Global Workspace Theory, its LIDA Model and the Underlying Neuroscience

Stan Franklin, Steve Strain, Javier Snaider, Ryan McCall, Usef Faghihi

Abstract

A biologically inspired cognitive architecture must draw its insights from what is known from animal (including human) cognition. Such architectures should faithfully model the high-level modules and processes of cognitive neuroscience. Also, biologically inspired cognitive architectures are expected to contribute to the BICA "challenge of creating a real-life computational equivalent of the human mind." One unified theory of cognition, Global Workspace Theory (GWT) has emerged as the most widely accepted, empirically supported theory of the role of consciousness in cognition. Recent experimental studies reveal rich cortical connectivity capable of supporting a large-scale dynamic network. We propose that brains in fact cyclically and dynamically form such a network according to GWT. The biologically inspired LIDA cognitive architecture implements GWT conceptually and computationally. Here we argue that the LIDA architecture's breadth, flexible motivations using feelings, explicit attention mechanism, and continual, incremental and online learning in several modalities provide a significant first step in the direction of the BICA challenge. We also measure LIDA against the architectural features listed in the BICA Table of Implemented Cognitive Architectures. Applying recent brain connectivity results, we go on to elucidate the relationship between LIDA and the underlying and motivating neuroscience, using the language of non-linear dynamics. In particular, we claim that LIDA's representations correspond to basins of attraction in the non-linear dynamics of neural activation patterns. In addition, we claim that the rhythms of LIDA's cognitive cycle and of its internal cognitive elements have definite psychophysiological corollaries in the oscillatory patterns observed in the human brain.

Motivation and summary

The BICA journal focuses on biologically inspired cognitive architectures. A cognitive architecture can be thought of as a computational formalism that implements a unified theory of cognition in the sense of Newell (1990). Such cognitive architectures must aspire to account for the full range of cognitive processing from sensory input to motor output. A biologically inspired cognitive architecture must draw its insights from what is known from animal (including human) cognition. That is, the structures and processes comprising the cognitive architecture should be tested against empirical studies of humans and other animals. Such studies are the products of cognitive science and cognitive neuroscience. To be truly biologically inspired, such architectures should faithfully model the high-level modules and processes of cognitive neuroscience, though they need not model the low-level neural representations and mechanisms. Whereas cognition in humans and other animals is implemented in brains, cognitive architectures typically do not attempt to model neural systems *per se*, but rather work from functional conceptual models. This stance leaves cognitive modelers, the designers of cognitive architectures, with the problem of explaining how their high-level structures and processes might correspond to those in an underlying neural system. Finally, biologically inspired cognitive architectures are expected to contribute to the BICA "challenge of creating a real-life computational equivalent of the human mind."[1]

Supported by considerable empirical evidence (e.g. Baars, 2002), one such unified theory of cognition, Global Workspace Theory (GWT) (Baars, 1988) has emerged as the most widely accepted theory of the role of consciousness in cognition (Connor & Shanahan, 2010; Dehaene & Naccache, 2001; Glazebrook & Wallace, 2009; Schutter & van Honk, 2004; Sergent & Dehaene, 2004; Seth, 2007; Shanahan & Baars, 2005; Wallace, 2007). Recent experimental studies reveal rich cortical connectivity capable of supporting a large-scale dynamic network (Hagmann et al., 2008; Shanahan, 2010; van den Heuvel & Sporns, 2011). We propose that brains in fact cyclically and dynamically form such a network according to GWT, allowing for highly flexible, rapid reorganization of the

[1] http://www.elsevier.com/wps/find/journaldescription.cws_home/727718/description#description

neural state in accordance with the demands of an open, unpredictable, and at times dangerous environment.

The biologically inspired LIDA[2] cognitive architecture (Franklin, Baars, Ramamurthy, & Ventura, 2005; Franklin & Patterson, 2006) implements GWT conceptually (Baars & Franklin, 2003; Baars & Franklin, 2007) and computationally (Snaider, McCall, & Franklin, 2011), as well as other theories from cognitive science and neuroscience including situated (embodied) cognition (Glenberg & Robertson, 2000; Varela, Thompson, & Rosch, 1991), perceptual symbol systems (Barsalou, 1999), working memory (Baddeley & Hitch, 1974), memory by affordances[3] (Glenberg, 1997), long-term working memory (Ericsson & Kintsch, 1995), and transient episodic memory (Conway, 2002). Here we argue that the LIDA architecture's breadth, flexible motivations using feelings, explicit attention mechanism, and continual, incremental and online learning in several modalities provide a significant first step in the direction of the BICA challenge mentioned above. We also measure LIDA against the architectural features listed in the BICA Table of Implemented Cognitive Architectures[4]. Applying the brain connectivity results referred to earlier, we go on to elucidate the relationship between LIDA and the underlying and motivating neuroscience, using the language of non-linear dynamics (Edelman & Tononi, 2000; Kelso, 1995; Skarda & Freeman, 1987). In particular, we claim that LIDA's representations can be thought of as corresponding to basins of attraction in the non-linear dynamics of neural activation patterns. In addition, we claim that the rhythms of LIDA's cognitive cycle and of its internal cognitive elements have definite psychophysiological corollaries in the oscillatory patterns observed in the human brain. We hold cognition to be a biological adaptation that self-organizes mental representations of the world across the interface between an individual and its environment as well as within the organism itself.

In order to specify the motivation and scope of our model, it is important that we assert the following:

1. Our hypotheses regarding the ways in which brains generate minds are quite speculative. There is a great deal of conflicting evidence, differing philosophical positions, and lack of theoretical convergence in the field of cognitive neuroscience. As a result of this, we find objections to our neural hypotheses equally speculative.
2. LIDA is model of minds, not a model of brains. Our studies of empirical neuroscience serves as a part of the theoretical basis for our model. The recently initiated experimental arm of our work will test our model's ability to replicate the results of psychological experiments as well as its suitability for the design of autonomous agents that use human-like approaches to problem solving. We have no intention of, nor will our work be capable of, directly testing our neuroscientific hypotheses. In other words, we neither require nor expect that LIDA's internal dynamics correspond to those of brains, only that its functions match the cognitive functions of minds.

Regarding the above points, one may ask 1) Why it is necessary to take intellectual positions that are inconclusive? and 2) Why assert hypotheses that won't be tested directly?

In response to the first objection, we note that little prior work has attempted to integrate a model of mind across as broad a theoretical base as we have attempted here. Nonetheless, we consider it important that such attempts begin. We are trying to construct a broad functional map of the mind, hoping to get the "continents" right, and not troubling excessively about the exact shape of the "coastline" (Bach 2009, "Seven Principles for a Synthetic Intelligence"). We note that a watertight aquarium serves well as a container for fish, but not at all for catching them. For the latter, a net with the right size holes is quite effective. There will certainly be places that we get it wrong. We believe our mistakes will be as informative as our successes.

In response to the second, we reiterate that results from *experimental* neuroscience is a *theoretical* part of our model. They inform us as to how evolution has implemented minds on brains. We wish to build an alternative implementation of the very same minds that nature provides. To that

---

[2] LIDA (Learning IDA) developed from IDA, a working software agent that that finds new jobs for sailors at the end of a tour of duty (Franklin, Kelemen, & McCauley, 1998).

[3] Gibson (1979) introduced the term *affordance*, which is often interpreted as meaning that information about the available uses of an object *exists* in the object itself. We are using it in the sense that the agent can *derive* such information from the object.

[4] http://bicasociety.org/cogarch/architectures.htm

end, it is quite helpful to attempt to understand and to integrate as well as possible, all that we can learn from the mind's biological implementation. Our hypotheses about low-level neural dynamics provide us a high-level theoretical inspiration.

Global Workspace Theory

Global Workspace Theory (Baars, 1988) was originally conceived as a neuropsychological model of conscious and unconscious processes, but has been broadened into a higher-level theory of human cognitive processing, supported by empirical evidence (Baars, 2002). GWT views the nervous system as a distributed parallel system with many different specialized processes. 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. GWT is therefore a theory of how consciousness functions within cognition.

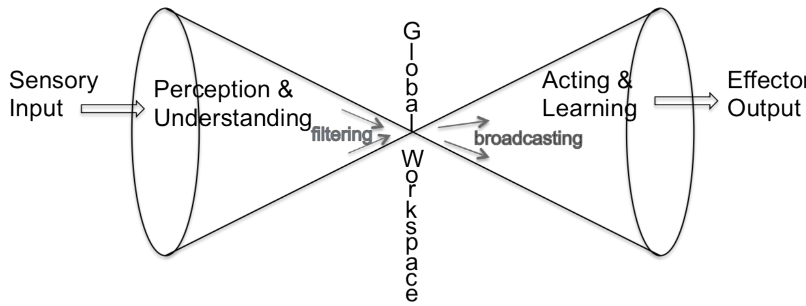

Figure 1 – GWT Cognitive Cycle

GWT postulates that the brain's multitude of relatively small, special purpose processes are almost always unconscious (Franklin & Baars, 2010). Communication between them is rare and occurs over a narrow bandwidth. Coalitions of processes can be added to the global workspace (and perhaps into consciousness). This limited capacity workspace serves to broadcast the contents of the coalition to all the unconscious processors, in order to recruit other processors to join in handling the current situation, or in solving the current problem. Thus consciousness in this theory enables the agent to deal with novel or problematic situations that cannot be dealt with efficiently, or at all, by habituated unconscious processes. In particular, consciousness provides access to appropriately useful resources, thereby solving the *relevance problem* (Franklin, 2003a), constituting a major function of consciousness in cognition. A second major function of consciousness in cognition is the enabling of learning, the encoding of knowledge about the past for use in the present. GWT supports the Conscious Learning Hypothesis: significant learning takes place via the interaction of consciousness with the various memory systems (Baars, 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 below for a discussion of cognitive cycles). Data flow according to GWT can be visualized as having an hourglass shape with sensory data coming in the top and flowing through the upper chamber. 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 bottom chamber. GWT is therefore a theory of how consciousness functions within cognition, first as a filter and then as a recruiter (please see Figure 1). Unconscious contexts influence this competition for consciousness (please see Figure 2). Thus GWT focuses on the function of consciousness in solving the relevance problem, that is, in providing the system with access to its internal resources that are most relevant to the current situation. Learning, which requires only attention, occurs with each conscious broadcast. In summary, GWT is a theory of how consciousness functions within cognition in the broad sense of the term.

Connectivity in brains and GWT

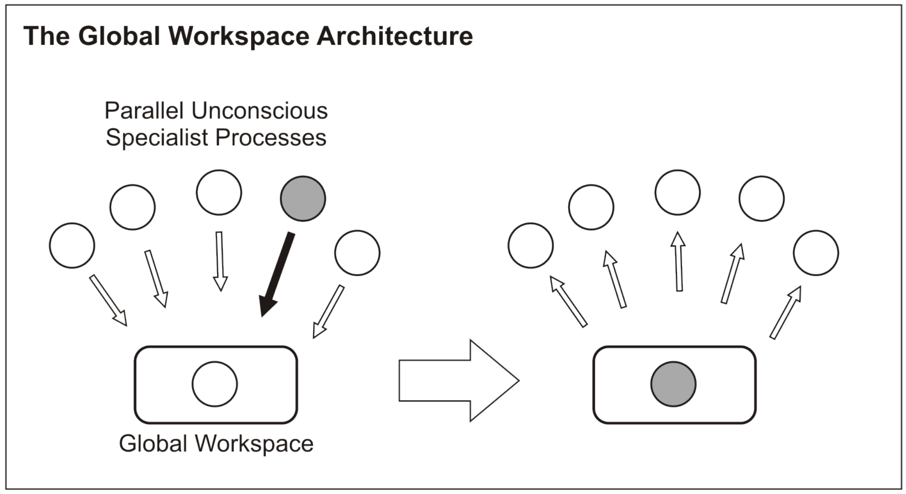

Figure 2 – The Global Workspace Architecture (adapted from Shanahan 2010)

Brains can be reasonably thought of as networks composed of nodes (neurons) and links (synapses). A brain *module* would be a cluster of neurons densely connected internally, but sparsely connected externally. Shanahan describes the sort of network structure that would allow a GWT architecture to be implemented within a brain (Shanahan, 2010). He argues that a *"small world" network* is needed. A small world network is one that is densely connected locally, sparsely connected globally, and has a short path between any two nodes. He further asserts that this small world property is needed at multiple scales of organization of modules and submodules, say from cell assembly modules (small processes in GWT), up to much larger cognitive modules. This implies that the network should be "hierarchically modular." Finally, these networks should be liberally provided with *connector hubs*, through which many paths pass from one module to another. These connector hubs, together with their interconnections, constitute the *connective core* of the network. Shanahan argues that a hierarchically modular small world network structure provided with a connective core constitutes the communicative infrastructure in brains needed to implement GWT. Such a network structure enables small processes to influence the global system by being broadcast.

Do brains have the kind of network infrastructure so described as needed to support GWT? At this point Shanahan's arguments become empirical, resting heavily on an extensive neuroscience literature. Hagmann and colleagues have mapped out "a dense network of fiber pathways linking all regions of the cerebral cortex." (2008) They have further found connector hubs linking these regions (structural modules) and constituting a connective core. More recent work has confirmed the existence of these connector hubs, both neocortical and subcortical, and that the regions they link are, indeed, modules with more dense internal connections than external (van den Heuvel & Sporns, 2011). These studies, and others, strongly suggest that the brain contains a hierarchically modular, small world network with a connective core that Shanahan claims "… is capable of globally disseminating the influence of a process or coalition of processes," in such a way that "only one coalition of processes at a time can take over the connective core, to the exclusion of its rivals." Thus the brain provides just the sort of network infrastructure required to enable GWT. Shanahan also suggests that the connective core produces a bottleneck in the transmission of information among processes. Although this may appear to be a negative consequence, he maintains that it is actually positive. This bottleneck helps to ensure that only one process takes control of limited resources (for example actuators), enforces that action selection is mediated by the winning coalition, and promotes serial processing, which is central for the chaining of certain mental operations, such as planning.

The LIDA architecture and its cognitive cycle

The LIDA model is a comprehensive, conceptual and computational model that covers a large portion of human cognition while implementing and fleshing out GWT. The model and its ensuing architecture are grounded in the LIDA cognitive cycle. The cycle is based on the fact that every autonomous agent (Franklin & Graesser, 1997), be it human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). The agent's "life" can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle is comprised of phases of understanding, attending and acting. Neuroscientists call this three-part process the action-perception cycle. A cognitive cycle can be thought of as a cognitive "moment." Sophisticated agents such as humans process (make sense of) the input from such sampling in order to facilitate their decision making. Higher-level cognitive processes are composed of many of these cognitive cycles, each a cognitive "atom."

Just as atoms have inner structure, the LIDA model hypothesizes a rich inner structure for its cognitive cycles (Baars & Franklin, 2003; Franklin, et al., 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 both external and internal features of its world. By a competitive process to be described below, it then decides what portion of the represented situation is most salient, that is, in the most in need of attention. This portion is broadcast, making it the current contents of consciousness, and enabling the agent to choose an appropriate action and execute it. Though GWT speaks to phenomenal consciousness, and thus to the "hard problem" of consciousness (Chalmers, 1996), the LIDA model follows Shanahan (2010) in taking the "post-reflective inner view" and doing "… without the habit of metaphysical thinking." More specifically, consciousness in the LIDA model refers to *functional consciousness* in which no assumption is made of the conscious broadcast being phenomenally conscious (Franklin, 2003b).

Figure 3 shows the process in more detail. The cycle begins with sensory stimuli from external and internal sources in the agent's environment. Low-level feature detectors in Sensory Memory begin the process of making sense of the incoming stimuli. These low-level features are passed on to Perceptual Associative Memory (also called recognition memory) where higher-level features, such as objects, feelings, events, categories, relations etc. are recognized. These entities, which have been recognized preconsciously, make up the percept that passed asynchronously to the workspace, where a model of the agent's current situation is assembled. This percept serves as a cue to two forms of episodic memory, transient and declarative. Responses to the cue (recalls) consist of remembered events from these two memory systems that were associated with the various elements of the cue. In addition to the current percept, the Workspace contains recent percepts and portions of the structures assembled from them that haven't yet decayed away.

A new model of the agent's current situation is assembled from the percepts, the recalls, and the undecayed parts of the previous model. This assembling process will typically require structure-building codelets[5]. These structure-building codelets are small, special purpose processors, each of which has some particular type of structure it is designed to build. To fulfill their task these codelets may draw upon Perceptual Associative Memory and even Sensory Memory, to enable the recognition of relations and situations. They may also draw on the Conscious Contents Queue, which stores the conscious contents of the past few seconds. The newly assembled model constitutes the agent's understanding of its current situation within its world. It has made sense of the incoming stimuli.

For an agent operating within a complex, dynamically changing environment, this Current Situational Model (CSM) may well be much too much for the agent to consider all at once in deciding what to do next. It needs to selectively attend to a portion of the model. Portions of the CSM compete for attention. These competing portions take the form of coalitions of structures from the CSM. Such coalitions are formed by attention codelets, whose function is to try to bring certain structures to consciousness. When one of the coalitions wins the competition, the agent has effectively decided on what to attend.

---

[5] The term codelet refers generally to any small, special purpose processor or running piece of computer code. The concept is essentially the same as Baars' processors (1988), Minsky's agents (1985), Jackson's demons (1987) or Ornstein's small minds (1986). The term was borrowed from Hofstadter and Mitchell (1995).

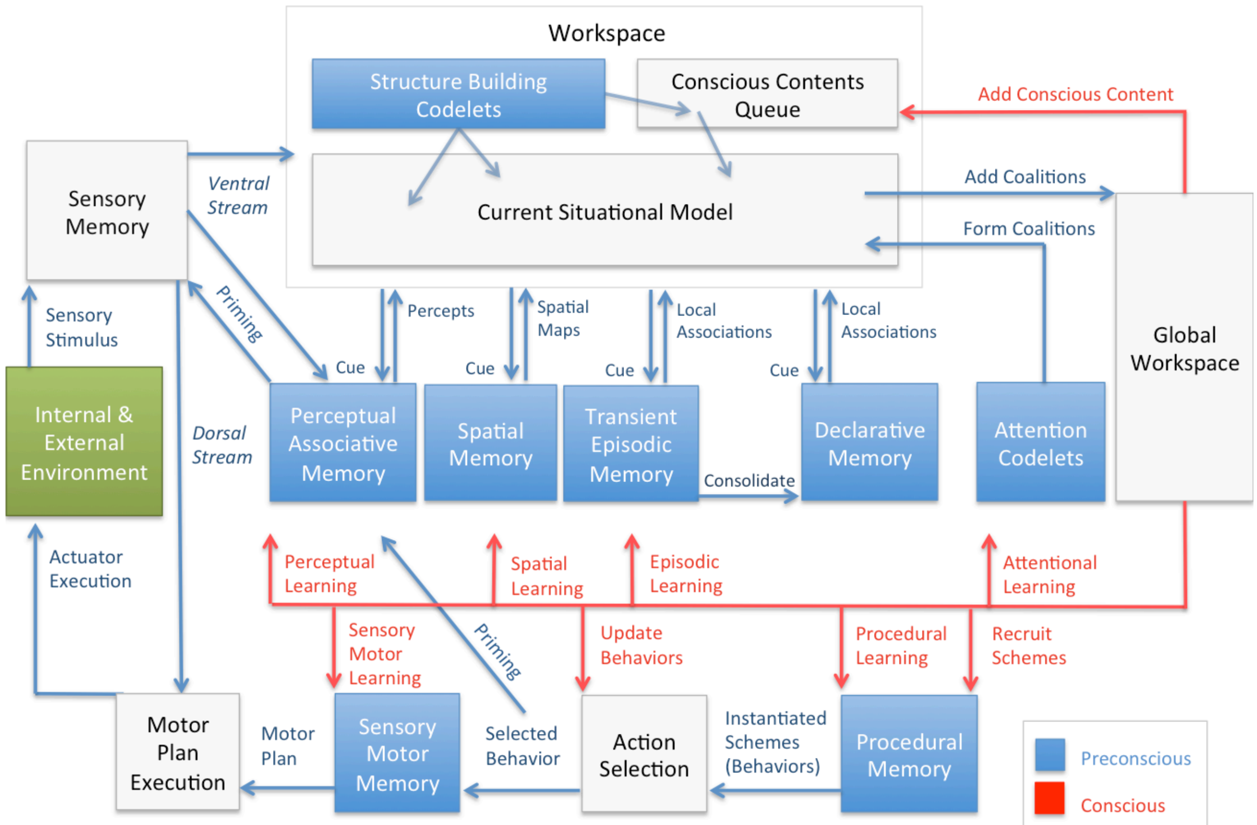

Figure 3 – The LIDA Cognitive Cycle Diagram

One purpose of this processing is to help the agent decide what to do next. To this end, a representation of the contents of the winning coalition is broadcast globally (hence the name Global Workspace Theory). Though the contents of this conscious broadcast are available globally, a primary recipient is Procedural Memory, which stores templates of possible actions including their contexts and possible results. It also stores an activation value that attempts to measure, for each such template, the likelihood that an action taken within its context produces the expected result. Templates whose contexts intersect sufficiently with the contents of the conscious broadcast instantiate copies of themselves with their variables specified to the current situation. Instantiated templates remaining from previous cycles may also continue to be available. These instantiations are passed to the action selection mechanism, which chooses a single action from one of them. The chosen action then goes to Sensory-Motor Memory, where it is executed by an appropriate algorithm, called a motor plan. The action taken affects the environment, and the cycle is complete.

Another function of LIDA's cognitive cycle is to facilitate learning. LIDA's multiple modes of learning all occur continually, simultaneously, and online using each global broadcast of the contents of consciousness (Franklin, et al., 2005; Franklin & Patterson, 2006). Perceptual learning is learning to recognize objects, categorizations, relationships, events, etc. As new objects, categories, and the relationships among them and between them and other elements of the agent's ontology are learned, nodes (objects and categories) and links (relationships) are added to Perceptual Associative Memory (Figure 3). Spatial learning refers to the building and updating of cognitive maps which serve to locate objects in the environment (Derdikman & Moser, 2010). Episodic learning is the encoding of information into episodic memory, the associative, content-addressable, memory for events -- the what, the where, and the when (Baddeley, Conway, & Aggleton, 2001; Franklin, et al., 2005). Relatively little studied by memory theorists, attentional learning refers to the learning of what to pay attention to. In the LIDA architecture attentional learning is the learning of new attention codelets and the updating and reinforcing of the existing ones. Procedural learning is the encoding of procedures for executing behaviors into Procedural Memory (Figure 3) It is the learning of new actions and action sequences with which to accomplish new tasks (D'Mello & Franklin, 2004). Here we must distinguish between action selection and action execution. LIDA's Procedural Memory is composed of schemes concerned with the selection of actions. Algorithms (motor plans) for the execution of actions are found in Sensory-Motor Memory where sensory-motor learning takes place.

The LIDA model hypothesizes that all human cognitive processing is via a continuing iteration of such cognitive cycles. These cycles occur asynchronously, with each cognitive cycle taking roughly 300 ms (Madl, Baars, & Franklin, 2011). These cycles cascade, that is, several cycles may overlap having their currently active processes at different stages of the cycle running simultaneously in parallel. This cascading must, however respect the serial nature of conscious processing necessary to maintain the stable, coherent image of the world it provides (Franklin, 2005; Merker, 2005). This cascading, together with the asynchrony, allows a rate of cycling in humans of five to ten cycles per second. A cognitive "moment" is thus quite short! There is considerable empirical evidence from neuroscience suggestive of and consistent with such cognitive cycling in humans (Doesburg, Green, McDonald, & Ward, 2009; Madl, et al., 2011; Massimini et al., 2005; Sigman & Dehaene, 2006; Uchida, Kepecs, & Mainen, 2006; Willis & Todorov, 2006). None of this evidence is conclusive, however.

LIDA and the BICA challenge

The BICA challenge is that of "*… creating a real-life computational equivalent of the human mind."* Meeting such a challenge will require a cognitive architecture that provides flexible motivations, an explicit attention mechanism, and continual, incremental and online learning in several modalities making it suitable for controlling cognitive software agents through a substantial developmental period (Weng et al., 2001). Some of the design principles underlying the LIDA architecture, and that differentiate it from other cognitive architectures, should make it eminently suitable for controlling agents aspiring to the BICA challenge. The LIDA architecture adheres to the principles of grounded cognition (Barsalou, 2008), which emphasize the importance of modal representations, situated action, and perceptual simulation. It employs artificial feelings and emotions that allow for flexible and sophisticated action selection, and for the modulation of learning (Franklin & Ramamurthy, 2006). The architecture includes an explicit functional consciousness mechanism that plays a major role in perceptual filtering, action selection, and learning. Each of the various modes of learning in the model follows the *principle of profligacy*. This means that new representations are added to the various memories at the slightest justification, that is, whenever they come to consciousness, where they will either survive due to future reinforcement or not because of decay. Finally, the Cognitive Cycle Hypothesis, implemented in the LIDA architecture, follows human cognitive functioning by means of similar flexible cognitive cycles occurring at a rate of some 5hz to 10hz (Franklin, et al., 2005). Building higher-level cognitive processes operating across multiple cognitive cycles should prove a useful strategy for developing cognitive software agents aspiring to the BICA challenge.

Cognitive architectures in general, and the LIDA architecture in particular, are complex from two points of view: the underlying theory tends to be inherently complicated and, consequently, any software implementation is also very complex (Snaider, et al., 2011). Cognitive architectures are typically composed of various modules with distinct functionalities, and in many cases, with different algorithmic implementations. This complicates the implementation of software agents based on them. Developers expend significant resources re-implementing common functionalities for each new agent implementation. Code reuse between architectures has been difficult in general because of lack of standardization and ill-defined modules. Software frameworks are ideal tools with which to solve many of the problems. They allow developers to focus on their particular algorithms instead of implementation details common to many agents. The architecture can be understood more quickly because the framework itself provides a higher level of abstraction than unitary code, speeding up the development, abstracting the complexity of the implementation, which, in turn, allows more effective and accurate communication between researchers.

The LIDA Framework (Snaider, et al., 2011), a generic and customizable computational implementation of the LIDA model, was designed with all these framework advantages in mind. It is implemented in Java, a strong and proven object-oriented language. Its design and implementation aim to simplify the process of implementing a cognitive architecture allowing the user to concentrate on the specifics of the application, hiding the complexities of the generic parts of the model. It also enforces good software practices that simplify the implementation process. It achieves a high level of abstraction, permitting a variety of customization options while maintaining a low level of coupling between modules. The LIDA Framework was conceived with multithreading support in mind — biological minds operate in a parallel manner, so should artificial ones.

The current implementation of the LIDA computation framework includes its generic structure, several tools, such as a customizable GUI and a logging mechanism, and implementations for all the modules of the LIDA model (please see Figure 3). Many of them are mature: Workspace, Episodic Memory, Global Workspace, and Procedural Associative Memory. Others, such as Procedural Memory, Action Selection and Sensory Motor Memory, have simple implementations and are the focus of current work. Finally, some of the modules, such as Sensory Memory, are domain dependent, and must be designed and implemented in each case. Nevertheless, the framework provides a template implementation of these modules. Episodic Memory has its learning mechanism already implemented. Other modules, such as Procedural Memory and Perceptual Associative Memory, have basic learning support, and new algorithms are forthcoming. The modular conception of the framework allows any module implementation to change without greatly affecting the others. We believe that tools like the LIDA Framework are as fundamental to addressing the BICA challenge as the underlying theory itself. Some LIDA-based agents have replicated experiments using the LIDA Framework thus demonstrating its effectiveness. For simple examples, see the Framework tutorial[6], which is part of the framework package; for more advanced ones, please see (Madl, et al., 2011; Madl & Franklin, 2012)

LIDA features and the BICA Table

In the following, we will give an assessment of the LIDA model against the features of the BICA Table of Implemented Cognitive Architectures (Samsonovich, 2010a; Samsonovich, 2010b). Many of the theoretical features of the LIDA model are already implemented in the LIDA computational Framework. Most others have been designed and are in the process of being implemented. Column 1 of the BICA Table contains a list of features proposed by developers of cognitive architectures to be at least potentially useful, if not essential, for the support of biologically inspired cognitive architectures. Subsequent columns are devoted to individual cognitive architectures with a cell describing how the column architecture addresses the row feature. The rest of this section is an expansion of the column devoted to LIDA in the BICA Table. Note that all the **Basic overview** features such as episodic memory listed in the Table's first column are detailed above. We will discuss the rest of the features in what follows:

- **Support for Common Components:** *The LIDA Model supports all features mentioned in this section, such as episodic and semantic memories. However, the auditory mechanism is not implemented in a LIDA-based agent as yet.*
- **Support for Common Learning Algorithms:** *the LIDA Model supports different types of learning such as episodic, perceptual, procedural, and attentional learning. However, the Bayesian Update and Gradient Descent Methods (e.g., Backpropagation) have not been implemented in a LIDA-based agent. The LIDA model does not include Bayesian update, but its learning can be thought of as an approximation to such methods. LIDA's stance on learning as continual incremental and online stands at odds with Backpropagation methods.*
- **Common General Paradigms Modeled:** *the LIDA Model supports features listed in this section such as decision-making and problem solving. However, perceptual illusions, meta-cognitive tasks, social psychology tasks, personality psychology tasks, motivational dynamics have not been implemented in a LIDA-based agent*.

[6] The LIDA Framework and its tutorial are open source for research projects. They can be downloaded from the authors' web site: http://ccrg.cs.memphis.edu/framework.html

- **Common Specific Paradigms Modeled columns:** 1) Stroop; 2) Task Switching; 3) Tower of Hanoi/London; 4) Dual Task; 5) N-Back; 6) Visual perception with comprehension; 7) Spatial exploration; 8) Learning and navigation; 9) Object/feature search in an environment; 10) Learning from instructions; 11) Pretend-play. *Although the Common Specific Paradigms Modeled features listed above are not implemented in LIDA, in principle LIDA is capable of implementing each of them. For instance Madl et al. (2011)* developed two LIDA-based agenta, the first of which matched its cognitive cycle to the human action-perception cycle achieving human-like reaction times on simple reaction time experiments. The second agent modeled phenomenal simultaneity within timeframes comparable to human subjects. Madl's work represented a first step in identifying and tuning the internal parameters of the LIDA model. Continuing the development of such parameterization, Faghihi et al. (in review) successfully replicated an attentional task (Van Bockstaele, Verschuere, De Houwer, & Crombez, 2010) with a LIDA-based agent. In this task, as in the human experiment, the subject was required to respond to a target that appeared after a cue on a computer monitor. The target appeared on either the same side as the cue (congruent) or on the opposite side (incongruent). Most recently, a LIDA-based agent replicated several attentional blink phenomena (Madl & Franklin, 2012).

Meta-Theoretical Questions:

1) Uses only local computations? *Yes, throughout the architecture with the one exception of the conscious broadcast which is a local process that is transmitted almost globally*;

2) Unsupervised learning? *Yes. The LIDA Model supports four different modes of online, unsupervised learning: perceptual, episodic, attentional and procedural*; *spatial learning is a topic of current research;*

3) Supervised learning? *While in principle possible for a LIDA agent, supervised learning per se is not part of the architecture*;

4) Can it learn in real time? *Yes (see above)*;

5) Can it do fast stable learning; i.e., adaptive weights converge on each trial without forcing catastrophic forgetting? *Yes. One shot learning in several modes occurs with the conscious broadcast during each cognitive cycle. With sufficient affective support and/or sufficient repeated attention, such learning can be quite stable*

6) Can it function autonomously? *Yes. A LIDA-based agent can, in principle, autonomously operate machines and drive vehicles, for example;*

7) Is it general-purpose in its modality; i.e., is it brittle? *A LIDA-based agent can, in principle, be developed to be general purpose and robust in real world environments;*

8) Can it learn from arbitrarily large databases; i.e., not toy problems? *While LIDA's perception is still nascent, it is capable in principle of processing real-world data. LIDA learns profligately with each cognitive cycle. However only common, recurring information that is reinforced will survive its continual decay process;*

9) Can it learn about non-stationary databases, i.e., environmental rules change unpredictably? *Yes, a LIDA-based agent is, in principle, capable of working properly in an unpredictable environment;*

10) Can it pay attention to valued goals? *Yes, through its explicit attention mechanism.*

11) Can it flexibly switch attention between unexpected challenges and valued goals? *Yes. A LIDA-based agent attends to what is currently most salient based on its situational awareness;*

12) Can reinforcement learning and motivation modulate perceptual and cognitive decision-making? *Yes;*

13) Can it adaptively fuse information from multiple types of sensors and modalities? *In principle, yes, but it has yet to be implemented in particular domains with multiple senses.*

Self-organizing Neurodynamics and the LIDA Model

Just as chemical theory must rest on the underlying physics, every comprehensive model of cognition must be grounded in the underlying neuroscience. However, one would not attempt to account for chemical reactions in terms of atomic and subatomic particles. Nor are cognitive processes easily explained in terms of the behavior of neurons and cell assemblies. A higher-level theory with a more abstract ontology of cognitive modules and processes is required. The LIDA Model is such a higher-level theory, fleshing out Global Workspace Theory, as well as providing a computational architecture for it.

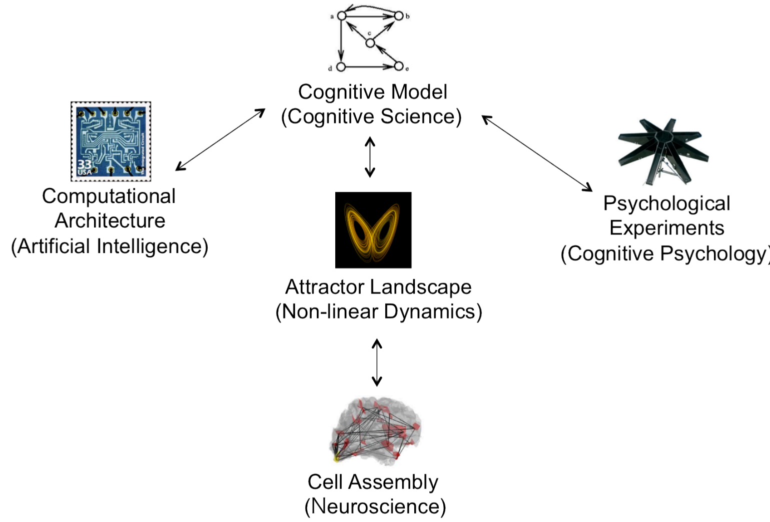

Figure 4 – Theoretical Levels of Abstraction

How is the grounding of the LIDA Model in neuroscience to be accomplished? Perceptual symbols (Barsalou, 1999) in the form of nodes and links in its perception module (Perceptual Associative Memory) comprise the common representational currency of the LIDA Model. To ground these perceptual symbols in the underlying neuroscience, we think of them as representing not neurons or cell assemblies, but rather wings of chaotic attractors in an attractor landscape (Freeman, 1999; Harter, Graesser, & Franklin, 2001; Skarda & Freeman, 1987). When perturbed by a previously learned exogenous stimulus such as one that may result from an inhalation, the spiking trajectory of a cell assembly, such as an olfactory bulb, falls into a wing of an attractor and so recognizes an odor. Thus we postulate non-linear dynamics as an intermediate theory serving to ground comprehensive cognitive models such as LIDA in the underlying neuroscience (see Figure 4).

Cognitive self-organization

According to the neural hypotheses of the LIDA Model, cognition flexibly composes itself within each cognitive cycle by integrating neural activity across spatial, temporal, and modal domains for the creation of high-level representations of objects, events, and plans that are relevant to an agent's current needs, wants, and interests. While this integration is occurring, each modality maintains its responsiveness to rapid and unpredictable changes in the environment, allowing for the possibility of a suitably sudden reorganization of cognitive priorities.

We propose phase-coupling[7] between high and low frequency EEG bands as a general neural mechanism for this integration. Furthermore, we propose that self-organizing dynamics constitutes 1) the bridge between the mind's cognitive representations and the neural dynamics that underlie them, as well as 2) the substrate for the rapid neurocognitive reorganizations, or phase transitions, that must frequently occur in response to sudden and unforeseen changes in environmental states (Buzsaki, 2006; Lewis, 2005).

Within the context of our neural hypotheses, we model both environmental and agent variables as chaotic oscillators in their respective state spaces, and to categorize these oscillators into high-energy and low-energy categories according to a scheme proposed by Barham and elaborated by Freeman (Barham, 1996; Freeman, 2003). Examples of low-energy oscillators are visual, auditory, or chemical stimuli, sensory signals, and neural activity patterns. Examples of high-energy oscillators are large objects or environmental features such as terrain, trees or walls; predator/prey populations or food/water supplies; and motor actuators and effectors.

As cognition develops over the lifetime of an agent, the brain establishes low-energy perceptual oscillators in its neurodynamics that couple with certain "meaningful" low-energy oscillators in the

[7]By "phase-coupling," we refer to a detectable relationship between the timing patterns of two signals.

environment. These low-energy environmental oscillators are meaningful by virtue of their tight association with high-energy environmental oscillators that impinge on the survival and success of the agent. Thus, the low-energy environmental oscillators serve as reliable indicators of the state of pertinent high-energy oscillators. For instance, the proximity of a predator is modeled as a particular state in a high-energy environmental oscillator, and the presence of a characteristic odor is modeled as a specific state of a low energy environmental oscillator. A neural activity pattern in the olfactory bulb constitutes a low-energy perceptual oscillator coupled to the odor signal. In this way, an agent can detect and respond to the presence of a significant object before becoming mechanically or thermodynamically coupled with it (Barham, 1996; Freeman, 2003).

The brain dynamically integrates the activity of its perceptual oscillators with the activity of its higher-order neural oscillators 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 for broadcast. This broadcast then drives action selection and diverse forms of learning, and the selected actions activate the high-energy oscillators that control the organism's action execution, altering the state of environmental oscillators with varying degrees of observability and predictability. As will be elaborated below, it is within this higher-order integration that we hypothesize phase-coupling of EEG activity to play a pivotal role.

To summarize, we view cognition as a biological adaptation that self-organizes across the interface between an individual and its environment as well as within the organism itself. Furthermore, we propose timing relationships in the form of phase-coupling between oscillators as a key characteristic of cognition's neurophysiological "structure."

EEG phase-coupling and cognition

Temporal organization in the high frequency domain of local field potential (LFP) activity, also known as gamma synchronization, appears to be a "glue" for the bottom-up assembly of activity patterns into coordinated representations at a relatively microscopic spatiotemporal scale (Holz, Glennon, Prendergast, & Sauseng, 2010; Tallon-Baudry, 2009). Empirical evidence continues to accumulate in support of the importance of temporal correlations in LFP signals for the encoding of cognitive variables. 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, 2009a), attention (Sauseng, Klimesch, Gruber, & Birbaumer, 2008), and perceptual organization (Doesburg, et al., 2009).

Cross-frequency-coupling (CFC) could serve as a mechanism for top-down association of synchronized gamma assemblies—each representing an element of a cognitive representation—into a composite entity, e.g. a list of remembered items being held in working memory for a specific task (as in Axmacher et al., 2010). We hypothesize that phase coupling involving large portions of association cortical areas is the neurophysiological realization of the global broadcast in GWT.

Cognitive self-organization results in the "emergence and stabilization of psychological or neural configurations that correspond with (or represent) conditions in the world." (Lewis 2005, p. 173) Thus, psychology and cognitive neuroscience both depict the mind as a system that reorganizes itself and its own representations in response to a dynamic, partially knowable, and unpredictable environment. The rapid, flexible, and highly effective temporal patterning observed across the hierarchy of brain rhythms presents a natural substrate for the biological implementation of mental experience.

Theta-gamma coupling and the LIDA Model

Cognitive tasks in humans and animals, such as perception and memory, modulate oscillatory brain activity in various frequency bands, including both the theta (5–8 Hz) and gamma (30–150 Hz) bands. As discussed above, evidence has been found of the cross-coupling of these frequencies. More specifically, cross-frequency coupling was detected as a strong correlation between theta phase and gamma power (Canolty et al., 2006). Moreover, the amplitude of the fast gamma oscillations was systematically modulated during the course of a theta cycle. "Importantly, the cross-frequency coupling was modulated by [the] tasks." (Jensen & Colgin, 2007) Even more recently such cross coupling in the hippocampus has been linked to memory recall (Tort, Komorowski, Manns, Kopell, & Eichenbaum, 2009b). Specifically, "[hippocampal] theta rhythms correlate with intake of sensory information during movements such as whisking and sniffing in rats and may temporally segment samples of stimuli from the environment, with each theta cycle providing a discrete unit for sensory information processing in the brain." (Colgin & Moser, 2010)

Within the LIDA Model, individual cognitive cycles run their course in roughly 300ms. Several such cognitive cycles, perhaps three, can cascade or overlap as long as the seriality of consciousness and action selection 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 3), always varying with the current situation or task. An attractive conjecture from the LIDA Model would be 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. 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. As previously mentioned, the LIDA model hypothesizes that high level cognition e.g. planning, problem solving, reasoning, meta-cognition are accomplished through multiple cognitive cycles. Such complex phenomena are achieved by multiple cognitive cycles working iteratively. LIDA invokes James' ideomotor theory to control multi-cyclic volitional deliberation (Franklin, 2000). However, it is currently not clear what neural mechanisms beyond those discussed in this paper are required to implement this.

Moreover, this view fits nicely with Global Workspace Theory in the following extension of this conjecture. Coalitions assemble when neural sub-processes, encoded as high-frequency gamma activity, synchronize within a theta cycle through theta-gamma coupling. This synchrony enhances the processes' ability to affect other brain regions, an effect known as communication through coherence (CTC) (Fries, 2005). When one coalition wins the competition for consciousness, in a "winner-take-all" fashion it dynamically establishes a large-scale synchrony (in the sense of Bressler and Kelso, 2001; Tononi and Edelman, 2000; and Sporns, 1998) across multiple cortical areas, which realizes the global broadcast. This neural "coalition" might dominate for multiple cycles before diminishing in strength and being supplanted by another. However, while controlling the global workspace, its composition would be fluid and flexible in accordance with its current situational model. The coalescence and dissolution of coalitions, and the sequence of broadcast, would both manifest as phase transitions within the non-linear neurodynamics of the neural system.

Conclusion

Developers of biologically inspired cognitive architectures are faced with several challenges in validating their models. They must draw their insights from what is known from animal (including human) cognition. The structures and processes comprising such a cognitive architecture should be tested against empirical studies of humans and other animals. To be truly biologically inspired, such architectures should faithfully model the high-level modules and processes of cognitive neuroscience, though they need not model the low-level neural representations and mechanisms. However, functional conceptual models and their high-level structures and processes must be consistent with of the underlying neural system. Finally, biologically inspired cognitive architectures are expected to contribute to the BICA "challenge of creating a real-life computational equivalent of the human mind."

Recent experimental studies reveal rich cortical connectivity capable of supporting a large-scale dynamic network. We propose that brains in fact cyclically and dynamically form such a network according to Global Workspace Theory (GWT), allowing for highly flexible, rapid reorganization of the neural state in accordance with the fluctuations of an open, unpredictable, and at times dangerous environment.

We presented the biologically inspired LIDA cognitive architecture, which implements GWT conceptually and computationally. It was argued that LIDA's breadth, flexible motivations using feelings, explicit attention mechanism, and continual, incremental and online learning in several modalities provide a significant first step in the direction of the BICA challenge. We measured LIDA against the architectural features listed in the BICA Table of Implemented Cognitive Architectures. Applying the brain connectivity results referred to earlier, we described the relationship between LIDA and the underlying and motivating neuroscience, using the language of non-linear dynamics. In particular, we claimed that LIDA's representations could be thought of as corresponding to basins of attraction in the non-linear dynamics of neural activation patterns. In addition, we claim that the rhythms of LIDA's cognitive cycle and of its internal cognitive elements have definite psychophysiological corollaries in the oscillatory patterns observed in the human brain.

References

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences, 107*(7), 3228-3233. doi: 10.1073/pnas.0911531107

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, B. J. (2003). The global brainweb: An update on global workspace theory. *Science and Consciousness Review*(2).

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

Baars, B. J., & Franklin, S. (2007). An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA. *Neural Networks, 20*, 955-961.

Baddeley, A., Conway, M., & Aggleton, J. (2001). *Episodic Memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The Psychology of Learning and Motivation* (pp. 47–89). New York: Academic Press.

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

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

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology, 59*, 617–645.

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.

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

Chalmers, D. J. (1996). *The Conscious Mind*. Oxford: Oxford University Press.

Colgin, L. L., & Moser, E. I. (2010). Gamma oscillations in the hippocampus. *Physiology, 25*(5), 319-329.

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. (2002). Sensory-perceptual episodic memory and its context: autobiographical memory. In A. Baddeley, M. Conway & J. Aggleton (Eds.), *Episodic Memory* (pp. 53–70). Oxford: Oxford University Press.

D'Mello, S. K., & Franklin, S. P. (2004). A cognitive architecture capable of human like learning. *Connection Science*.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79*, 1–37.

Derdikman, D., & Moser, E. I. (2010). A manifold of spatial maps in the brain. *Trends in Cognitive Sciences, 14*(12), 561-569. doi: 10.1016/j.tics.2010.09.004

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
Edelman, Gerald M., & Tononi, G. (2000). *A Universe of Consciousness*. New York: Basic Books.
Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.
Faghihi, U., McCall, R., & Franklin, S. (in review). A Computational Model of Attentional Learning in a Cognitive Agent.
Franklin, S. (2000). Deliberation and Voluntary Action in 'Conscious' Software Agents. *Neural Network World, 10*, 505–521
Franklin, S. (2003a). An Autonomous Software Agent for Navy Personnel Work: A Case Study. In D. Kortenkamp & M. Freed (Eds.), *Human Interaction with Autonomous Systems in Complex Environments: Papers from 2003 AAAI Spring Symposium* (pp. 60–65). Palo Alto: AAAI.
Franklin, S. (2003b). IDA: A Conscious Artifact? *Journal of Consciousness Studies, 10*, 47–66.
Franklin, S. (2005). Evolutionary Pressures and a Stable World for Animals and Robots: A Commentary on Merker. *Consciousness and Cognition, 14*, 115–118.
Franklin, S., & Baars, B. (2010). Two Varieties of Unconscious Processes. In E. Perry, D. Collerton, H. Ashton & F. LeBeau (Eds.), *New Horizons in the Neuuroscience of Consciousness* (pp. 91–102). Amsterdam: John Benjamin.
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., & Graesser, A. C. (1997). Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents *Intelligent Agents III* (pp. 21–35). Berlin: Springer Verlag.
Franklin, S., Kelemen, A., & McCauley, L. (1998). IDA: A Cognitive Agent Architecture *IEEE Conf on Systems, Man and Cybernetics* (pp. 2646–2651): IEEE Press.
Franklin, S., & Patterson, F. G. J. (2006). The LIDA Architecture: Adding New Modes of Learning to an Intelligent, Autonomous, Software Agent *IDPT-2006 Proceedings (Integrated Design and Process Technology)*: Society for Design and Process Science.
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.
Freeman, Walter J. (1999). *How Brains Make Up Their Minds*. London: Weidenfeld & Nicolson General.
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.
Glazebrook, J. F., & Wallace, R. (2009). Small worlds and Red Queens in the Global Workspace: An information-theoretic approach. *Cognitive Systems Research, 10*(4), 333-365. doi: 10.1016/j.cogsys.2009.01.002
Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences, 20*, 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*, 379–401
Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the Structural Core of Human Cerebral Cortex. *PLoS Biol, 6*(7), e159. doi: 10.1371/journal.pbio.0060159
Harter, D., Graesser, Arthur C., & Franklin, S. (2001). Bridging the gap: Dynamics as a unified view of cognition. *Behavioral and Brain Sciences, 24*, 45–46.
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.
Holz, E. M., Glennon, M., Prendergast, K., & Sauseng, P. (2010). Theta-gamma phase synchronization during memory matching in visual working memory. *Neuroimage, 52*(1), 326-335.
Jackson, J. V. (1987). Idea for a Mind. *Siggart, Newsletter, 181*, 23–26.
Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences, 11*(7), 267-269.
Kelso, J. A. S. (1995). *Dynamic Patterns: The Self Organization of Brain and Behavior*. Cambridge, MA: MIT Press.

Lewis, M. D. (2005). Bridging emotion theory and neurobiology through dynamic systems modeling *BEHAVIORAL AND BRAIN SCIENCES, 28*, 169–245.
Madl, T., Baars, B. J., & Franklin, S. (2011). The Timing of the Cognitive Cycle. *PLoS ONE, 6*(4), e14803.
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.
Massimini, M., Ferrarelli, F., Huber, R., Esser, Steve K., Singh, H., & Tononi, G. (2005). Breakdown of Cortical Effective Connectivity During Sleep. *Science, 309*, 2228–2232.
Merker, B. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition, 14*, 89–114.
Minsky, M. (1985). *The Society of Mind*. New York: Simon and Schuster.
Newell, A. (1990). *Unified Theories of Cognition*. Cambridge MA: Harvard University Press.
Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience & Biobehavioral Reviews, 34*(7), 1023-1035.
Ornstein, R. (1986). *Multimind*. Boston: Houghton Mifflin.
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.
Samsonovich, A. V. (2010a). Toward a Unified Catalog of Implemented Cognitive Architectures. *Proceeding of the 2010 Conference on Biologically Inspired Cognitive Architectures*, 195-244.
Samsonovich, A. V. (2010b). 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.
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.
Schutter, D. J. L. G., & van Honk, J. (2004). Extending the global workspace theory to emotion: Phenomenality without access. *Consciousness and Cognition, 13*(3), 539-549. doi: 10.1016/j.concog.2004.05.002
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,Äì6), 374-384. doi: 10.1016/j.jphysparis.2005.09.006
Seth, A. (2007). Models of consciousness. *Scholarpedia, 2*(1), 1328.
Shanahan, M. (2010). *Embodiment and the Inner Life*. Oxford: Oxford University Press.
Shanahan, M. P., & Baars, B. J. (2005). Applying Global Workspace Theory to the Frame Problem. *Cognition, 98*(2), 157-176.
Sigman, M., & Dehaene, S. (2006). Dynamics of the Central Bottleneck: Dual-Task and Task Uncertainty. *PLoS Biol., 4*.
Skarda, C., & Freeman, Walter J. (1987). How Brains Make Chaos in Order to Make Sense of the World. *Behavioral and Brain Sciences, 10*, 161–195.
Snaider, J., McCall, R., & Franklin, S. (2011, August 3-6, 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.
Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Frontiers in bioscience, 14*, 321-332.
Tort, A. B. L., Komorowski, R. W., Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009a). Theta–gamma coupling increases during the learning of item–context associations. *PNAS, 106*(49), 20942-20947

Tort, A. B. L., Komorowski, R. W., Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009b). Theta–gamma coupling increases during the learning of item–context associations. *Proceedings of the National Academy of Sciences, 106*(49), 20942-20947.
Uchida, N., Kepecs, A., & Mainen, Zachary F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nature Reviews Neuroscience, 7*, 485–491.
Van Bockstaele, B., Verschuere, B., De Houwer, J., & Crombez, G. (2010). On the costs and benefits of directing attention towards or away from threat-related stimuli: A classical conditioning experiment. *Behaviour Research and Therapy, 48*, 692-697.
van den Heuvel, M. P., & Sporns, O. (2011). Rich-Club Organization of the Human Connectome. *Journal of Neuroscience, 31*(44), 15775-15786. doi: 10.1523/JNEUROSCI.3539-11.2011
Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.
Wallace, R. (2007). Culture and inattentional blindness: a global workspace perspective. *J Theor Biol, 245*(2), 378-390. doi: S0022-5193(06)00473-5 [pii] 10.1016/j.jtbi.2006.10.006
Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., & Thelen, E. (2001). Autonomous mental development by robots and animals. *Science, 291*, 599-600.
Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science, 17*, 592–599.