions within cognition, first as a filter and then as a recruiter and a modulator of learning.

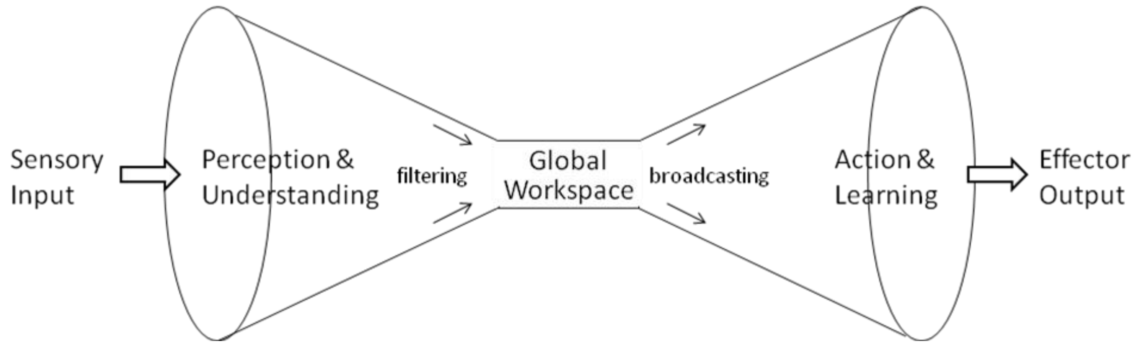

Figure 2. Global Workspace Theory within the Action-Perception cycle.

Frequent published assertions about "unconscious learning" might make our Conscious Learning Hypothesis appear untenable. However, closer examination reveals that all such refer to either priming or implicit learning. Subliminal priming can occur when a stimulus is presented for too short a time for conscious awareness to occur, and followed by some masking stimulus. The priming stimulus, though unconscious, can affect subsequent decisions (Eimer & Schlagecken, 2003; Silverta, Delplanquea, Bouwalerha, Verpoorta, & Sequeira, 2004; Tulving & Schacter, 1990; Yang, Xu, Du, Shi, & Fang, 2011). This kind of masked priming effect involves too little content over too short a time period for us to be comfortable referring to it as learning.

Implicit learning is often misleadingly defined as learning that takes place without intention or awareness (Jimenez, 2003). In each instance of implicit learning, the learning subject must be conscious during the learning experience. Typically some skill is learned (for example, recognizing well-formed sequences of letters) while some underlying enabling pattern (the rule by which well-formed sequences are constructed) remains implicit (Cleeremans, et al., 1998; Jimenez, 2003; Reber, 1967; Sun, Slusarz, & Terry, 2005). Implicit learning is, indeed, learning, but it does not occur in the absence of consciousness. We refer to such learning as "consciously-mediated" (Franklin & Baars, 2010; Franklin, et al., 2012).

We do contend that commitments to learning will be critical to any AGI agent, since not all eventually needed recognitions, knowledge, skills, etc. can be built in. However, we make no such contention about LIDA's particular mechanisms of learning.

4.7 Comprehensive Decay of Representations and Memory

Level of importance to the LIDA Model: Significant
Level of importance to AGI: Undetermined

Essentially, every representational entity in the LIDA model has a numeric activation value, which may have different meanings for different entities. To be clear, modules are not considered representational entities though they may contain instantiated representations and memory. Entities in one of the long-term memories typically have both a base-level activation measuring the past usefulness of the entity, and a current activation that measures its current relevance. In LIDA, all activations decay. Decay rates vary with the type of entity and with the type of activation. The decay rate of a given entity will often decay in an inverse exponential relationship to its activation. If the activation of an entity decays sufficiently, that entity (memory trace) is removed from the LIDA memory module. On the other hand, due to the exponential decay, saturated entities may decay so slowly as to seem permanent.

Many human memory researchers are happy with the idea of memory traces decaying away (Cansino, 2009; Sims & Gray, 2004). Others believe that long-term memory traces never decay away, but rather are unable to be reached due to retrieval failure (cue-dependent forgetting) (Armstrong & Mewhort, 1995; Bjork & Bjork, 1988; Miller & Matzel, 2006; Shiffrin, 1970).

In addition to decay, representations in LIDA can change as a result of interference by later cognitive activities. Memory loss due to interference is not controversial among human memory researchers (Brainerd & Dempster, 1995; Chandler, 1991). Neither is the concept of seemingly permanent long-term episodic memory storage (Craik, Routh, Broadbent, & Craik, 1983; Stickgold & Walker, 2005).

Again our commitment to decay in LIDA is a fairly significant one, but we make no such claim to its general usefulness or importance to AGI research. It may not even happen uniformly in humans.

4.8 Profligacy in Learning

Level of importance to the LIDA Model: Subsidiary
Level of importance to AGI: Undetermined

As noted in the section on learning via consciousness above, learning occurs in a profligate manner during each of LIDA's cognitive cycles as a result of the conscious broadcast. In Figure 1, the solid lines depict opportunities for this multimodal learning to occur in each of the long-term memory systems using the contents of the conscious broadcast. Such learning can be instructionalist, the learning of new entities, or selectionist, the reinforcement of existing entities by modifying their base-level activation. Thus learning can occur at roughly five to ten times a second in each of the modalities (Madl, et al., 2011). We refer to this process as profligate learning, the rapid learning of everything that comes to consciousness. Of course, most of what is so learned rapidly decays away (see the previous section). Recall that long-term memories are, in fact, only potentially long-term. Given sufficient reinforcement, a particular memory may acquire enough base-level activation that its decay rate essentially drops to zero. In AI, such common profligate learning algorithms are often referred to by the term generate and test[5] (e.g., Kaelbling, 1994), or in common language, trial and error.

Though learning via generate and test has certainly proved itself useful to AI research, we see no reason to claim that profligate learning must be equally useful to AGI research.

4.9 Feelings as Motivators and Modulators of Learning

Level of importance to the LIDA Model: Very Significant
Level of importance to AGI: Significant

The LIDA model employs artificial feelings and emotions that allow for flexible and sophisticated action selection, and for the modulation of learning (Franklin & Ramamurthy, 2006). Feelings in humans include hunger, thirst, various sorts of pain, hot or cold, the urge to urinate, tiredness, depression, etc. Implemented biologically as somatic markers (Damasio, 2008), feelings typically attach to response options and therefore bias the agent's choice of action.

Emotions, such as fear, anger, joy, sadness, shame, embarrassment, resentment, regret, guilt, etc., are taken to be feelings with cognitive content (Johnston, 1999; Panksepp, 2005). One cannot simply feel shame, but shame at having done something—the cognitive content. Feelings, including emotions, are nature's means of implementing motivations for actions in humans and other animals. Feelings and emotions give us the ability to make an almost immediate assessment of situations (Roseman & Smith, 2001; Smith & Kirby, 2001). They allow us to determine whether a given state of the world is beneficial or detrimental.

---

[5] Learning is generated each time content comes to consciousness, and its usefulness is continually tested by decay as time passes.

Feelings are represented in the LIDA Model as particular kinds of nodes in its Perceptual Associative Memory and elsewhere. Each feeling node constitutes its own identity; for example, distress at not enough oxygen is represented by one node, relief at taking a breath by another. Each feeling node has its own valence, always positive or always negative. The current activation of the node measures its arousal. Those feeling nodes with sufficient activations, along with their incoming links and object nodes, become part of the current percept and are passed to the Workspace.

These feeling nodes play a major role in contributing activation to those coalitions that contain them, thereby increasing the coalition's likelihood of selection in the Global Workspace. Any feeling nodes that belong to the winning coalition become part of the conscious broadcast, the contents of consciousness.

Any feeling node in the conscious broadcast that also occurs in the context of a scheme in Procedural Memory adds to the current activation of that scheme, increasing the likelihood that it will be instantiated into the action selection mechanism. It is here that feelings play an additional role as implementation of motivation by adding to the likelihood of a particular action being selected.

Learning is both a function of attention and of feelings. Feelings in the conscious broadcast modulate learning. Up to a point, the more intense a feeling node's activation, also known as its affect, the greater the learning. Beyond the optimal point, more affect begins to interfere with learning (Yerkes & Dodson, 1908).

While some form of quite flexible motivation will be a critical commitment to any AGI agent, there may well be ways to bring this about other than by feelings and emotions.

4.10 Asynchrony

Level of importance to the LIDA Model: Very significant
Level of importance to AGI: Undetermined

The LIDA cognitive cycle diagram above (Figure 1) consists of modules, represented by labeled boxes, and their associated processes, represented by labeled arrows. The modules mostly store memories of some kind or another, either (potentially) long-term or shorter term. The process operating on the Global Workspace waits for an appropriate global trigger before broadcasting the current contents of consciousness (black arrows in the diagram), thus assuring the seriality and coherence of consciousness. Baars' recent notion of a dynamic Global Workspace (dGW) (Baars, Franklin, & Ramsøy, 2013) suggests that conscious contents arise from a winner-take-all binding coalition among competing and cooperating signal streams emanating from a selected region of the cortico-thalamic core. A winning coalition can ignite a ~100ms global broadcast to widely distributed receiving networks (Gaillard et al., 2009). This suggestion does not contradict the existence of such global triggers in brains.

All of the other processes of the LIDA model operate asynchronously in response to their local conditions. Thus multiple processes may be active simultaneously throughout the LIDA architecture, giving rise to overlapping cognitive cycles running in parallel. In other words, a cognitive cycle does not need to finish before another begins—in fact, it typically won't. These cycles are an emergent property of the architecture. In neither brains, nor in the computational implementations of LIDA-based agents, are cognitive cycles explicitly evolved, or built into the architecture. This may differ in other cognitive architectures.

Though LIDA, and brains, work well without a system clock, we see no reason to expect this must be true of AGI agents in general.

4.11 Transient Episodic Memory

Level of importance to the LIDA Model: Very Significant
Level of importance to AGI: Significant

Following Conway (2001), the LIDA model hypothesizes that humans have a content-addressable, associative, transient episodic memory with a decay rate measured in hours or a day. In our theory, a conscious event is stored in transient episodic memory following a broadcast from the Global Workspace. Although this is a minority opinion in the psychological literature, we justify the need for it as follows. We note that such a transient episodic memory will be needed by any agent who must keep track of repetitive similar events whose important features differ slightly from event to event, for example, where in a parking garage an automobile is parked daily. Such events cannot be successfully tracked by long-term episodic memory that is subject to interference.

In spite of this commitment being ignored by most human memory researchers, we suspect that an AGI agent will be confronted with repetitive similar events with important features differing only slightly from event to event. Thus this will be a significant commitment for AGI research.

4.12 Consolidation

Level of importance to the LIDA Model: Subsidiary
Level of importance to AGI: Subsidiary

A corollary to the hypothesis of the previous section says that conscious contents can only be encoded (consolidated) in long-term Declarative Memory via Transient Episodic Memory (see Figure 1). Though still somewhat controversial, there is much evidence for consolidation (Born & Wagner, 2006; Graves, Heller, Pack, & Abel, 2003; McGaugh, 2000; Nadel, Hupbach, Gomez, & Newman-Smith, 2012; Remondes & Schuman, 2004; Stickgold & Walker, 2005; Wamsley, Tucker, Payne, Benavides, & Stickgold, 2010; Wiedemann, 2007; Zhang, 2009).

A commitment to some way of producing long-term episodic memory will surely be needed by AGI agents, but this need not happen via consolidation.

4.13 Non-linear Dynamics Bridge to Neuroscience

Level of importance to the LIDA Model: Very Significant
Level of importance to AGI: Undetermined

The LIDA model is a model of mind, not of the underlying brain. LIDA is a biologically inspired model, which incorporates the useful, functional aspects of the brain into the model, e.g. Global Workspace Theory, while, hopefully, avoiding those features idiosyncratic to the brain. In addition, such a model of mind should be consistent with the only known existing implementations of highly sophisticated minds; namely those generated by biological brains. Thus, the model needs to avail itself of existing neuroscientific evidence in order to account for the observed connection between minds and brains.

We propose that non-linear, self-organizing dynamics constitutes a bridge or mapping, in principle, between the mind's cognitive representations and the neural dynamics that underlie them. For example, Fuster's cognits (2006) are proposed to represent cognitive entities neurally. The brain dynamically integrates the activity of its perceptual oscillators with the activity of its higher-order neural oscillators[6] (Barham, 1996; Freeman, 2003) so as to subserve the application of memory, deliberation, and goals to the present state of the environment and brain. In keeping with Global Workspace Theory, a subset of this integrated oscillatory activity is selected (see Asynchrony above) for broadcast. This broadcast then drives action selection and several forms of learning, and the selected actions activate oscillators that control the organism's action execution. Furthermore, we propose timing relationships in the form of phase-coupling between oscillators as a key characteristic of cognition's neurophysiological "structure" (Strain, Franklin, Heck, & Baars, in preparation).

This commitment is to a particular explanation of the relationship between LIDA and neuroscience. It need play no role in AGI research in general.

4.14 Theta Gamma Coupling from the Cognitive Cycle

Level of importance to the LIDA Model: Significant
Level of importance to AGI: Undetermined

This conceptual assumption is about the relationship of the LIDA model to the underlying empirical neuroscience evidence. Cross-frequency-coupling (CFC), in which high-frequency (gamma) activity organizes within low-frequency (theta) response patterns, is implicated in a variety of cognitive contexts (Canolty & Knight, 2010), including declarative memory (Nyhus & Curran, 2010; Osipova et al., 2006; Sederberg, Kahana, Howard, Donner, & Madsen, 2003), working memory (Sauseng, Griesmayr, Freunberger, & Klimesch, 2010; Tort, Komorowski, Manns, Kopell, & Eichenbaum, 2009), attention (Sauseng, Klimesch, Gruber, & Birbaumer, 2008), and perceptual organization (Doesburg, Green, McDonald, & Ward, 2009). More specifically, cross-frequency coupling was detected as a strong correlation between theta phase and gamma power (Canolty et al., 2006).

Within the LIDA model, individual cognitive cycles run their course in roughly 300-500ms. Several such cognitive cycles, perhaps three, can cascade or overlap as long as the quasi-seriality of consciousness is preserved. Thus, such cascading cognitive cycles would be expected to occur at a theta band rate of 5-8hz. But within each cognitive cycle one finds a bevy of modules and processes contributing to the activity of the cycle (see Figure 1), always varying with the current situation or task. An attractive conjecture from the LIDA Model is that this internal activity of cognitive cycles gives rise to gamma frequency amplitude modulations during the course of a theta cycle corresponding to a cognitive cycle (Strain, et al., in preparation). If so, the observed phenomenon of theta-gamma coupling on some underlying cell assembly, that is, the modulating of each theta cycle by gamma frequency amplitude variations, could be interpreted as corresponding to the activity of a LIDA process correlated with that cell assembly. Variations in theta phase would correspond to variations in cognitive cycle lengths and in the details of the overlap during the cascading of cognitive cycles.

This commitment concerns only the relationship between LIDA's emergent cognitive cycle and the emergent theta-gamma cycles in brains. It may well have no role at all to play in AGI research in general.

[6] From Barham (1996), we construe "oscillator" as referring broadly to the activity pattern of any fluctuating entity that has a measurable physical energy. Examples range from the low-energy activity of sensory receptors, neurons and neural ensembles, to high-energy patterns such as the variation of a predator population, the presence or absence of a physical obstacle, or even weather patterns and solar cycles. In this scheme, an organism's low-energy oscillators allow it to respond to the activity of high-energy environmental oscillators before becoming "thermodynamically coupled" to them (Barham, 1996).

5. Summary and Conclusions

We have approached the problem of artificial general intelligence by integrating existing models and evidence into the broad-based LIDA model comprising the known cognitive principles of perception, memory, understanding, attention, consciousness, action selection and execution. In order to accomplish this, we have taken tentative theoretical stances in each of these subfields, many of which would be controversial if asserted dogmatically. As a rule, our stances do not represent staunchly held positions in the various debates, so much as operational commitments that enable the broad modeling necessary to artificial general intelligence research. If any one of our tentative commitments were to prove false, it would likely alter our model significantly.

To summarize, the research philosophy for the LIDA model is strongly defined by an adherence to *systems-level modeling* of multiple aspects of cognition within a single unified model. LIDA is *biologically inspired*, using biology as a guide as well as a restriction on its functionality. It adheres to *embodied (situated) cognition*, using grounded modal representations. LIDA views *cognitive cycles as cognitive atoms*, suggesting high-order processes are built out of simpler action-perception cycles. Nonetheless, there are no serial timekeepers defining such cycles; rather the seriality emerges from a bevy of *asynchronous* processes. LIDA is strongly coupled with *Global Workspace Theory*, which suggests the need for a filter on the agent's representation of the current situation selecting the most salient portion to be used to select the next action. Additionally, *learning* is hypothesized to primarily occur *via* this *consciousness* selection. Since consciousness moments appear to be frequent (every 100ms), then *learning is profligate,* which should be offset by *comprehensive decay of representations and memory,* leaving only the most frequent and salient representations. In order to have an autonomous agent with an agenda, LIDA, sticking with biological inspiration, calls upon *feelings as motivators and modulators of learning.* LIDA requires *Transient Episodic Memory*, for repetitive similar events whose important features differ slightly from event to event. A corollary to this suggests that long-term Declarative Memory should be produced via the *consolidation* of Transient Episodic Memory. Finally, LIDA grounds itself to neuroscience in two ways: 1) It proposes a *non-linear dynamics bridge* between mental representations and the neural dynamics that underlie them, and 2) it suggests *theta gamma coupling from the cognitive cycle*.

As mentioned, the cognitive-cycle-as-cognitive-atom-hypothesis and the Global Workspace Theory of consciousness are two firmly held commitments in the LIDA Model. LIDA's modules contain distributed asynchronous processes with a quasi-seriality imposed by periodic conscious broadcasts according to GWT. This seriality defines a cycle. Asynchrony predicts that one such cycle may begin before the previous one is completed, making empirical detection and measurement of such cycles an extremely difficult methodological problem. However, the model itself offers a computational platform, the recently completed LIDA Computational Framework (Snaider, McCall, & Franklin, 2011) on which to test the Model's numerous hypotheses. Implementations of a model of mind should replicate data from experimental psychology, and/or produce reasoning, learning, and behavior comparable to that of biological organisms including humans. Ongoing research will reinforce validated assumptions and mandate changes to those producing results inconsistent with empirical data.

The authors state these commitments here hoping to generate discussion of these, as well as the commitments of other cognitive modelers, as a means of furthering research in artificial general intelligence. It would seem that a consensus as to commitments might well lead to a greater similarity of AGI architectures, with the possibility of inter-usability of their modules.

Acknowledgements

We are greatly indebted to the members of the Cognitive Computing Research Group (CCRG) at the University of Memphis for their many contributions to the various commitments of the LIDA model described here. We also offer our thanks to Phil Pavlik for reviewing the manuscript and providing several very useful suggestions. Finally, we thank the three JAGI editors for their insightful reviews that so helped to improve this paper.

References

Anderson, J. R. (2007). Using brain imaging to guide the development of a cognitive architecture. In W. D. Gray (Ed.), Integrated models of cognitive systems (pp. 49-62). New York: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An Integrated Theory of the Mind *Psychological Review, 111*(4), 1036-1060.

Armstrong, I. T., & Mewhort, D. (1995). Repetition deficit in rapid-serial-visual-presentation displays: Encoding failure or retrieval failure? *Journal of Experimental Psychology: Human Perception and Performance, 21*(5), 1044.

Augustenborg, C. C. (2010). The Endogenous Feedback Network: A new approach to the comprehensive study of consciousness. *Consciousness and Cognition*. doi: 10.1016/j.concog.2010.03.007

Baars, B., Franklin, S., & Ramsøy, T. (2013). Global workspace dynamics: Cortical “binding and propagation” enables conscious contents. *Frontiers in Consciousness Research, 4*, 200. doi: 10.3389/fpsyg.2013.00200

Baars, Bernard J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

Baars, Bernard J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Science, 6*, 47–52.

Baars, Bernard J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Science, 7*, 166–172.

Bach, J. (2008). Seven principles of synthetic intelligence. In P. Wang, G. Goertzel & S. Franklin (Eds.), *Artificial General Intelligence 2008: Proceedings of the First AGI Conference* (pp. 63-74). Amsterdam: IOS Press.

Bach, J., Goertzel, B., & Iklé, M. (Eds.). (2012). *Artificial General Intelligence: 5th International Conference*. Oxford, UK: Springer.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation*. New York: Academic Press.

Barham, J. (1996). A dynamical model of the meaning of information. *BioSystems, 38*, 235-241.

Barsalou, L. W. (1999a). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(04), 577-609.

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

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol., 59*, 617-645.

Bjork, E. L., & Bjork, R. A. (1988). On the adaptive aspects of retrieval failure in autobiographical memory.

Block, N. (2007). Consciousness, Accessibility and the Mesh between Psychology and Neuroscience. *Behavioral and Brain Sciences, 30*, 481-548.

Boltea, A., & Goschke, T. (2008). Intuition in the context of object perception: Intuitive gestalt judgments rest on the unconscious activation of semantic representations. *Cognition, 108*(3), 608-616. doi: 10.1016/j.cognition.2008.05.001

Born, J., & Wagner, U. (2006). Memory Consolidation during Sleep: Role of Cortisol Feedback. *Annals of the New York Academy of Sciences, 1032*, 198 - 201. doi: 10.1196/annals.1314.020

Brainerd, C. J., & Dempster, F. N. (1995). *Interference and inhibition in cognition*: Academic Press.

Bullock, T. H. (1993). Goals and Strategies in Brain Research: The Place of Comparative Neurology. In T. H. Bullock (Ed.), *How Do Brains Work?: Papers of a Comparative Neurophysiologist How Do Brains Work?: Papers of a Comparative Neurophysiologist*. Boston: Birkhauser.

Campanella, J., & Rovee- Collier, C. (2005). Latent learning and deferred imitation at 3 months. *Infancy, 7*(3), 243-262.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science, 313*(5793), 1626-1628.

Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences, 14*(11), 506-515. doi: 10.1016/j.tics.2010.09.001

Cansino, S. (2009). Episodic memory decay along the adult lifespan: A review of behavioral and neurophysiological evidence. *International Journal of Psychophysiology, 71*(1), 64-69.

Chamizo, V. D., & Mackintosh, N. (1989). Latent learning and latent inhibition in maze discriminations. *The Quarterly Journal of Experimental Psychology, 41*(1), 21-31.

Chandler, C. C. (1991). How memory for an event is influenced by related events: Interference in modified recognition tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 115–125.

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: news from the front. *Trends in Cognitive Sciences, 2*(10), 406-416.

Connor, D., & Shanahan, M. (2010). A computational model of a global neuronal workspace with stochastic connections. *Neural Netw, 23*(10), 1139-1154. doi: S0893-6080(10)00143-7 [pii]10.1016/j.neunet.2010.07.005

Conway, Martin A. (2001). Sensory–perceptual episodic memory and its context: autobiographical memory. *Philos. Trans. R. Soc. Lond B., 356*, 1375–1384.

Conway, M. A. (2002). Sensory–perceptual episodic memory and its context: Autobiographical memory. In A. Baddeley, M. Conway & J. Aggleton (Eds.), *Episodic Memory*. Oxford: Oxford University Press.

Craik, F. I., Routh, D. a., Broadbent, D., & Craik, F. (1983). On the Transfer of Information from Temporary to Permanent Memory [and Discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 302*(1110), 341-359.

Cutsuridis, V., Hussain, A., & Taylor, J. G. (2011). *Perception-Action Cycle: Models, Architectures, and Hardware* (Vol. 1): Springer.

Daum, M. M., Sommerville, J. A., & Prinz, W. (2009). Disentangling embodied and symbolic modes of social understanding. *European Journal of Social Psychology, 39*(7), 1214-1216.

de Garis, H., & Goertzel, B. (2009a). Report on the 2nd International Conference on Artificial General Intelligence (AGI-09). *AI Magazine*, 115~116.

de Garis, H., & Goertzel, B. (2009b). Report on the First Conference on Artificial General Intelligence (AGI-08). *AI Magazine, 30*, 121-123.

de Vega, M., Glenberg, A., & Graesser, A. (Eds.). (2008). *Symbols and Embodiment: Debates on meaning and cognition*. Oxford: Oxford University Press.

Dennett, D. (2005). Sweet Dreams : Philosophical Obstacles to a Science of Consciousness. Cambridge, MA: MIT Press.

Dennis, J. L., & Schutter, G. (2004). Extending the global workspace theory to emotion: Phenomenality without access. *Consciousness and Cognition, 13*(3), 539-549.

Designing Intelligent Robots: Reintegrating AI II. (2012). *AAAI Spring Symposium 2013* Retrieved December 21, 2012, from http://people.csail.mit.edu/gdk/dir2/index.html

Dijkstra, T. M. H., Schöner, G., & Gielen, C. C. A. M. (1994). Temporal stability of the action-perception cycle for postural control in a moving visual environment. *Experimental Brain Research, 97*(3), 477-486.

Doesburg, S., Green, J., McDonald, J., & Ward, L. (2009). Rhythms of Consciousness: Binocular Rivalry Reveals Large-Scale Oscillatory Network Dynamics Mediating Visual Perception. *PLoS ONE, 4*(7), e6142. doi: 10.1371/journal.pone.0006142

Drescher, Gary L. (1991). Made-Up Minds: A Constructivist Approach to Artificial Intelligence. Cambridge, MA: MIT Press.

Edelman, Gerald M., & Tononi, G. (2000). *A Universe of Consciousness*. New York: Basic Books.

Eimer, M., & Schlagecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology, 64*, 7–26.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211-245.

Faghihi, U., & Franklin, S. (2012). The LIDA Model as a Foundational Architecture for AGI. In P. Wang & B. Goertzel (Eds.), *Theoretical Foundations of Artificial General Intelligence* (pp. 105-123). Paris: Atlantis Press.

Faghihi, U., McCall, R., & Franklin, S. (2012). A Computational Model of Attentional Learning in a Cognitive Agent. *Biologically Inspired Cognitive Architectures, 2*, 25-36.

Franklin, S. (1997). Autonomous Agents as Embodied AI. *Cybernetics and Systems, 28,* 499–520.

Franklin, S., & Baars, B. J. (2010). Spontaneous remembering is the norm: What integrative models tell us about human consciousness and memory. In John H. Mace (Ed.), *The Act of Remembering: Toward an understanding of how we recall the past*. Oxford: Blackwell.

Franklin, S., Baars, B. J., Ramamurthy, U., & Ventura, M. (2005). The Role of Consciousness in Memory. *Brains, Minds and Media, 1*, 1–38.

Franklin, S., & Ramamurthy, U. (2006). Motivations, Values and Emotions: Three sides of the same coin *Proceedings of the Sixth International Workshop on Epigenetic Robotics* (Vol. 128, pp. 41–48). Paris, France: Lund University Cognitive Studies.

Franklin, S., Strain, S., Snaider, J., McCall, R., & Faghihi, U. (2012). Global Workspace Theory, its LIDA model and the underlying neuroscience. *Biologically Inspired Cognitive Architectures, 1*, 32-43. doi: 10.1016/j.bica.2012.04.001

Franks, N. R., Hooper, J. W., Dornhaus, A., Aukett, P. J., Hayward, A. L., & Berghoff, S. M. (2007). Reconnaissance and latent learning in ants. *Proceedings of the Royal Society B: Biological Sciences, 274*(1617), 1505-1509.

Freeman, W. J. (2002). The limbic action-perception cycle controlling goal-directed animal behavior. *Neural Networks, 3*, 2249-2254.

Freeman, W. J. (2003). A neurobiological theory of meaning in perception. Part 1. Information and meaning in nonconvergent and nonlocal brain dynamics. *International Journal of Bifurcation and Chaos, 13*, 2493–2511.

Fuster, J. (2006). The cognit: a network model of cortical representation. *International Journal of Psychophysiology, 60*, 125-132.

Fuster, J. M. (2002). Physiology of executive functions: The perception-action cycle.

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in Cognitive Sciences, 8*(4), 143-145.

Fuster, J. M., & Bressler, S. L. (2012). Cognit activation: a mechanism enabling temporal integration in working memory. *Trends in Cognitive Sciences, 16*(4), 207-218.

Gaillard, R., Dehaene, S., Adam, C., Clémenceau, S., Hasboun, D., et al. (2009). Converging intracranial markers of conscious access. *PLoS Biology, 7*(3), e1000061. doi: 10.1371/journal.pbio.1000061

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*(01), 1-19.

Glenberg, A. M., & Robertson, D. A. (2000). Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language, 43*(3), 379-401.

Goertzel, B., & Pennachin, C. (2007). *Artificial General Intelligence*. Berlin: Springer.

Goertzel, B., & Wang, P. (Eds.). (2007). Advances in artificial general intelligence : concepts, architectures and algorithms. Amsterdam; Washington, DC: IOS Press.

Graves, Laurel A., Heller, Elizabeth A., Pack, Allan I., & Abel, T. (2003). Sleep Deprivation Selectively Impairs Memory Consolidation for Contextual Fear Conditioning. *Learning & Memory, 10*, 168–176.

Gunzelmann, G., Gluck, K. A., Van Dongen, H. P. A., O'Connor, R. M., & Dinges, D. F. (2005). A Neurobehaviorally Inspired ACT-R Model of Sleep Deprivation: Decreased Performance in Psychomotor Vigilance. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society* (pp. 857-862). Mahwah, NJ: Lawrence Erlbaum Associates.

Hofstadter, D. R., & Mitchell, M. (1995). The Copycat Project: A model of mental fluidity and analogy-making. In K. J. Holyoak & J. Barnden (Eds.), *Advances in connectionist and neural computation theory, Vol. 2: logical connections* (pp. 205–267). Norwood N.J.: Ablex.

Izhikevich, E. M., & Edelman, G. M. (2008). Large-Scale Model of Mammalian Thalamocortical Systems. *PNAS, 105*, 3593-3598.

Jimenez, L. (2003). *Attention and implicit learning*: John Benjamins Publishing Company.

Johnston, V. S. (1999). *Why We Feel: The Science of Human Emotions*. Reading, MA: Perseus Books.

Kaelbling, L. P. (1994). Associative Reinforcement Learning: A Generate and Test Algorithm. *Machine Learning, 15*(3), 299-319.

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive Architectures: Research Issues and Challenges. *Cognitive Systems Research, 10*(2), 141-160. doi: 10.1016/j.cogsys.2006.07.004

Lewin, K. (1951). Field theory in sociall science: selected theoretical papers. New York: Harper & Row.

Longo, M. R. (2009). What's embodied and how can we tell? *European Journal of Social Psychology, 39*(7), 1207-1209.

Madl, T., Baars, B. J., & Franklin, S. (2011). The Timing of the Cognitive Cycle. *PLoS ONE, 6*(4), e14803. doi: 10.1371/journal.pone.0014803

Madl, T., & Franklin, S. (2012, April 13-15). *A LIDA-based Model of the Attentional Blink*. Paper presented at the 11th International Conference on Cognitive Modeling, Berlin.

Maes, P. (1989). How to do the right thing. *Connection Science, 1*, 291–323.

McGaugh, J. L. (2000). Memory--a Century of Consolidation. *Science, 287*(5451), 248-251. doi: 10.1126/science.287.5451.248

Miller, R. R., & Matzel, L. D. (2006). Retrieval failure versus memory loss in experimental amnesia: definitions and processes. *Learning & Memory, 13*(5), 491-497.

Nadel, L., Hupbach, A., Gomez, R., & Newman-Smith, K. (2012). Memory formation, consolidation and transformation. *Neuroscience & Biobehavioral Reviews*.

Neisser, U. (1976). Cognition and Reality: Principles and Implications of Cognitive Psychology San Francisco: W. H. Freeman.

Newell, A. (1973). You can't play 20 questions with nature and win: Pro- jective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience & Biobehavioral Reviews, 34*(7), 1023-1035.

Osipova, D., Takashima, A., Oostenveld, R., Fernández, G., Maris, E., & Jensen, O. (2006). Theta and gamma oscillations predict encoding and retrieval of declarative memory. *The Journal of neuroscience, 26*(28), 7523-7531.

Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition, 14*, 30–80.

Pfeifer, R., & Bongard, J. C. (2006). How the body shapes the way we think: a new view of intelligence: MIT press.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855–863.

Remondes, M., & Schuman, Erin M. (2004). Role for a cortical input to hippocampal area CA1 in the consolidation of a long-term memory. *Nature, 431*, 699–703.

Roseman, I. J., & Smith, C. A. (2001). Appraisal theory: Overview, assumptions, varieties, controversies *Appraisal processes in emotion: Theory, methods, research* (pp. 3-19). New York: Oxford University Press.Samsonovich, A. V. (2008, November 7–9). *Biologically Inspired Cognitive Architectures: Papers from the AAAI Fall Symposium*. Paper presented at the AAAI Fall Symposia, Arlington, Virginia.

Samsonovich, A. V. (2010). Toward a Unified Catalog of Implemented Cognitive Architectures. In A. V. Samsonovich, K. R. Jóhannsdóttir, A. Chella & B. Goertzel (Eds.), *Proceeding of the 2010 Conference on Biologically Inspired Cognitive Architectures* (pp. 195-244). Amsterdam: IOS Press.

Samsonovich, A. V., & Johannsdottir, K. R. (Eds.). (2011). Biologically Inspired Cognitive Architectures 2011 - Proceedings of the Second Annual Meeting of the BICA Society (Vol. 233). Amsterdam: IOS Press.

Samsonovich, A. V., Jóhannsdóttir, K. R. J., Chella, A., & Goertzel, B. (Eds.). (2010). *Biologically Inspired Cognitive Architectures 2010 - Proceedings of the First Annual Meeting of the BICA Society* (Vol. 221). Amsterdam: IOS Press.

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neuroscience & Biobehavioral Reviews, 34*(7), 1015-1022.

Sauseng, P., Klimesch, W., Gruber, W. R., & Birbaumer, N. (2008). Cross-frequency phase synchronization: a brain mechanism of memory matching and attention. *NeuroImage, 40*(1), 308-317.

Schmidhuber, J., Thorisson, K. R., & Looks, M. (2011, August 3-6). *Artificial General Intelligence , Proceedings*. Paper presented at the 4th International Conference, AGI 2011, Mountain View, CA, USA.

Sederberg, P. B., Kahana, M. J., Howard, M. W., Donner, E. J., & Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. *The Journal of Neuroscience, 23*(34), 10809.

Sergent, C., & Dehaene, S. (2004). Neural processes underlying conscious perception: Experimental findings and a global neuronal workspace framework. *Journal of Physiology-Paris, 98*(4-6), 374-384. doi: 10.1016/j.jphysparis.2005.09.006

Shanahan, M. P. (2006). A Cognitive Architecture that Combines Internal Simulation with a Global Workspace. *Consciousness and Cognition, 15*, 433-449.

Shiffrin, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science; Science*.

Silverta, L., Delplanquea, S., Bouwalerha, H., Verpoorta, C., & Sequeira, H. (2004). Autonomic responding to aversive words without conscious valence discrimination. *Int. J. Psychophysiol., 53*, 135–145.

Sims, C. R., & Gray, W. D. (2004). *Episodic versus semantic memory: An exploration of models of memory decay in the serial attention paradigm*. Paper presented at the 6th international conference on cognitive modeling (ICCM-2004), Pittsburgh, PA.

Sloman, A. (Ed.). (1999). *What Sort of Architecture is Required for a Human-like Agent?* Dordrecht, Netherlands: Kluwer Academic Publishers.

Smith, C. A., & Kirby, L. D. (2001). Toward delivering on the promise of appraisal theory. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, Methods, Research* (pp. 121-138). New York: Oxford University Press.

Snaider, J., McCall, R., & Franklin, S. (2009). *Time Production and Representation in a Conceptual and Computational Cognitive Model*. Paper presented at the AAAI Fall Symposium on Biologically Inspired Cognitive Architecture, Washington, DC.

Snaider, J., McCall, R., & Franklin, S. (2011). *The LIDA Framework as a General Tool for AGI*. Paper presented at the The Fourth Conference on Artificial General Intelligence (Springer Lecture Notes in Artificial Intelligence), Mountain View, California, USA.

Stickgold, R., & Walker, Matthew P. (2005). Memory consolidation and reconsolidation: what is the role of sleep? *Trends Neurosci., 28*, 408–415.

Strain, S. F., Franklin, S., Heck, D. H., & Baars, B. J. (in preparation). Brain rhythms, cognitive cycles and mental moments.

Sukthankar, G. (2000). *Face recognition: a critical look at biologically-inspired approaches*: Carnegie Mellon University, the Robotics Institute.

Sun, R., & Franklin, S. (2007). Computational Models of Consciousness: A Taxonomy and some Examples. In P. D. Zelazo & M. Moscovitch (Eds.), *Cambridge Handbook of Consciousness* (pp. 151–174). New York: Cambridge University Press.

Sun, R., & Naveh, I. (2004). Simulating organizational decision-making using a cognitively realistic agent model. *Journal of Artificial Societies and Social Simulation, 7*(3).

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review, 112*(1), 159-192.

Taylor, J. G. (2011). A Review of Models of Consciousness. In V. Cutsuridis, A. Hussain & J. G. Taylor (Eds.), *Perception-Action Cycle: Models, Architectures, and Hardware* (pp. 335-357).

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *Biological Bulletin, 215*, 216-242,.

Tort, A. B. L., Komorowski, R. W., Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009). Theta–gamma coupling increases during the learning of item–context associations. *PNAS, 106*(49), 20942-20947

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301-306.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT press.

Wallace, R. (2005). Consciousness: A Mathematical Treatment of the Global Neuronal Workspace Model. New York: Springer.

Wamsley, E. J., Tucker, M., Payne, J. D., Benavides, J. A., & Stickgold, R. (2010). Dreaming of a learning task is associated with enhanced sleep dependent memory consolidation. *Current Biology*.

Wang, P., Goertzel, B., & Franklin, S. (2008). *Artificial General Intelligence 2008*. Amsterdam: IOS Press.

Wiedemann, C. (2007). Memory consolidation...while you are sleeping. [10.1038/nrn2084]. *Nat Rev Neurosci, 8*(2), 86-87.

Yang, J., Xu, X., Du, X., Shi, C., & Fang, F. (2011). Effects of unconscious processing on implicit memory for fearful faces. *PLoS ONE, 6*(2), e14641.

Yerkes, R. M., & Dodson, J. D. (1908). The Relationship of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurology and Psychology, 18*, 459–482.

Zhang, Q. (2009). A computational account of dreaming: Learning and memory consolidation. *Cognitive Systems Research, 10*(2), 91-101