Elsevier Editorial System(tm) for
Biologically Inspired Cognitive Architectures
Manuscript Draft

Manuscript Number:

Title: A LIDA Cognitive Model Tutorial

Article Type: Invited Article

Keywords: cognitive model; attention; memory; emotion; action selection; decision making

Corresponding Author: Prof. Stan Franklin, Ph.D,

Corresponding Author's Institution: University of Memphis

First Author: Stan Franklin, Ph.D,

Order of Authors: Stan Franklin, Ph.D,; Tamas Madl; Steve Strain; Usef Faghihi; Daqi Dong; Sean Kugele; Javier Snaider; Pulin Agrawal; Sheng Chen

Abstract: Over a decade in the making and described in some seventy-five published papers, the LIDA cognitive model is comprehensive, complex, and hard to "wrap one's head around". Here we offer, in tutorial fashion, a current, relatively complete and somewhat detailed, description of the conceptual LIDA model, with pointers to more complete accounts of individual processes in the literature. These descriptions also include some features of the workings of the LIDA model that have not been published previously.

The tutorial begins with several short sections designed to ease the reader into the LIDA model. These are followed by an account of the conceptual commitments of the LIDA model. We also include a brief introduction to the LIDA computational model via the LIDA Framework, with pointers to its own tutorial. This is followed by sketches of several of the LIDA based agents developed with the help of the Framework. The tutorial ends with a section on current research activity, which includes a table showing which aspects of the LIDA conceptual model have currently been implemented computationally.

Suggested Reviewers: John Laird
laird@umich.edu

Ben Goertzel
ben@goertzel.org

Opposed Reviewers: Pat Langley

A LIDA Cognitive Model Tutorial

Stan Franklin, Tamas Madl, Steve Strain, Usef Faghihi, Daqi Dong, Sean Kugele, Javier Snaider, Pulin Agrawal, Sheng Chen

Abstract

Over a decade in the making and described in some seventy-five published papers, the LIDA cognitive model is comprehensive, complex, and hard to "wrap one's head around". Here we offer, in tutorial fashion, a current, relatively complete and somewhat detailed, description of the conceptual LIDA model, with pointers to more complete accounts of individual processes in the literature. These descriptions also include some features of the workings of the LIDA model that have not been published previously.

The tutorial begins with several short sections designed to ease the reader into the LIDA model. These are followed by an account of the conceptual commitments of the LIDA model. We also include a brief introduction to the LIDA computational model via the LIDA Framework, with pointers to its own tutorial. This is followed by sketches of several of the LIDA based agents developed with the help of the Framework. The tutorial ends with a section on current research activity, which includes a table showing which aspects of the LIDA conceptual model have currently been implemented computationally.

1 Introduction

Cognitive models come in several varieties, conceptual, mathematical, and computational. Their function is to explain cognitive representations and processes, and to predict their outcomes. They spawn hypotheses that serve to guide experimentation. Most cognitive models attempt to model some single type of cognitive process, say perception, attention, memory, emotion, decision making, action selection, etc., or some narrow range within one of these. The much rarer systems-level model (cognitive architecture) attempts the full range of activities from incoming stimuli to outgoing actions, together with the full range cognitive processes in between.

LIDA is a systems-level cognitive model. It is conceptual and partly computational. It attempts to model minds, be they human, animal or artificial (Franklin, 1995, p. 412), which we take to be control structures for autonomous agents (Franklin & Graesser, 1997). We think of minds as being implemented as virtual machines running on top of underlying devices such as brains or computers

(Sloman & Chrisley, 2003). In addition to providing explanations and producing hypotheses, we aspire that LIDA act as a cognitive prosthesis to aid in thinking about, and understanding, individual cognitive activities and their processes. It should do so by providing a useful cognitive ontology (Franklin & Ferkin, 2006).

After providing a synopsis of the LIDA model (Section 2) and a brief account of its cognitive cycle (Section 3), we quickly summarize its conceptual commitments (Section 4). LIDA's modules and their interactions are then described in some detail (Section 5). These descriptions include details of the workings of the LIDA model that have not been published previously. We continue by describing several LIDA based software agents (Section 6) and the computational framework on which they are based (Section 7). We conclude with a discussion of current and future work. (Section 8).

2 A brief synopsis of the LIDA cognitive model

2.1 LIDA's definition of mind

"An autonomous agent is a system situated in and part of an environment, which senses that environment and acts on it over time in accordance with its own agenda, so as [it may affect] what it senses in the future." (Franklin & Graesser, 1997)

As a cognitive model, LIDA seeks to describe mental phenomena in terms of concepts with explanatory and predictive power. At the heart of the LIDA model is a technical definition of *mind as a control structure for an autonomous agent*. The primary function of an autonomous agent is to continually and iteratively answer the question, "What do I do next?"

Such an agent may be biological or artificial; when we speak of minds as biological or artificial, we will do so exclusively in terms of these technical definitions of autonomous agent and of mind. Many of the concepts found in LIDA's particular ontology of cognitive processes, found throughout this paper, may be usefully traced back to these definitions.

2.2 LIDA's cognitive cycle

Every animal must frequently sample its environment, external or internal, and act appropriately in response. The LIDA model's *cognitive cycle*, taken from the action-perception cycle of the psychologists and neuroscientists (Cutsuridis, Hussain, & Taylor, 2011; Dijkstra, Schöner, & Gielen, 1994; Freeman, 2002; Fuster, 2004; Fuster, 2002; Neisser, 1976), enables just such frequent (~10 hz) sampling and responding (Madl, Baars, & Franklin, 2011). One can think of the cognitive cycle as a cognitive atom of which higher-level cognitive processes, deliberation, reasoning, problem solving, planning, imagining, etc., are comprised. Each cognitive cycle can be divided into three phases, a *perception and understanding phase*, an *attention phase*, and an *action and learning phase*. (See Figure **1**) Using incoming sensory data, memories, etc., the first phase updates its understanding of the current situation. The attention phase then filters the content of this understanding for saliency, and broadcasts this conscious content globally in accordance with *Global Workspace Theory* (GWT)(Baars, 1988). Although cognitive cycles may overlap, partially operating in parallel, conscious broadcasts occur in sequence. The third phase selects and executes an appropriate response, and also learns into a bevy of memory systems.

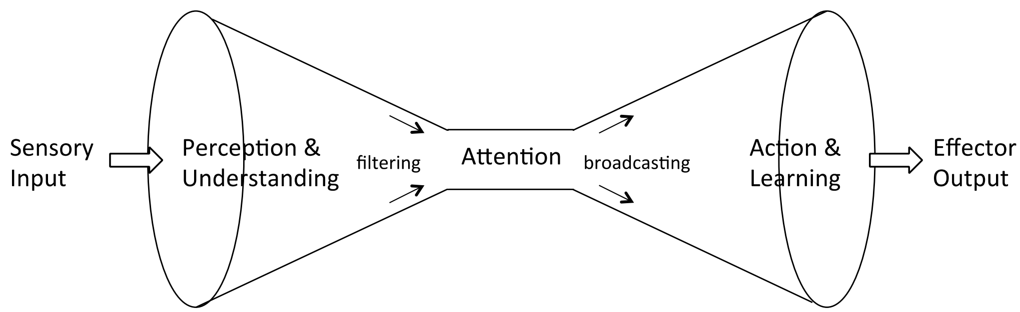

Figure 1. The LIDA cognitive cycle phase diagram

3 A quick trip through LIDA's cognitive cycle

Though an individual cognitive cycle is very brief in humans, ~200-500 ms (Madl, et al., 2011), it is quite complex, consisting of more than a dozen interacting modules (See Figure 2). Though we describe the LIDA model, and its cognitive cycle, in terms of modules, we make no commitment to a modular structure of the underlying brain. Keep in mind that LIDA models minds, not brains (see Section 2.1). Though the modules are represented with sharp boundaries in the figure, they actually interact considerably, pointing back from one to another to access needed data elements. Also note that the LIDA model, unlike procedural computer programs, does not execute its computations serially. Its processes, excepting only consciousness and action selection (see Section 4.9), are completely asynchronous. Its various memory systems range from quite short term to very long term. Its processes fall into one of three categories: *never conscious*, *pre-conscious* (possibly to come to consciousness), or *conscious* (Franklin & Baars, 2010).

The LIDA cognitive cycle begins with sensory stimuli, both external and internal, coming to *Sensory Memory* where it is represented, and engages early feature detectors. The resulting content involves both the Current Situational Model, and Perceptual Associative Memory. The latter serves as recognition memory, producing a percept that is made available to the Current Situational Model. Using both the percept and the incoming content, together with remaining content which has not yet decayed away, the Current Situational Model continually updates itself by cueing *Perceptual Associative Memory, Spatial Memory, Transient Episodic Memory* and *Declarative Memory,* and using the returning local associations. Further updating is produced in the *Workspace*[1] by *Structure Building Codelets*[2] (Hofstadter & Mitchell, 1995) that build preconscious thoughts (Franklin & Baars, 2010) using material from the Current Situational Model and the *Conscious Contents Queue*. All of this comprises the perception and understanding phase of the model.

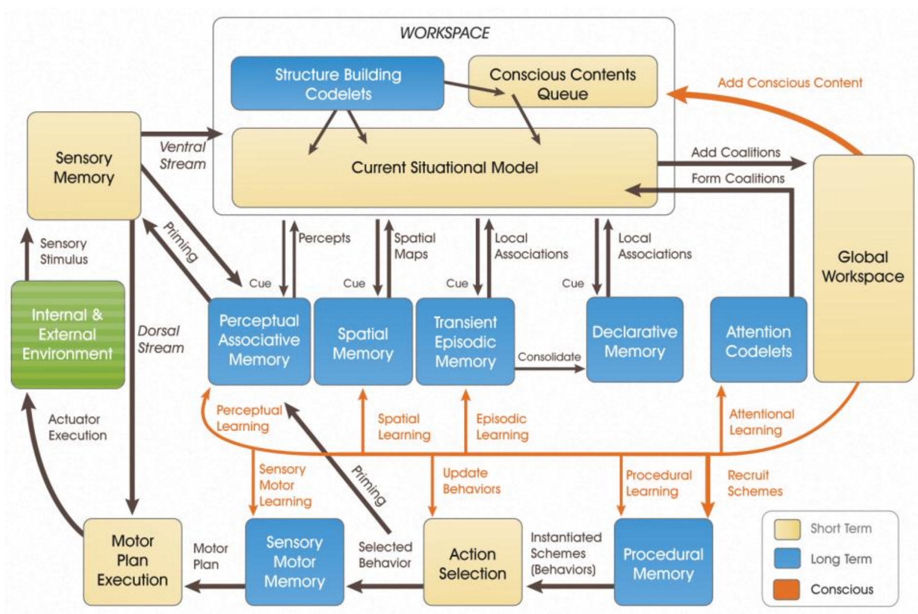

Figure 2. The LIDA cognitive cycle

---

[1] This preconscious analog of working memory is not to be confused with the Global Workspace (described below), a consciousness mechanism based on Baars' theory (1988).
[2] A codelet is a small piece of code that keeps watch waiting for conditions to be ripe for it to act in pursuit of its one specific task.

In the service of the attention phase of the cognitive cycle, each *attention codelet* continually surveys the Current Situational Model on the lookout for content that it would like to bring to consciousness. Upon finding such, it creates a *coalition*, which in the Global Workspace (Baars, 1988) engages in a competition for consciousness. The winning, the most salient, coalition has its content *broadcast* globally, whereby it becomes conscious in the functional sense[3] (Franklin, 2003) (See Figure **2**), completing the attention phase of the cognitive cycle.

The third, the action and learning, phase of the LIDA cognitive cycle allows almost every LIDA module to select that part of the conscious contents of the cycle that is appropriate for it to learn, that is, fitting for its underlying data structure. Procedural Memory, the memory of what to do when, uses conscious contents to instantiate behaviors that might be suitable as responses to the incoming stimuli. The Action Selection module chooses one such behavior that is then submitted to Sensory Motor Memory for the creation or selection of an appropriate motor plan, which can then be executed. This completes the LIDA cognitive cycle.

4 Conceptual commitments of the LIDA model

Before discussing the individual LIDA modules in more detail, we will briefly describe the various conceptual commitments to which we attempt to adhere while designing the LIDA model (Franklin, Strain, McCall, & Baars, 2013).

4.1 Systems-level cognitive modeling

Cognitive scientists use conceptual, mathematical and computational models to explain and predict cognitive phenomena. For the most part, and for utilitarian reasons, these models are limited to some restricted function of cognition, such as memory, perception, attention, action selection, or some subset of one of these. Though these limited models have proved exceedingly useful, the problem of discovering the relationships between them is often a difficult one. As a result, researchers from various disciplines such as social psychology (Lewin, 1951), artificial intelligence (Newell, 1973), memory research (Hintzman, 2011), cognitive modeling (Langley, Laird, & Rogers, 2009), and neuroscience (Bullock, 1993) have argued for the necessity of systems-level cognitive modeling.

LIDA is a systems-level model attempting to account for the perception of incoming stimuli, all of the concurrent and resulting internal processing, culminating with the selection and execution of an appropriate action. For its relationship to other, related disciplines, please see Section 4.12 below and Figure 3.

[3] The LIDA model makes no claims regarding phenomenal consciousness.

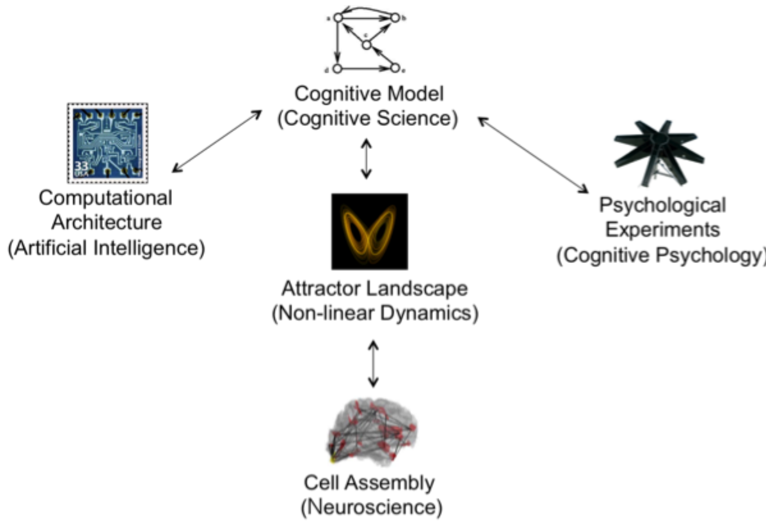

Figure 3. The LIDA model's family tree (reprinted from (Franklin, Strain, Snaider, McCall, & Faghihi, 2012))

4.2 Embodied (Situated) Cognition

Embodied cognition asserts that bodies as well as brains, the body mind relationship, affects all cognitive processing (de Vega, Glenberg, & Graesser, 2008). Situated cognition argues for the influence of the environment on cognitive processing. The LIDA model conforms to both these strictures by employing only perceptual symbols (Barsalou, 1999), and completely avoiding the use of amodal symbols. The labels that appear in our descriptions of the *conceptual LIDA model*[4], and the diagrams thereof, are strictly for the use of the reader, and play no causative role in the model itself.

Embodied and situated cognition closely intersect with the so-called *enactive model of cognition* (Varela, Thompson, & Rosch, 1991), a brainchild of the phenomenology of Husserl, Heidegger, and Merleau-Ponty. All three models are closely related to *dynamical systems theory* (Thelen & Smith, 1994; Van Gelder, 1998), *Freeman's neurodynamics* (Dreyfus, 2009; Freeman, 1999) and *interactionism* (Clark, 1999; Dewey, 1896; Gallagher, 2009; Oyama, 2000; Von Uexküll & Mackinnon, 1926). Important common features are the continual and mutual interaction between agent and environment, the active rather than passive role of the agent's internal processes, and the lack of actual separation between perception and action. As will be discussed in more detail at specific points below (see Sections 4.9, 4.12, and 4.13), we feel that the LIDA model is resonant with the core ideas of the embodied, situated, and enactive views (Franklin, Madl, D'Mello, & Snaider, 2014; Franklin, et al., 2012).

[4] We subdivide the LIDA model into conceptual and computational sub-models. The discussion in Section 3 primarily relates to the conceptual model; that in Section 7 primarily to the computational model, and Section 5 to both.

4.3 Cognitive Cycles as Cognitive Atoms

Using salient information from the contents of the conscious broadcast, together with never conscious processing, each cognitive cycle selects and executes an appropriate response. We refer to this single cycle process as consciously mediated action selection. Higher-level action selection (decision making), such as making breakfast, requiring a sequence of actions, can be implemented by multiple cognitive cycles. Some such decision making is deliberative, employing partially conscious processing. Other such higher-order partly conscious cognitive processing is implemented in the LIDA model by sequences of cognitive cycles. These include planning, imagining, reasoning, day dreaming, volitional memory retrieval, etc.

4.4 Global Workspace Theory

Psychologists and neuroscientists have given various definitions of attention, and have ascribed different functions to it. Posner (Posner & Fan, 2004) suggests three separate functions of attention with distinct underlying brain networks. These functions include 1) *alerting*: "maintaining an alert state"; 2) *orienting*: "focusing our senses on the information we want" (e.g., your focus on reading this document); and 3) *executive attention*: "the ability to manage attention towards goals and planning". In each of the three attention functions suggested by Posner, there must be an attentional mechanism to select and bring the most urgent, salient information to the consciousness. The *selective* part of LIDA's attentional mechanism is very briefly described in the following.

The attention phase of the LIDA cognitive cycle is taken directly from *Global Workspace Theory* (GWT) (Baars, 1988, 2002), where attention is defined as the process of bringing content to consciousness. That definition is adopted for the LIDA model. Hypothesizing a parallel distributed nervous system composed of a bevy of specialized processes, GWT has coalitions of these processors competing for consciousness, with the contents of the winning coalition broadcast globally. These most salient conscious contents, collectively referred to as the *conscious* or *global broadcast*, are used for learning and for action selection. Our LIDA cognitive model can be viewed as a specification and fleshing out of GWT (Franklin, et al., 2013; Franklin, et al., 2012), along with a number of other psychological and neuropsychological theories (Baddeley & Hitch, 1974; Barsalou, 1999; Conway, 2001; Ericsson & Kintsch, 1995).

4.5 Learning via Consciousness

Taken from GWT, the LIDA model supports the *Conscious Learning Hypothesis*: significant learning takes place via the interaction of consciousness with the various memory systems (Baars & Franklin, 2003; Franklin, Baars, Ramamurthy, & Ventura, 2005). Following each conscious broadcast, every memory module in LIDA updates itself incorporating appropriate material from the conscious broadcast. Thus consciousness is both necessary and sufficient for significant learning in unimpaired humans. Substantiated claims for subliminal learning so far have turned out to be due to unconscious priming (Boltea & Goschke, 2008; Eimer & Schlagecken, 2003) which is too limited in both scope and duration to be considered significant learning. Also note that in all cases of implicit learning (Cleeremans, Destrebecqz, &

Boyer, 1998; Jimenez, 2003) and latent learning (Campanella & Rovee - Collier, 2005; Chamizo & Mackintosh, 1989; Franks et al., 2007) learning subjects must be conscious when learning takes place.

4.6 Comprehensive Decay of Representations and Memory

Each LIDA module is composed of processes operating on structured representations of internal or external entities. The fundamental data type of these representations is the digraph, consisting of nodes and links[5]. More complex structures are built from these. Each represented entity has one or more numerical variables attached to it, for example a *base-level activation* measuring its past usefulness, or a *current activation* tracking its relevance to the current situation. All of these numerical variables decay, with many of their various decay rate functions sigmoid. An entity decays away (is removed from the system–forgotten[6]) when its appropriate variable, for example its base-level activation, falls below a threshold. On the other hand, because of sigmoidal decay rate functions, some entities decay so slowly that they never seem to decay away.

LIDA's conceptual commitment to decay accords with one of the four general requirements for a self-organizing system: Such a system must be dissipative (see Section 4.12, which describes LIDA's conceptual commitment to cognition as a self-organizing dynamical system). Decaying away is also necessary to make profligacy in learning computationally tractable.

4.7 Profligacy in Learning

As described in Section 4.5, learning in LIDA takes place in every memory system (see the red arrows in Figure 2) with each conscious broadcast. As we have seen in Section 3, such broadcasts occur with each cognitive cycle, that is, very frequently (at ~10hz in humans (Madl, et al., 2011)). Thus learning in LIDA is profligate, happening in every possible system at every possible opportunity. LIDA learns in both an *instructionist* manner in which new entities are represented, and in a *selectionist* manner in which the base-level activation (or other appropriate variable) is reinforced. New entities are generated whenever possible, and are reinforced (tested) whenever they come to consciousness again. Such entities remain in the system so long as their reinforcement outstrips their decay. Hence learning in LIDA is a generate and test algorithm (Kaelbling, 1994). One can also view learning in LIDA as Darwinian, with a new population being generated with each cognitive cycle, and its fitness tested in the same cycle. (In Section 5.3 we'll see

---

[5] The newly initiated Vector LIDA project (Snaider & Franklin, 2014b) will have high dimensional vectors as the fundamental data type. It will still use nodes and links, which will be represented as vectors (Snaider & Franklin, 2014a).

[6] Representations relevant in the current situation can decay away within tens of seconds, and will be removed from the Workspace, but can still persist in long-term memory (declarative, transient episodic, or spatial). On a longer time scale, it can decay away altogether, being removed from all memory systems.

another way in which the LIDA model is Darwinian.) Note also that long-term memories should more accurately be called potential long-term memories, since many may decay away quite quickly.

4.8 Feelings are Motivators and Modulators of Learning

In humans feelings include appetitive drives such as thirst and hunger, temperature preferences such as too hot or too cold, various sensing of pain, feeling tired, feeling depressed, etc. LIDA models such feelings as motivators and as modulators of learning (Franklin & Ramamurthy, 2006). As motivators, feelings enable action selection that is both sophisticated and flexible. LIDA treats emotions as feelings with cognitive content (de Spinoza, 1883; Johnston, 1999; Panksepp, 2005). These include anger, joy, sadness, fear, guilt, regret, envy, shame, resentment etc.

Representations of emotions in LIDA, and their association of an emotion with aspects of the current situation, are consistent with appraisal theory (Scherer, 2001) – see Figure 4. Briefly, a specific type of Structure Building Codelet (see Section 5.3.2), called an Appraisal Codelet, can propose and link an emotion PAM node to an existing node structure, based on its relevance, implications, the agent's coping potential, among other factors. The connection to appraisal theory is explained in more detail in (Franklin, et al., 2014).

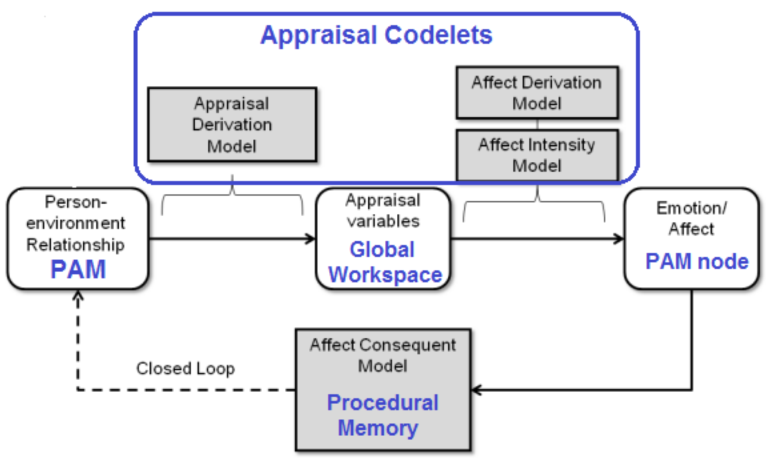

Figure 4. Connection between components of LIDA (illustrated in blue / bold font) and of affordance theory, based on (Marsella, Gratch, & others, 2010) (illustrated in black / light font). Reprinted from (Franklin, et al., 2014).

As motivators for action selection feelings, including emotions, allow rapid evaluation of situations, including whether one is helpful or harmful with respect to the agent's goals (see Sections 5.7 and 6). As modulators of learning, feelings (affect) is a major determiner of learning rate, producing an inverted U effect (Yerkes & Dodson, 1908). Feelings are represented as nodes in Perceptual Associative Memory (see Figure 2), and occur and play a central role in the determination of activation values throughout the model.

4.9 Asynchrony[7]

The cascading cognitive cycles are serial in regard to the conscious broadcast as required for the seriality and coherence of consciousness. Recently Baars has proposed a *dynamic* Global Workspace (Baars, Franklin, & Ramsøy, 2013) under which a winning coalition emerges (ignites) from some place in the cortico-thalamic core, producing a global broadcast of conscious contents via a ~100ms broadcast to receiving neural networks widely distributed over the brain (Gaillard et al., 2009). The other serial process in the LIDA cognitive cycle is the selection of a single behavior by the Action Selection module (see Figure 2) for execution. All other processes of modules in the LIDA model respond to their local conditions in a completely independent and asynchronous manner. Thus multiple processes run simultaneously, making the model highly parallel. In fact, it can be thought of as a multi-agent system (Doran, Norman, Franklin, & Jennings, 1997; McCauley & Franklin, 2002; Watson, Mills, & Buckley, 2011). Conceptually, there is no system clock, and rather than being implemented in the architecture, LIDA's overlapping cognitive cycles emerge from the asynchronous operation of multiple independent processes acting in parallel in response to local conditions.

Asynchrony in the LIDA model accommodates the possibility of algorithmic behavior more complex than that of a data pipeline in the information processing paradigm. Such a pipeline is closed along its length from input to output, and thus will always produce the same output for a given input; its flow is sequentially dependent, meaning that distal processes must wait for proximal processes to finish before they can begin; it is inactive in the absence of input; the activity of its internal processes cannot alter its shape or points of connection at either of or between its ends; and the relationship between input and output can be updated only at a periods commensurate with the time required for information to flow through the entire pipeline . These constraints, while available if desired, can be removed in the LIDA model, particularly in the Workspace, the content of which may be modified by the various memories and by processes known as structure building codelets (See Section 5). Thus, LIDA's asynchrony allows for the possibility of process features emphasized in embodied cognition (Section 4.2) and self-organizing dynamics (Section 4.12).

4.10 Transient Episodic Memory

We humans are often confronted by, and must remember, events that are repetitive with many significant features remaining almost constant, for example, where we park a car in a parking garage on a daily basis. The major features of the parking

---

[7] There are numerous technical senses of the term "asynchrony" (e.g. see https://en.wikipedia.org/wiki/Asynchrony). We use the term in the sense of asynchronous input-output in computer science; in other words, asynchronous processes are not in general required to wait on input from other processes to continue their own operations. In particular, this use is distinct from that of neuroscience, where it refers to a lack of temporal correlation between neural activity patterns.

garage will reinforce themselves each day, while the individual parking spot will interfere from one day to the next. Thus long-term episodic memory cannot be expected to handle such a situation. For this reason, the LIDA model postulates a *Transient Episodic Memory* (see Figure 2) whose traces decay within a few hours or a day (Conway, 2001; Franklin, et al., 2005). Thus we can often remember what we had for lunch yesterday, but not on the same day of the week two weeks ago. Though Transient Episodic Memory has been mostly ignored by memory researchers, we think it quite necessary for human episodic memory functioning.

4.11 Consolidation

The LIDA model also postulates that all episodic memories are learned into Transient Episodic Memory, and that those that have not yet decayed away are consolidated into Declarative Memory (see Figure 2) at some offline time (Franklin, et al., 2005). There is much evidence for such consolidation (Born & Wagner, 2006; Daoyun & Wilson, 2006; Haist, Gore, & Mao, 2001; McGaugh, 2000; Nadel, Hupbach, Gomez, & Newman-Smith, 2012; Remondes & Schuman, 2004; Stickgold & Walker, 2005; Walker, Brakefield, Hobson, & Stickgold, 2003).

4.12 Nonlinear Dynamics Bridge to Neuroscience

As mentioned several times already, LIDA is intended to model minds, not brains (see Section 1 and Figure 3). However it is critical that any systems-level cognitive model such as LIDA (see Section 4.1) be consistent with known neuroscientific evidence, so as to account for the relationship between minds and brains, since biological brains are the only known examples of sophisticated minds. We concur with Fuster that the gap between LIDA's cognitive representations and the underlying neurodynamics can be bridged by non-linear dynamics that exhibit self-organization. We support Fuster's proposal that cognitive entities are represented neurally by cognits (2006). The activity of the brains perceptual oscillators is integrated with that of its higher-order neural oscillators (Barham, 1996; Freeman, 2003) allowing the application of various memory systems, of deliberation, and of goals to the current state of the brain and its environment. The globally broadcast subset of such integrated oscillatory activity (Baars, et al., 2013) enables action selection and the several forms of learning, thus activating oscillators that effect action execution (see Figure 2). In addition, the phase-coupling of oscillators effect timing relationships that are characteristic of the neurophysiological structure of cognition (See Section 4.13; and Strain, Franklin, Heck, & Baars, in preparation).

A key feature of non-linear systems is their resistance to reductionist approaches. (Strogatz, 2014) How then can a model that reduces cognitive processes into small codelets, such as LIDA, capture the essential behavior of a self-organizing system? No currently implemented LIDA agent (see Section 7) exhibits self-organization of its processes, although several that might are currently under development (Section 9). Thus we find it necessary to justify our model by explaining how it can support such complex dynamics and the attendant dynamical phenomena. We claim that LIDA can, in principle, accommodate such dynamics according to criteria enumerated by Kelso (1995); in other words, a sufficiently sophisticated LIDA agent would self-organize its cognitive processes in the way described by van Gelder (1998). In Kelso's view (1995), a system with the following properties is likely to exhibit spontaneous self-organization:

1. A large number of components interacting weakly and nonlinearly;
2. Dissipative, far-from-equilibrium thermodynamics;
3. Reciprocal influence and coordination between patterns of activity and the components that form the patterns; and
4. At least three mutually interacting levels, namely a component level, a collective level at which patterns (aka attractors) may emerge, and a context or task level that acts as a boundary condition on the other levels (Kelso 1995).

Regarding 1), a LIDA agent fully implemented with the Model's cognitive functions will have a great number of mutually interacting codelets operating in its Workspace (Section 5.3); 2), The Workspace will be a system with a continuous influx of informational energy (and thus far-from-equilibrium), open system, reciprocally interacting with various memory modules (Section 5.2), with activity that will decay over time unless cognitively reinforced in some way (thus dissipative; see also Section 4.6); 3), It will also feed, via the action of attention codelets (Section 5.4), the Global Workspace (GW) (Section 5.5), the broadcast of which will modulate (and potentially, through learning mechanisms, modify) all preconscious modules;[8] and 4), the Workspace possesses processes that operate on three scales (in order of increasing timescale): the codelet timescale, the GW broadcast timescale, and the timescale defined by the agent's currently active goal-related, task-related, and environmental constraints (the agent's "sense of time").

A more concrete connection of LIDA's processes to non-linear dynamics, based on Dynamic Field Theory (Erlhagen & Schoner, 2002; Schöner, 2008), has been outlined in (Franklin, et al., 2014). Briefly, representations in each of the modules in LIDA's cognitive cycle can be implemented using neural populations which represent dimensions characterizing their features, and which are governed both by input activations and the activations of neighboring neurons. While beyond the scope of the LIDA Model's purpose to model of mind as a control system for an autonomous agent, this would allow a mathematical formulation of the dynamics of these representations, as well as making a connection to empirical neuroscience (Franklin, et al., 2014).

4.13 Theta Gamma Coupling and the Cognitive Cycle

LIDA models minds rather than brains. Why then does LIDA care about brains? In brief, LIDA shares with certain other theories the view that brain and mind are different aspects of the same dynamical system. As a model of mind, what does LIDA have to say about brains? In summary, LIDA's requirement for brains follows: The dynamical organization of brain activity must align with the temporal structure of the corresponding cognitive processes. Neural dynamic patterns at multiple temporal levels have been shown to have cognitive significance, and so the processes of LIDA must have a parallel temporal structure. In previous work we

---

[8] The GW broadcast will also modulate the executive modules (Section 5.6-5.7), but this effect of the broadcast is not immediately pertinent to the present discussion.

have shown how LIDA's macroscopic structure, the cognitive cycle, relates to neural activity on the scale of 100s of milliseconds (Madl, et al., 2011), Below we will explain and discuss *theta-gamma coupling*, a neural phenomenon that correlates well with LIDA's cognitive processes at the mesoscopic[9] scale of 10s of milliseconds.

LIDA subscribes to the embodied view of cognition (see Section 4.2), which views mind and brain as different aspects of the same whole. LIDA's most direct connection to neural theory is via dynamic Global Workspace Theory (dGWT) (Baars, et al., 2013) (see Section 4.9). dGWT constructs out of recent neuroscientific evidence a general specification for the neural implementation of a cyclically recurring global broadcast. Although both dGWT and LIDA address the relationship between mind and brain, dGWT neurally grounds a psychological theory (GWT; see Section 4.4), while LIDA seeks a general theory of cognition based on GWT (and consequently, on dGWT as well).

On this view, a brain rhythm phenomenon known as *theta-gamma coupling* offers an interpretation that elaborates the connection between Freeman's neurodynamics and LIDA's cognitive cycle. Theta-gamma coupling is a type of *cross-frequency coupling* (CFC), a measurable brain state in which activity with a neural signature in the low frequency range becomes correlated with high frequency activity. In particular, *phase-amplitude coupling* refers to a CFC structure in which an amplitude burst of fast frequency activity occurs at a particular phase of a slow wave.

An illustrative metaphor is eating a meal at a certain time of day. The solar cycle can be said to be the slow wave and the time of day its phase, with the behavior modeled as a fast frequency wave that peaks during the meal and goes to zero in between. The wave representing the eating activity can then be said to be phase-amplitude coupled to the solar cycle.

CFC is measured by spectral analysis of raw EEG signals (eg Voytek, D'Esposito, Crone, & Knight, 2013). CFC, especially the subtype known as theta-gamma coupling[10], empirically differentiates task successes from non-successes within a broad range of cognitive functions, including declarative memory, working memory, attention, perceptual organization, spatial memory, and perceptual organization

---

[9] These terms of scale and their meanings are from Freeman's neurodynamical theory (2003). Note that in brains they connote a typical spatial scale as well as a temporal one; however, LIDA makes no claims regarding the spatial organization of cognitively relevant brain activity, since cognitive processes in the abstract are organized independently of physical space. Thus we limit our concern to the temporal structure of neural activity.

[10] Due to lack of clear standards for identifying various frequency bands (see Steriade, 2006 for a review), we choose to adopt a usage of the very common "theta-gamma coupling" as being more or less equivalent—for our purposes—to the more general "cross-frequency coupling." In other words, with said term we refer not to specific frequency bands (which are defined differently in different decades, cortical regions, species, and labs) but to the temporal association between a slow wave and a fast wave, which we believe to be the electrical signature of the kind of cognitive processing hypothesized by Freemanian neurodynamics, dGWT, and LIDA.

(Canolty & Knight, 2010; Doesburg, Green, McDonald, & Ward, 2009; Osipova et al., 2006; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010).

A cognitive cycle in LIDA is hypothesized to last on the order of 200-500 ms (Madl et al. 2011). However, in keeping the asynchrony discussed in Section 4.9 above, a cognitive cycle starts before the previous one finishes. In fact, the cycle is not a true cycle in the classical sense of mathematics or physics; rather, it is a recurrent pattern that is roughly defined by the average time that would be necessary for a nervous impulse to traverse a path through cortex from receptor to effector. This approximate length for a cognitive cycle correlates with the period of a typical theta (or delta or even alpha) wave. Dynamic GWT proposes that the corticothalamic system implements the global broadcast by means of distributed activity organized using theta-gamma coupling. Extending this hypothesis, we suggest that a broadcast's cognitive content is represented by synchronous gamma activity within a theta-gamma couplet.[11]

Similarly, since each broadcast originates as a coalition built by attention codelets (Section 3), we view these coalitions as theta-gamma couplets as well. Coalitions that fail to win the competition for consciousness in the Global Workspace will nonetheless continue their activity within a neural assembly and produce electrical activity that is not organized in synchrony with that of the broadcast. Thus, a theta-gamma couplet would represent a coalition containing a bundle of cognitive content (gamma activity) organized within an activity pattern (theta activity) commensurate with the bandwidth of the broadcast (roughly defined by the average length of a cognitive cycle). Winning the broadcast would give the coalition/couplet access to a "megaphone" that can be transmitted across the cortex according to Pascal Fries' communication-through-coherence mechanism (Bastos, Vezoli, & Fries, 2015; Fries, 2005; Landau & Fries, 2012)

5 LIDA's individual modules and their interactions

This section describes LIDA's modules and processes conceptually, both processes internal to a single module, and processes between modules. (Computational information about them will be found in Section 7 below.) Referring to Figure 2, each module is described in its own subsection, beginning with Sensory Memory in the upper left and proceeding in a roughly clockwise direction around the figure. Do keep in mind that LIDA is a massively parallel system with the processes of each

[11] The role of synchronization as a conceptual binding mechanism for gamma activity has been hypothesized by numerous neuroscientists (Buzsaki, 2006; Gray, König, Engel, & Singer, 1989; Holz, Glennon, Prendergast, & Sauseng, 2010; Jensen & Colgin, 2007; Osipova, et al., 2006; Tallon-Baudry, 2009). The work of many others has implicated the role of theta in organizing synchronized gamma activity (e.g. (Canolty et al., 2006; Clayton, Yeung, & Kadosh, 2015; Doesburg, et al., 2009; Doesburg, Green, McDonald, & Ward, 2012; Jensen & Colgin, 2007; Lisman & (2005), 2005; Lisman & Buzsaki, 2008; Lisman & Jensen, 2013; Nakatani, Raffone, & van Leeuwen, 2014; Voytek et al., 2015)).

module operating independently and asynchronously in response to current local conditions (see Section 4.9). Exceptions to this rule are the Global Workspace, where conscious broadcasts must occur serially, and Action Selection, where a single behavior must be chosen during each cognitive cycle. Since the LIDA model is relatively young, some modules are more developed than others, both conceptually and computationally. In particular, the modules at either end, Sensory Memory and Motor Plan Execution must depend heavily on the sensors and actuators of a particular agent, and so can be less fully described.

Many of the modules described in the subsections below are memory systems of one sort or another that store information from the past for potential use in the present. In the LIDA model memory systems are taken from those of humans (Anderson & Bower, 1973; Baddeley & Hitch, 1974; Broadbent, Squire, & Clark, 2004; Conway, 2001; Ericsson & Kintsch, 1995; Mayes & Roberts, 2002; Quillian, 1966; Rugg & Yonelinas, 2003; Schacter & Tulving, 1994; Tulving, 1983; Tulving & Markowitsch, 1998). One way of cutting up the memory pie, but by no means the only one, is illustrated in Figure 5. The diagram starts with the shortest term memory systems on the left, increasing to the longest term on the right. Otherwise what distinguishes one system from another is a difference in the structure of the information remembered, their typical data structure. These differences will be specified in the subsections below.

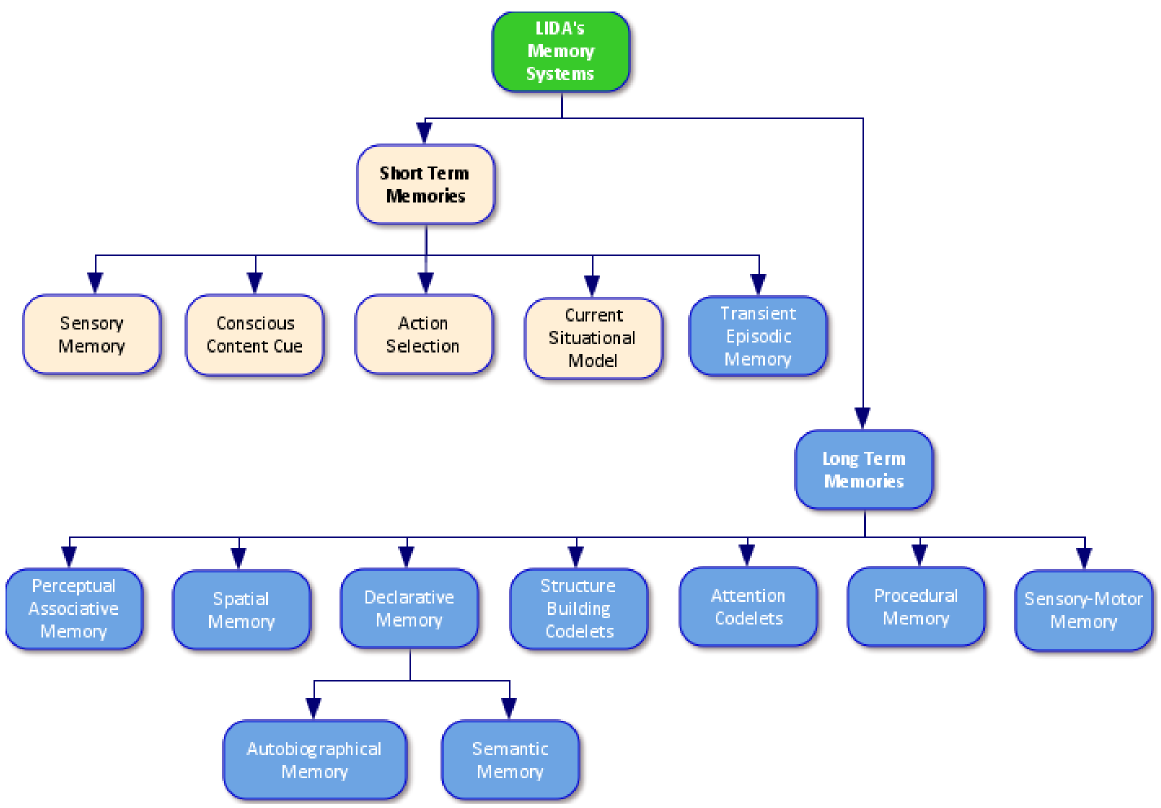

Figure 5. Memory Systems in LIDA

5.1 Sensory Memory

Incoming stimuli from each of the agent's various sensors are represented, and remembered briefly (some tens of milliseconds in humans), in LIDA's Sensory Memory. Early feature detectors, mostly in each sensory modality, process these representations. The resulting sensory information is passed simultaneously to both Perceptual Associative Memory and to the Current Situational Model in the preconscious Workspace.

The concrete implementation of Sensory Memory is still an open question, and several lines of research are being pursued. On one hand, deep learning approaches show promising visual recognition performance, comparing favorably to humans on some datasets (He, Zhang, Ren, & Sun, 2015), have been argued to learn brain-like representations (Khaligh-Razavi & Kriegeskorte, 2014), and have been used to interface LIDA to a realistic robotic simulator (Madl, Franklin, Chen, Montaldi, & Trappl, to appear). Another kind of Sensory Memory is being implemented as a set of Hierarchical Temporal Memory's (HTM) Cortical Learning Algorithms (CLA) regions (Hawkins, Ahmad, & Dubinsky, 2011). A CLA region is claimed to be a spatial and temporal pattern recognizer by Hawkins et al. It is modeled after the cortical regions in brain. Like cortical regions, CLA regions can be assembled into hierarchies (Felleman & Van Essen, 1991), for the performance of more complex pattern recognition. Such pattern recognition elements can be employed as feature detectors of LIDA's Sensory Memory. The Sensory Memory can be equipped with a set of these CLA regions. Each region would be specific to a certain kind of feature or pattern in the input, for example, shape of object, color of the object, characteristic sound of the object, and so forth. This can be done by selecting an appropriate set of sensors whose output will be fed to a CLA region for a particular kind of feature detection. A hierarchy of CLA regions can be used if the feature is very complex (Agrawal & Franklin, 2014).

5.2 Perceptual, Spatial and Episodic Memory systems

In the following subsections longer-term memory systems are described that feed into the preconscious Workspace and its Current Situational Model (see Figure 2).

5.2.1 Perceptual Associative Memory

Derived from the Slipnet in the CopyCat architecture (Hofstadter & Mitchell, 1995), LIDA's Perceptual Associative Memory (PAM) is the model's long-term (but see Section 4.7) recognition memory. Previously known entities are recognized and become part of the percept (see below).

PAM currently represents incoming sensory information using node and links, with nodes representing feature detectors, objects, feelings, actions, events, categories, concepts, etc., and the links representing relations between them, for example feature-of, category membership, inhibition, causation, thematic roles (Fillmore, 1968), etc. Each node has a base-level activation, a current activation (see Sections 4.6 and 4.7); some in addition have base-level incentive salience (McCall, 2014). Current activation passes along links each in an appropriate manner.

A node with no incoming link is considered to be on the frontier of PAM. The conceptual depth of a node in PAM is the minimal length of a chain of links beginning with a frontier node, and ending with the given node. The decay rate of the base-level activation of a PAM node decreases with its conceptual depth.

The distinctive, but by no means the only, higher-level data structure of PAM is the event (Chandler, 1991, 1993; Hohman, Peynircioğlu, & Beason-Held, 2012; Zacks, Kurby, Eisenberg, & Haroutunian, 2011; Zacks, Speer, Swallow, Braver, & Reynolds, 2007). In LIDA the event data structure consists of an event node together incoming thematic role links from thematic role nodes (Fillmore, 1968; McCall, Franklin, & Friedlander, 2010). Figure 6 illustrates an event with agent, action, and object thematic roles. Other thematic roles include beneficiary, source, destination, location, and instrument (Sowa, 1991, 2014). Though the illustration has labels, they are only for the convenience of the reader. In PAM links are typed but not labeled. Nodes in some modules are sometimes typed, but never labeled. Meanings arise in a grounded fashion from the network connections (Barsalou, 2008; Fuster, 2006) (see Section 4.2).

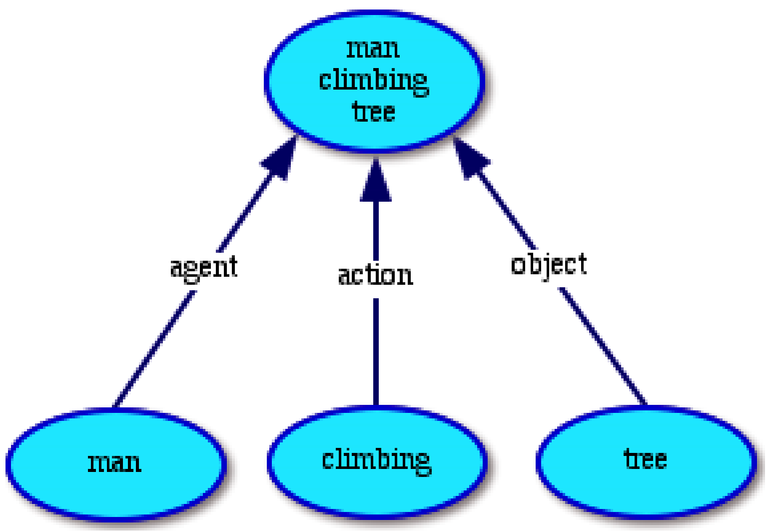

Figure 6. Event data structure example

In addition to coming from Sensory Memory, structures in PAM can be activated by cues from the Current Situational Model. Data structures in PAM whose total activation (some function of base-level and current activation) is over threshold have copies instantiated into the Current Situational Model as part of the percept.

Each conscious broadcast offers PAM an opportunity to learn new entities, and to reinforce the base-level variables of various entities (see Section 4.7). Entities also decay regularly (see Section 4.6).

5.2.2 Spatial Memory

Perceiving, representing and storing its own position and the positions of important objects in its environment are vital abilities for any embodied agent. Spatial Memory refers to the part of the memory systems that encodes, stores and recalls spatial information about the environment and the agent’s orientation. In LIDA, spatial representations (Figure 7) are first built in the Workspace. In addition to the identity of recognized entities, represented as PAM nodes, their relative positions to the agent can also be obtained perceptually (e.g. calculating depth-information from stereo disparity). These relative positions are represented as 'Egocentric Spatial Vectors' between the self representation and object representation, which are special kinds of PAM links containing position information. In addition, there is an allocentric spatial grid – a grid of PAM 'place nodes' representing specific locations in the environment – constructed and maintained by spatial structure building codelets (see Section 5.3.2). In addition to updating and maintaining egocentric links, these codelets also connect perceived objects to their corresponding place node, and update these connections during movement. There are clear neural correlates in brains corresponding to these two types of spatial information, among others in the hippocampal-entorhinal complex for allocentric and the precuneus for egocentric representations – see (Madl, Franklin, Chen, & Trappl, 2013).

Specific egocentric representations are transient and temporary, and do not need to be stored long-term. However, allocentric representations, if and when they become conscious, are stored in Spatial Memory, one of LIDA's long-term memory systems based on Sparse Distributed Memory. This long-term storage is not yet implemented computationally (work is underway to map grids of place nodes and associated objects to a concise graph representation, which can be efficiently encoded in Extended Sparse Distributed Memory (Snaider & Franklin, 2011)). Conversely, long-term Spatial Memories are also cued whenever relevant objects appear in the Workspace, helping to recall previously encountered allocentric maps.

Apart from storage and representation, the inference of accurate spatial positions from noisy data also presents significant challenges. Both the agent's own position, and that of significant objects around it, are uncertain and have to be inferred from inexact measurements. In robotics, probabilistic approaches have become very popular and successful to tackle this problem. We have found evidence in prior work that the assumption of statistically near-optimal use of information can partially explain the firing of hippocampal place cells (which represent allocentric spatial information) (Madl, Franklin, Chen, Montaldi, & Trappl, 2014), which is in line with the 'Bayesian brain' hypothesis (Knill & Pouget, 2004), and makes the probabilistic approach plausible for cognitive models of spatial memory as well. For this reason, path integration (self-movement) information, and distance information, are integrated in a Bayesian fashion when estimating positions (Madl, et al., to appear).

5.2.3 Transient Episodic Memory

Episodic memory is memory for events (episodes), often expressed as the what, the where, and the when (Tulving, 1983; Tulving, 2002). It is typically thought of as long-term, possibly lasting a lifetime. As pointed out, argued for and described in Section 4.10, the LIDA model includes a shorter-term version, Transient Episodic Memory (Conway, 2001; Franklin, et al., 2005) whose unreinforced memories last a few hours or a day in humans. It will not be described further here.

New events can be learned with each conscious broadcast, and old ones reinforced (see Section 4.7). Events may also decay away (see Section 4.6). During some offline time the as yet undecayed memories in Transient Episodic Memory are consolidated into Declarative Memory (see Section 4.11).

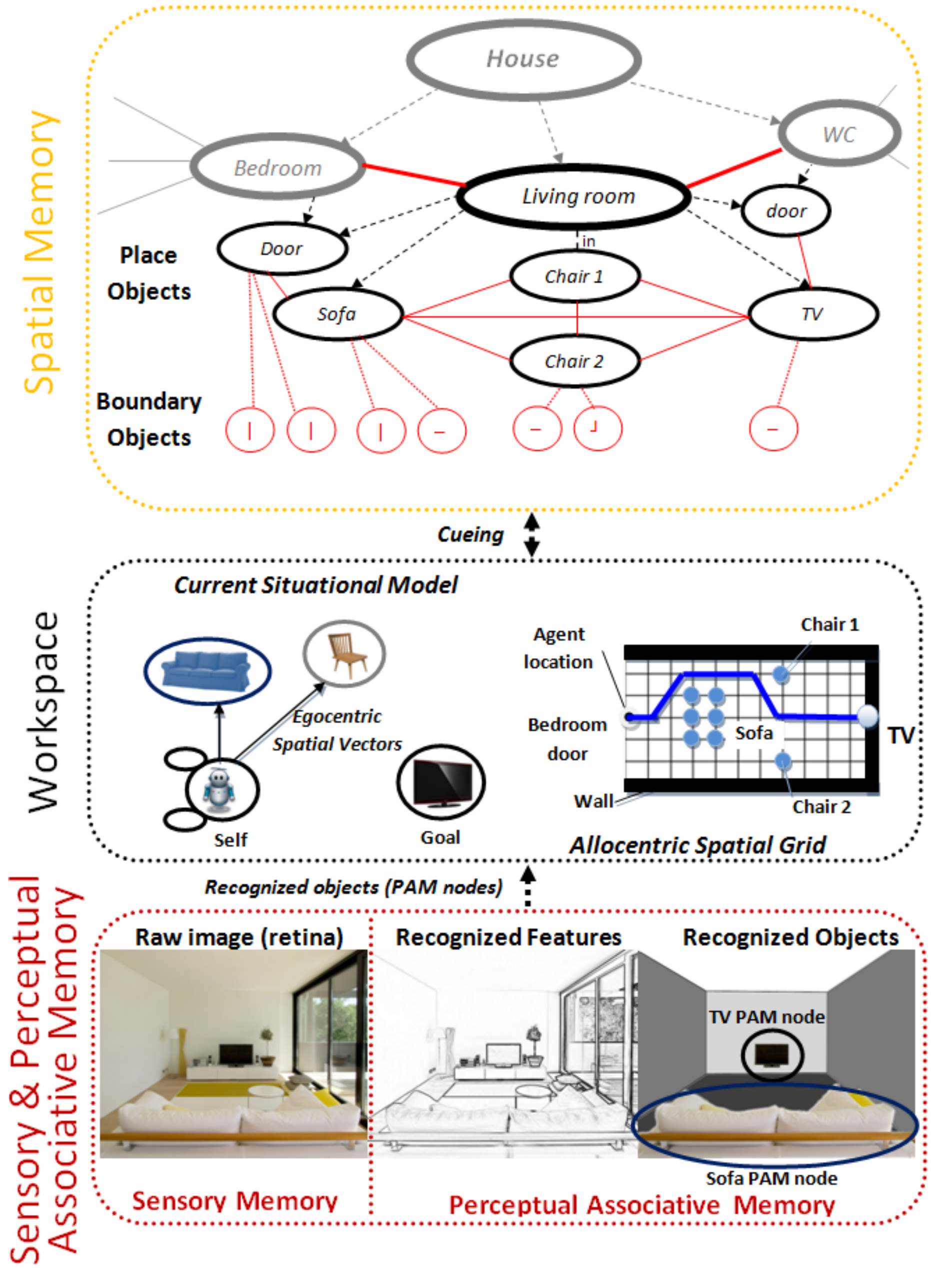

Figure 7. Recognized percepts (from Sensory Memory & Perceptual Associative Memory) are used to construct temporary egocentric (self-centered) and allocentric (world-centered) spatial representations in the Workspace, which in turn can be stored in, or can cue previous representations from, long-term Spatial Memory

5.2.4 Declarative Memory

Long term episodic memories of events, some capable of lasting a lifetime in humans, are stored in LIDA's Declarative Memory system. Rather than being learned from conscious broadcasts, as are other memories in LIDA, here they come to Declarative Memory from Transient Episodic Memory via offline consolidation (see Section 4.11). At this consolidation time, REM sleep in humans, whatever memories that have not decayed away in Transient Episodic Memory are consolidated into Declarative Memory. This consolidation includes the creation of memory traces for new events, and the reinforcing of traces of past events that have newly made their way to Transient Episodic Memory with sufficient affect to not have decayed away. How can this later situation occur? Suppose Event A arrives in the Current Situational Model (see Section 5.3.1) from Declarative Memory via local association (see Figure 2.) with sufficient affect to come to consciousness during a subsequent cognitive cycle. Then Event A will be learned into Transient Episodic Memory. If it does not soon decay away, it may be consolidated, in this case reinforced, in Declarative Memory.

In addition to the memory of full events with what, where and when, referred to as Autobiographical Memory (see Figure 5), Declarative Memory also contains traces that have lost their where and when to interference, while retaining their what in the form of facts, rules, etc. These are referred to as Semantic Memory (see Figure 5).

5.3 Preconscious Workspace

Unlike the long-term memory PAM (but like Baddeley's working memory (Baddeley & Hitch, 1974)), LIDA's Workspace (see Figure 2) is short-term, with latency measured in tens of seconds. Like PAM (but unlike Baddeley's working memory which requires consciousness (Baddeley, 1992)), LIDA's Workspace is preconscious in that its representations (data structures) are not conscious, but any of them can come to consciousness during a conscious broadcast (Franklin & Baars, 2010). In the following subsections we describe the Workspace's two modules, the Current Situational Model and the Conscious Contents Queue, and the Structure Building Codelets that process them.

5.3.1 Current Situational Model

Repeatedly taking in internal and external sensory information both directly from sensory memory, and from percepts from PAM (see Figure 2), LIDA's Current Situational Model (CSM) continually updates itself so as to keep track of the LIDA agent's current situation. Input from PAM comes in the form of node and link structures, while Sensory Memory input may have to be translated into such node and link structures by structure building codelets (see Section 5.3.2).

Structures arriving in the CSM automatically cue each of the attached long-term memory systems, PAM, Spatial Memory, Transient Episodic Memory, and Declarative Memory, resulting in local associations from each of them as appropriate (see Figure 2). Each of these local associations is itself an incoming structure, and so cues the long-term memory systems again sometimes producing new local associations. Thus new structures are continually added to the CSM.

Simultaneously, structures decay away at varying rates in ranges of tens of seconds. At any given time the content of the CSM can be quite extensive and complex.

Content of the CSM can be real, representing what is really occurring in the agent's internal and external environment, or virtual, including memories, desires, plans, imaginings, etc. Thus structures in the CSM, whether originating there through the efforts of structure building codelets (see Section 5.3.2), or having been instantiated from structures in PAM, must carry some designation as real or some form of virtual. We sometimes speak of a real scene and a virtual scene in the CSM (see Figure 8).

We humans can construct at least visual and auditory virtual images in our minds. These are produced from known entities from PAM, and must bring with them, in addition to their node/link structure, representations that can be used to produce these virtual images. These representations are illustrated in Figure 8 by rectangles hanging from nodes.

LIDA's CSM can also contain more complex structures such as plans, itineraries story plots, melodies, etc. (see Figure 8) Structure building codelets and attention codelets can find their concerns as substructures of one of these complex structures.

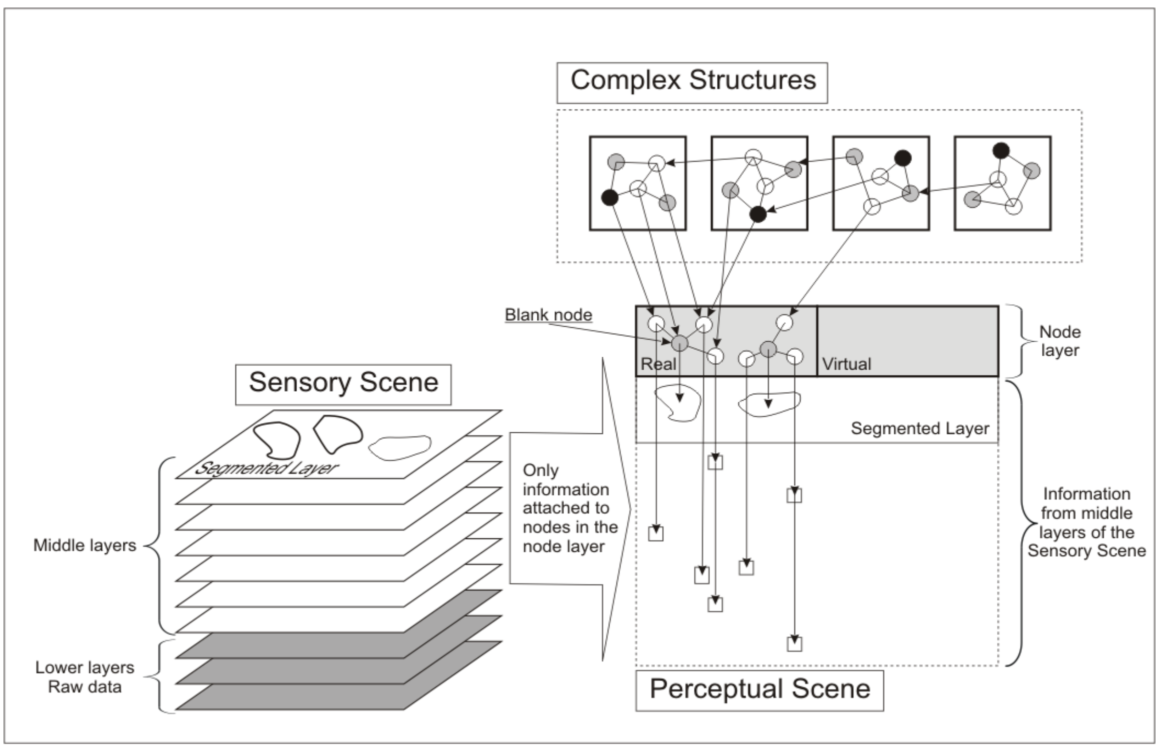

Figure 8. The Perceptual Scene and Complex Structures in the Current Situational Model.

5.3.2 Structure Building Codelets

In addition to arriving as percepts and local associations, various structures in the CSM can be created by structure building codelets. Structure building codelets are special purpose processes that support an agent's ability to recognize relationships between concepts and objects; for example, similarity, causality, etc. Structure building codelets continually monitor the CSM (and the Conscious Contents Queue) looking for content of interest. If this content is found, then the codelet will perform an action that will result in modifications to the CSM. Possible actions include creating new associations (links), creating new content (such as category nodes), or removing previous associations and content. For example, a structure building codelet that specializes in categorization might add an is-a-member-of link between an object node and a category node, while another with a different specialization might add an affordance link (Gibson, 1979) from an object node to an action node. Yet another may produce an option (see Section 6.2). Some structure building codelets required for spatial navigation are described in (Madl, et al., to appear).

Structure building codelets must also assign a current activation to each new structure it creates. This activation is a function of the current activations of its various raw materials (i.e., the preexisting structures of interest in the CSM), how well the raw materials match the concerns of the structure building codelet, and the base-level activation of the structure building codelet itself. The base-level activation of a structure building codelet is determined by how successful it has been in building structures that are consciously broadcast. As mentioned previously, the conscious broadcast is received by all LIDA modules (including the structure building codelets). When a structure building codelet recognizes content it built in the conscious broadcast, it will receive a small increase to its base-level activation. As a result, structure building codelets that consistently create "useful" structures will have higher base-level activations; structure building codelets that fail to create useful structures will slowly lose base-level activation, and may eventually be discarded.

Note that structure building codelets are profligate, just as learning is (see Section 4.7). That is, a structure building codelet will produce a structure of the type it is concerned with whenever it finds the appropriate raw materials in the CSM or the Conscious Contents Queue (described below). Thus many more structures are built than can possibly come to consciousness and, hence, be learned into some memory. The ones that are unlearned simply decay away in the few tens of seconds granted to CSM entities. Thus we can once again (see Section 4.7) think of the LIDA model as being Darwinian in nature, with only the fittest structures surviving (Rosenbaum, 2014).

The concept of the structure building codelet was inspired by the Copycat Project (Hofstadter & Mitchell, 1995) and follows in the tradition of Minsky's "The Society of Mind" (Minsky, 1985), which contends that intelligence emerges not from a single, monolithic and complex process, but through the interactions of a "society" of smaller processes.

5.3.3 Conscious Contents Queue

The Conscious Contents Queue (CCQ) (Snaider, McCall, & Franklin, 2010) is a very short-term memory system (we hypothesize it to last about three seconds in humans) that stores the past few tens of conscious contents (see Section 5.5). A newly broadcast conscious content is added to the end of the queue, pushing off the conscious content at the front of the queue. The Conscious Contents Queue is misnamed in that it is not actually a queue, since structure building codelets can help themselves to data from any point within, and not just from what pops off the front (see Figure 9). For example, a causation building codelet finding Event 1 newly in the CSM and an appropriate Event 2 recently in the CCQ, might create a causal link from Event 2 to Event 1 (Snaider, et al., 2010). Apart from causal links, this Queue can also be used to estimate the duration of events (by counting how many previous conscious broadcasts, stored in the Queue, contain an event), and is central to time perception (Madl, Franklin, Snaider, & Faghihi, 2015). However, probably the most important function of the CCQ is the grounding of time related concepts. In the same way that PAM nodes for sensory concepts, such as "red", are grounded in sensory memory, time concepts, such as "one second", are grounded in the CCQ.

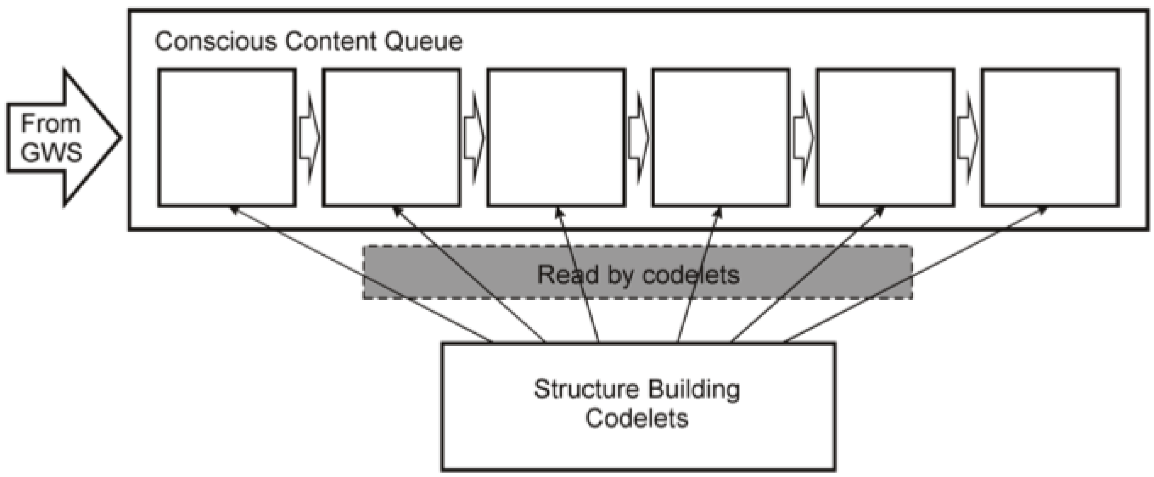

Figure 9. The Conscious Contents Queue

5.4 Attention Codelets

As we have seen in Section 5.3.1 at any time the contents of the CSM can be both complex and quite extensive. There can be an awful lot going on in a LIDA-based agent's world at any given time, too much for the agent to deal with at once. In phase two of the cognitive cycle as described in Section 2 and Figure 1, attention acts as a saliency filter, choosing the most salient (important, urgent, insistent, novel, unexpected, loud, bright, moving, etc.) structures to compete to become contents of the global broadcast. (See Section 4.4) This attention saliency filter is implemented by LIDA's Attention Codelets. Like the structure building codelets, each attention codelet keeps continual watch over the CSM looking for some structure that meets that codelet's particular concern for saliency.

Upon finding a suitable structure in the CSM, the codelet incorporates it into a coalition, which is then moved to the Global Workspace to compete for consciousness. The term "coalition" was chosen (Baars, 1988) since an attention codelet can include more than one structure in a coalition, and can also combine forces with other attention codelets to create a joint coalition.

The codelet(s) must also assign an activation to the new coalition, on the basis of which it will compete for consciousness. The amount of this activation depends on four factors (Madl & Franklin, 2012): the activation of the structures incorporated into the coalition, the base-level activation of the attention codelet, how well the structures match the particular concerns of the codelet, and a fourth that needs more explanation. When a winning coalition (see Section 5.5) has a particular strong activation, it drives the whole Attention Codelets module into a refractory period from which it gradually recovers. The earlier in this period, the less activation any attention codelet will assign to a new coalition. The base-level activation of the attention codelet factor insures that very salient structures, such as sudden motion in the visual periphery, an unexpected loud noise, etc. are incorporated into coalitions with high activation.

An attention codelet that successfully forms a winning coalition receives the resulting conscious broadcast, and reinforces its base-level activation. In theory, new attention codelets are formed from old ones using material in a conscious broadcast. As yet we have not developed this form of attentional learning.

There are at least four kinds of attention codelets. The *default attention codelet* observes the Current Situational Model in the Workspace, trying to bring the most activated structure to the Global Workspace. Thus it can be concerned with a broad spectrum of content, but its maximum activation is low. *Specific attention codelets* are codelets with specific concerns that have been learned. Each tries to bring particular Workspace content to the Global Workspace. *Expectation codelets*, mostly created during action selection, try to bring the result (or non-result) of the agent's recently executed action to consciousness. *Intention codelets* are attention codelets that bring to consciousness any coalition that can help the agent reach its current goal. When the agent makes a volitional decision (see Section 6.2), an intention codelet is generated.

5.5 Global Workspace

Attention codelets move their coalitions into the Global Workspace (see Section 5.4) where they compete to have their structures become the contents of the global broadcast (see Section 3), that is, they are broadcast to almost the entire LIDA model (see the orange arrows in Figure 2). The competition is a particularly simple one; the coalition with the highest activation wins. But the competition cannot be held continuously, so the question is when to hold it? The Global Workspace is one of two LIDA modules that do not operate completely asynchronously (see Section 4.7). But it does not operate on a clock either. Rather we have experimented with four different triggers, each of which can start the competition (Kaur, 2011).

The first trigger is a simple threshold on activation. When any coalition arrives with an activation over threshold, a competition is begun, with that strongly activated coalition becoming the winner. This trigger insures that structures with extraordinarily high salience have a high probability of coming to consciousness, and thus becoming the content of a global (conscious) broadcast (see Sections 3 and 4.4).

The second trigger occurs when the sum of the activations of the coalitions in the Global Workspace exceeds a collective threshold. This trigger is useful in those situations where a lot of activity of moderate saliency is occurring, but nothing of exceptional saliency.

A third trigger ensues when no new coalition arrives in the Global Workspace for a specified period of time. This trigger would apply to a very stable situation with little going on.

The fourth, and default, trigger happens when there has been no conscious broadcast for a specified period of time. Even say during meditation in humans when purposefully nothing of any saliency is occurring, consciousness does not cease. Rather something of relatively little saliency is broadcast.

Though the LIDA model includes the Global Workspace as a separate module, one must not infer that there is a corresponding place in brains where such a competition takes place. Rather recent work (Baars, et al., 2013) suggests that the competition for consciousness can occur throughout the cortico-thalamic core (see Section 4.9).

5.6 Procedural Memory

In the LIDA model Procedural Memory is the memory of what to do under a certain circumstance to achieve some goal. Following Drescher (1991, 1998), the basic data structure of Procedural Memory is the scheme[12], consisting of a context, an action, a result, and a base-level activation intended to measure the likelihood of the action, taken in the scheme's context, achieving the scheme's result. Both the context and the result are structures composed of nodes and links. An action in a scheme, represented by a node, can be a simple action such as reach, point to, pick up, turn to the right, etc., or a sequence, or even a stream with and/or branches, of such actions.

On receipt of a conscious broadcast, any scheme whose context overlaps significantly with the content of the broadcast instantiates a copy of itself, called a behavior, with its variables specified according to the conscious content. If the action of the scheme is a stream, we refer to the instantiated scheme as a behavior stream. The activation assigned to the instantiated behavior depends on the activation of the conscious content, on the base-level activation of the scheme, on the degree of coincidence of the conscious contents with the context of the scheme, and on the closeness of the scheme's result to any of the agent's goals (see Section 6.2). The instantiated behavior or behavior stream is then passed on to LIDA's Action Selection module.

If a behavior is selected and executed, and that event subsequently comes to consciousness, selectionist learning is triggered and the base-level activation of the scheme that generated the behavior is reinforced. If such a behavior comes from a behavior stream, the scheme that generated the behavior stream is reinforced. Instructionist learning takes place when the conscious content suggests that a new scheme be constructed from an old one, typically by adding or deleting structure from either a context or a result. The new scheme is assigned a base-level activation depending on that of the old, and on the activation of the conscious content. The old scheme remains in Procedural Memory as is.

---

[12] Drescher called it a schema. We altered that to scheme so as not to conflict with the different usage of 'schema' by psychologists.

5.7 Action Selection

LIDA's Action Selection mechanism, an enhanced form of Maes' behavior net (Maes, 1989), allows sophisticated, flexible action selection. Nodes in the network are instantiated behaviors that have come from Procedural Memory either singly or as part of a behavior stream. When a coalition is brought to consciousness, schemes in Procedural Memory will look at the contents to see if they match either the context or the desired result(s). Matching schemes will be self-recruited and instantiated as behaviors in Action Selection, where they compete for execution. A special deliberative scheme will always compete to start the volitional action selection process (see Section 6.2 below).

The successor (forward), predecessor (backward) and conflictor links in the Action Selection network are defined in terms of the contexts (preconditions) and results (add and delete lists) of the behaviors (nodes) in the network. Activation along a successor link strengthens its sink behavior, if the source result satisfies a sink precondition. Activation along a predecessor link strengthens its sink behavior if a source precondition is satisfied by a sink result. Along a conflictor link, activation from a source behavior inhibits the sink behavior since it can undo one of the source's context conditions. Several conditions can factor into which behavior is chosen. The selected behavior must match the appropriate context (preconditions)—for example, if an agent throws a ball with the expectation that it would be caught, there should be another agent present that is capable of catching the ball. Then, it must have at least the current threshold level of activation (see below). Selection may also be influenced by leftover behaviors from previous cycles that have not decayed away. A selected behavior whose action is external is passed to Sensory Motor Memory for execution, or in the case of an internal action, such as volitional action selection (see Section 6.2), is sent back to the Current Situational Model to take an internal action, such as to set up a competition for conscious decision-making.

Since behaviors filter into Action Selection asynchronously, a trigger system must determine when an action is chosen. There are three possible triggers similar to those of the Global Workspace (see Section 5.5): 1) *If a behavior is above a certain threshold level.* Any action of moderate interest to the agent should satisfy this requirement. 2) *If the total activation of all the behaviors in the Action Selection network is above a certain threshold level.* This can occur if there are many actions an agent can choose from, but nothing of great interest. 3) *If no behavior has been executed within a certain amount of time.* For example, during volitional action selection, multiple contests can be held without executing any behavior. In this case, the threshold levels of the actions should be gradually lowered to facilitate the deliberating process.

When a behavior is finally chosen, an expectation codelet is sent to the Attention Codelets to observe the Current Situational Model for the results of the action performed. This codelet should record both expected and unexpected outcomes, enabling the agent to build new schemes with additional context or result items should a coalition built by the expectation codelet come to consciousness. If the action produced the desired result, the scheme that produced the chosen behavior will receive an increase in base-level activation, and all schemes that could have been chosen with the same expected result will also be given a slight boost. However, if the action produced an undesirable result, the scheme will receive a decrease in base-level activation. A desirable result is defined as one with total associated feelings of positive valence, while an undesirable result is defined as one with total associated feelings of negative valence. Thus, the agent should continue to modify schemes so as to increase the probability of desirable results. Schemes with overwhelmingly undesirable results should eventually not be recorded in Procedural Memory, but might remain in Semantic Memory (facts) or Episodic Memory (events); however, it is normal for a scheme to have a mix of positive and negative results.

5.8 Sensory Motor Memory and Motor Plan Execution

Action execution in LIDA refers to a LIDA agent transforming a selected goal-directed action, the selected behavior, into low-level executable actions, motor commands, and executing them.

When an agent has selected an action, it understands what it will do before the execution begins; but normally this understanding is not executable, because the needed detailed environmental information is not yet available. Milner & Goodale have proposed a hypothesis in their work on the two visual systems (Goodale & Milner, 1992; Milner & Goodale, 2008), the ventral and dorsal streams, providing "vision for perception" and "vision for action" respectively[13]. Regarding action execution, they suggest that the dorsal stream "is critical for the detailed specification and online control of the constituent movements that form the action" (Milner & Goodale, 2008, p. 775).

In LIDA, action execution is modeled by the Sensory Motor System (SMS) (Dong & Franklin, 2015b), using two LIDA modules: Sensory Motor Memory and Motor Plan Execution (see Figure 2). Two other LIDA modules, Action Selection and Sensory Memory, provide input information to the SMS. Action Selection forwards a selected behavior, while the Sensory Memory sends data through a dorsal stream channel providing the most current detailed environmental information. The SMS sends out motor commands to the agent's actuators for appropriate movement. Within the SMS, two data structure types have been implemented—the Motor Plan Template (MPT), and the Motor Plan (MP)—and three types of processes have been modeled: online control, specification, and MPT selection.

A MP is designed based on the subsumption architecture (Brooks, 1991), a type of reactive motor control mechanism. In the subsumption architecture, 1) the sensory data is directly linked to the selection of motor commands that drive the actuators; 2) it decomposes a robot's control architecture into a set of task-achieving behaviors; and 3) it does not maintain internal models of the world. The MP generates motor commands as the output of the SMS to the environment (using actuators), while environmental data from the dorsal stream channel from Sensory

[13] In the LIDA model, the concept of ventral and dorsal streams for the transmission of visual information has been extended to multimodal transmission.

Memory directly influence the generation process. These cyclically occurring processes are called the online control process of the SMS.

A motor command (MC) is applied to an agent's actuator. Every MC has two components: a motor name, and a command value. The motor name indicates which motor of an actuator the MC specifically controls, while the command value of a MC encodes the extent of the command applied to the motor.

A set of MCs is prepared inside a Motor Plan (MP), and bound with fixed command values. In order to specify a MC's command value before the execution begins, a Motor Plan Template (MPT) and a specification process are created in the SMS. A MPT is an abstract motor plan that resides in Sensory Motor Memory. It has a set of motor commands that are not yet bound with the command values, whereas after a specification process, the motor commands are bound with specific values using the sensory data sent from Sensory Memory, instantiating the MPT into a concrete MP.

As the SMS's initial process, a MPT selection acts to select and initiate a MPT for an incoming selected behavior before the MPT is specified into a concrete motor plan. The MPT selection chooses one MPT from others associated with the selected behavior; it connects action selection to action execution.

Recently we have addressed the learning process in action execution (Dong & Franklin, 2015a). We implemented a model of sensorimotor learning in LIDA using the concept of reinforcement learning. This learning helps an agent generate effective motor commands in a certain context using past experiences. Following Global Workspace Theory, the learning is cued by the agent's conscious content, the most salient portion of the agent's understanding of the current situation (See Figure 2).

6 Modes of action selection

Every autonomous agent (Franklin & Graesser, 1997), be it human, animal or artificial, must iteratively and frequently answer the fundamental question "what do I do next." Thus, according to LIDA's definition in Section 2, action selection is a (the?) fundamental activity of autonomous agents. LIDA-based agents make such selections using one of four modes: consciously mediated action selection, volitional decision making, alarms, and automatized action selection. The first two of these modes correspond to Kahnemann's System 1 and System 2 (Faghihi, Estey, McCall, & Franklin, 2015; Kahneman, 2011). Sloman has proposed three levels of cognitive processes, the reactive, deliberative, and metacognitive (1999). Our consciously mediated action selection occurs as a reactive process à la Sloman, while volitional decision making is a deliberative process. Metacognitive decision making is envisioned as being implemented via deliberative processing in LIDA, but has yet to be implemented. Each of our modes will be described in turn in the following subsections.

6.1 Consciously Mediated Action Selection

During each of LIDA's cognitive cycles (see Section 3), that is in humans five to ten times a second, there's the opportunity for an action to be selected (Madl, et al.,

2011) and it's execution begun (Dong & Franklin, 2015b). Occurring every cycle or two, these actions are selected making extensive use of the contents of consciousness, but the selection process itself is never conscious (Franklin & Baars, 2010). We refer to this so common and frequent mode of action selection as being *consciously mediated*. For example, a thirsty agent reaching for a glass of water on the table may well have performed consciously mediated action selection, having been consciously aware of the thirst and of the location of the glass of water. Almost all of our speech in every day life is consciously mediated.

6.2 Volitional Decision Making

In contrast to consciously mediated action selection, some action selection occurs via processing that is itself partly conscious (deliberative). Following Global Workspace Theory (Baars, 1988) the LIDA model hypothesizes that such volitional decision making is accomplished using William James' ideomotor theory (Franklin, 2000; James, 1890). Such volitional decision making typically takes place over many, many cognitive cycles.

The major players (processes) in the LIDA version of ideomotor theory are a timekeeper, a proposer, an objector, and a supporter. The process begins with an option coming to consciousness (say the agent is thirsty, and the option is "let's have a beer.") This conscious option may instantiate several schemes (see Section 5.6) for accomplishing it (having a beer). It will also recruit and instantiate a deliberation scheme capable of implementing ideomotor theory. Perhaps one of the schemes effecting the option wins out; perhaps not. In the latter case, perhaps the deliberation scheme wins. Then a timer corresponding to the option will start in the Current Situational Model.

If no objection comes to consciousness, and the timer runs out, the option is converted to a goal, and will typically come to consciousness. There it will recruit and instantiate schemes to bring about the goal, and send these behaviors to Action Selection. If before the timer runs out an objection comes to consciousness ("its too early for a beer"), then the timer is turned off. If a supporter comes to consciousness ("oh, it's not that early"), then the timer is turned back on. Another objector can arise, or not. Or perhaps, instead, another proposer enters the fray ("let's drink water"). Another timer, timing the new proposal is turned on, and the process continues.

Notice that this process occurs over multiple cognitive cycles using consciously mediated action selection in such a way as the conscious contents are part of the decision making process. Note also that deliberative decision making makes direct use of consciously mediated action selection, rather than being separate "systems" a la Kahneman (Faghihi, et al., 2015; Kahneman, 2011).

6.3 Alarms

Many drivers have experienced another car suddenly swerving in front of them, and experiencing having already pressed the brake and turned the wheel while becoming conscious of the other car. Following Sloman (1998; 2001), such unconsciously selected actions are referred to in the LIDA model as *alarms*. In an alarm situation, one selects, and executes action(s) to deal with a dangerous situation prior to becoming conscious of the situation.

If an event is recognized by Perceptual Associative Memory (PAM) as an alarm, with perhaps a stop in the Current Situational Model to gather more details of the situation that came through Sensory Memory, it will proceed directly to Procedural Memory, skipping the consciousness broadcast. Then a scheme associated with that alarm will recruit itself, and trigger an action selection contest immediately due to the high activation of the original alarm. This type of action selection is an immediate, learned reaction that bypasses the conscious broadcast, and thus saves time.

Learning alarms, that is, learning to bypass attention, is a form of attentional learning (Chun & Jiang, 1999; Häkkinen, 2010; Liddell et al., 2005; Mateo, 2010; Miller & Fu, 2007; Ogawa & Yagi, 2002). Once we have learned that a situation is dangerous, it can influence our decision-making, reaction time and intensity, and attentional process (LeDoux, 2000; Rolls, 2000; Sloman, 1998; Squire & Kandel, 2000)

6.4 Automatized Action Selection

When walking down an empty sidewalk, a person looks ahead, sees that the way is clear, and then for the next few steps can attend to something else while each step calls the next directly. In the LIDA model we refer to this process as *automatized action selection* (Franklin, 2003), and think of it as a trivial application of pandemonium theory (Jackson, 1987). As yet this mode of action selection in LIDA is purely conceptual, having not been implemented.

7 LIDA-based Agents

The subsections below are devoted to descriptions of the various LIDA-based agents that have been implemented to date by members of the Cognitive Computing Research Group at the University of Memphis. One of these is a software agent with a real world task, while the others are all simulations of behavioral or neuroscience studies. A few other such have been contributed by researchers outside of our group (Becker, Fabro, Oliveira, & Reis, 2015; Hernes, 2014). Many of these LIDA-based agents were implemented using the LIDA Framework described in Section 8.

7.1 IDA

A software agent IDA (Intelligent Distribution Agent) (Franklin, Kelemen, & McCauley, 1998), the forerunner of the LIDA model developed for the US Navy (McCauley & Franklin, 2002), is presented here for historical reasons. IDA respects most of the conceptual commitments described in Section 4. Its architecture incorporates most of the modules found in LIDA's cognitive cycle (see Figure 2).

"Distribution" is the Navy's term for the process of assigning new billets (jobs) to a sailor at the end of his or her tour of duty. This process is carried out by Navy personnel called detailers, who communicate with the sailors in their community (under their jurisdiction) via either telephone or email. IDA was developed to automate the task of the detailer, communicating and negotiating with sailors using email in unformatted English. IDA must also query existing Navy databases for personnel records, requisition lists (needed job), etc. In choosing jobs to offer a sailor, IDA must consider the current needs of the Navy, the Navy's personnel policies, and the sailor's preferences. IDA was tested and accepted by the Navy, and a commercial software firm was employed to adapt IDA to the Web.

IDA's environment is the Internet. It senses and outputs only text (ascii characters). Its architecture is based on a slightly simplified version of the LIDA cognitive cycle. IDA is functionally conscious (Franklin, 2003). Completely hand crafted, IDA employs no learning.

7.2 Timing agent and Allport agent

The LIDA Timing agent and Allport agent were developed to test three aspects of LIDA's cognitive cycle, by means of comparison with human data: its duration, and its discrete conscious broadcasting mechanism (Madl, et al., 2011), and its ability to attend to images in a rapid serial presentation paradigm (Madl & Franklin, 2012).

The first agent operated in a very simple environment, consisting only of a light (which could be red or green) and a button. The agent had the simple task of pressing the button as soon as it became conscious of the light turning green – similarly to standard reaction time tests. The durations of each phase of the cognitive cycle were adjusted according to neuroscientific evidence, to 80-100ms for visual perception, an additional 100-200ms for the understanding/attending phase, and 60-110ms for the action selection phase. Average cognitive cycle length in this simulation was 283ms. The agent did not account for temporal expectation (human subjects engaging motor circuits before pressing the button – being 'on the brink of pressing it' – and just waiting for the green light can accelerate reaction times). Please note that more complex tasks require multiple cognitive cycles, which can overlap, allowing much faster sampling of the environment (up to ~10Hz).

The LIDA Allport agent was developed to verify whether LIDA, despite of its discrete consciousness mechanism, can still account for empirical findings which seem to favor a continuous mechanism of conscious perception. Specifically, Allport (1968) has developed a paradigm where subjects are seated in front of a screen which displays a single horizontal line in one of 12 possible positions, moving (changing position) upwards. They are asked to adjust the speed of this line until they arrive at the threshold of being able to consciously perceive movement, in two tasks. In both tasks, they first start with a slow line and increase its speed, arriving at time τ_1 of no perceived change, and subsequently start with a rapid line and slowly decrease its speed until the brink of seeing movement again at time τ_1 (the 'speed' of the line is measured by the time τ it spends in one position before jumping to the next). The first task allows lines to traverse the entire screen, and the second task simply leaves the lower half of the screen blank for exactly the duration that a line would take to traverse that half. Allport argued that if consciousness were to be discrete, two different cycle times τ_1 and τ_2 would necessarily arise in the second task. Using an analogy from a cinema, if consciousness consisted of discrete 'frames', like a 20th century film, movement cannot be perceived if it only falls within the duration of a frame and doesn't extend beyond it. In the second task, this can happen at two times τ, when the line traverses the upper half of the screen within a 'frame duration', and when it traverses the entire screen within a 'frame duration'. Thus, discrete consciousness should lead to different τ_1 and τ_2. However, Allport found statistically indistinguishable times, and concluded that consciousness must thus be continuous. Using the LIDA Allport agent, we could show that a discrete consciousness mechanism, too, can produce the same result – almost equal τ_1 and τ_2 – provided that old contents of consciousness persist for a time in new broadcasts ('frames'), until they decay away (Madl, et al., 2011)

7.3 Attentional blink agent

The LIDA Attentional blink agent (Madl & Franklin, 2012) reproduced a known phenomenon of human attention during a rapid serial visual presentation paradigm. In this paradigm, subjects are asked to attend to, and report, two 'targets' belonging to a specific class of stimuli within a rapidly changing sequence of 'distractor' stimuli (e.g. two letters or 'targets' within a stream of digits or 'distractors'). Subjects easily identify and report both targets if they are half a second or more apart. Somewhat counterintuitively, they also find it easy to report targets coming right after one another, even if the delay between them is as short as 100ms, but have trouble perceiving and reporting the second target if there is a distractor in between the targets. As an example, denoting targets with T and distractors with D, subjects will usually correctly report both targets in TTDDDD and TDDDDT, but will almost always fail to report the second target in TDTDDD. The inability to perceive and report the second target shortly after the first has been dubbed 'attentional blink' in the literature. The Attentional Blink agent aimed to reproduce this paradigm, based on LIDA's attention and consciousness mechanisms, and on the hypothesis that there is an attentional resource which gets temporarily depleted when looking out for and attending to the target (corresponding to the locus coeruleus-norepinephrine system, and operating on a matching timescale as that of this system in the brain). This assumption of a limited attentional resource which takes some time (about 400ms) to recharge allowed this agent to accurately reproduce human performance in this paradigm (Madl & Franklin, 2012).

This agent has the advantage of being more general than most other computational cognitive models of the attentional blink, being part of a systems-level cognitive architecture, as opposed to focusing on this single phenomenon (with the exception of the Threaded Cognition model, which is based on the ACT-R cognitive architecture (Taatgen, Juvina, Schipper, Borst, & Martens, 2009)). A further difference from other models includes the competition for consciousness between targets and distractors (thus, both of their saliencies influence the outcome). Finally, since LIDA's GWT-based consciousness mechanism is consistent with oscillatory synchrony-based accounts (see Section 4.13), it is also consistent with the implicated importance of oscillatory activity in the attentional blink (Janson & Kranczioch, 2011).

7.4 Attentional learning agent

Attentional learning is learning to what to attend (Estes, 1993; Gelman, 1969; Kruschke, 2010; Vidnyánszky & Sohn, 2003; Yoshida & Smith, 2003). In the following we will give a conceptual explanation of attentional learning in LIDA, followed by a brief description of a LIDA-based agent capable of attentional learning.

Attention in the LIDA model is primarily implemented by attention codelets that are stored in Attentional Memory (ATM) (Figure 5) (labeled Attentional Codelets in Figure 2). As described in Section 4.7 above, two kinds of attentional learning may occur each time a conscious broadcast comes to ATM. In *selectionist learning,* the attention codelet that wins the competition to bring a coalition to consciousness has its base-level activation strengthened. In *instructionalist learning*, according to the context of the agent's current task a new attention codelet is created from the winner with a more specific content of concern. According to the principle of profligacy (see Section 4.7), each conscious broadcast can lead to selectionist and/or instructionalist learning in each mode. Thus, learning occurs with the least provocation, but learned entities decay away unless they are later reinforced.

We will first consider instructionalist attentional learning. During LIDA's cognitive cycles (see Figure 2), percepts from Perceptual Associative Memory and local associations from Spatial Memory, Transient Episodic Memory, and Declarative Memory continually enter the preconscious Workspace's Current Situational Model. Such content can be acted upon by structure building codelets, and by attention codelets, which detect events or other structures salient to them. The default attention codelet responsible for creating coalitions of content happening for the first time is a primitive, built-in attention codelet, which competes among other attention codelets to bring the most activated content to consciousness. When a coalition created by this attention codelet wins the competition for consciousness. ATM's attentional learning mechanism then creates a new *specific attention codelet*. This new codelet's concern is set to be the most highly activated part of the winning coalition. The new specific attention codelet will have an initial base-level activation based on the default attention codelet's base-level activation and the coalition's current activation. In this way, an attention codelet is created in ATM for each broadcast of conscious content for which there is not already a dedicated attention codelet.

In some LIDA-based agents, we humans for example, the agent comes with primitive, built-in default attention codelets, such as described in the previous paragraph, for each of a number of types of salience, say among motion, brightness, loudness, unexpectedness, novelty, importance, urgency, insistency, etc. Instructional learning produces new attention codelets built from these default codelets as described above, allowing the agent to learn to what to attend in each of these types of saliency.

Selectionist learning occurs for an existing attention codelet when its coalition wins the competition for consciousness. That is, the base-level activation of the attention codelet in the winning coalition gets reinforced.

Expectation codelets (see Section 5.4) have their base-level activation adjusted whenever their coalition wins the competition for consciousness. Satisfied expectations result in increases, unsatisfied in decreases. If no similar attention codelet exists already then this expectation codelet is learned as a new attention codelet.

To simulate a human experiment, a LIDA agent was created (Faghihi, McCall, & Franklin, 2012) according to a human experiment realized by Van Bockstaele's (Van Bockstaele, Verschuere, De Houwer, & Crombez, 2010). The agent's task was to respond to an on-screen target that appeared either on the same side as a cue presented previously, or on the opposite side of the cue. The cues and targets were presented using two white rectangles on the computer screen. One white rectangle was on the left, and the other on the right. The cue consisted in one of the white rectangles being briefly recolored to either green or pink, at random. The target was a black rectangle randomly presented inside one of the white rectangles, and displayed until the subject responded by pressing a key to indicate whether the target was located left or right (For more information the readers are referred to (Faghihi, et al., 2012)).

For example:

Situation 1, CONGRUENT trial:

a) Both white rectangles are presented for 1000ms.
b) The cue appears in the place of the LEFT white rectangle for 200ms.
b) 20 ms break (both white rectangles empty).
c) The target appears in the LEFT white rectangle.

Situation 2, INCONGRUENT trial:

a) Both white rectangles are presented for 1000ms.
b) The cue appears in the place of the LEFT white rectangle for 200ms.
b) 20 ms break (both white rectangles empty).
c) The target appears in the RIGHT white rectangle.

For this experiment, the attention agent would respond "left" if the target appeared on the left or "right" if the if the target appeared on the right.

The LIDA agent's reactions for congruent trials were 360ms on average, whereas the average reaction time for incongruent trials was 380ms. This performance was similar to that found in human participants in Van Bockstaele et al. (2010). The experimenters concluded that the 20ms difference in reaction time was due to the fact that the cues attract attention, and thus targets appearing on the same side as the cue elicit a faster reaction time than targets appearing on the side opposite from the cue.

In this experiment, both instructionalist and selectionist learning occurred for each conscious broadcast. Whenever the default attention codelet was responsible for creating the winning coalition, a new attention codelet was acquired (in an instructionalist manner) with its content of concern equal to that of the broadcast. If a non-default attention codelet is responsible for a winning coalition, its base-level activation is reinforced. The default attention codelet's base-level activation was already saturated.

7.5 Medical Agent X (MAX)

Medical Agent X (MAX) is an agent under development to replicate cognitive functions relevant to medical diagnosis (Strain & Franklin, 2011; Strain, Kugele, &

Franklin, 2014). The initial implementation of MAX will focus on the diagnostic reasoning process known as differential diagnosis, in which a ranked list of possible causes for a patient's condition is generated, and a line of investigation is developed to rule in or rule out the identified possibilities. Another important problem is the extraction of clinical information from natural language medical records. These two processes are related, since diagnostic investigation typically requires significant research into the prior medical record.

Much of our earlier work on LIDA has focused on cognitive processes that occur within a single cognitive cycle; MAX's reasoning will require multiple cognitive cycles. MAX must "sense" and "perceive" clinical information in various forms, apply medical knowledge to generate relevant hypotheses, and select actions to evaluate, compare, and refine those hypotheses. If initial work, involving hand-coded medical knowledge, is successful, future work would include the development of learning mechanisms for MAX.

While the other agents in the LIDA bestiary--with the lone exception of LIDA's precursor, IDA, described above-- replicate human psychological phenomena for comparison with experimental studies, MAX seeks to test LIDA's conceptual model by applying it to a real-world problem with current human performance as the benchmark. We have termed this the *engineering fork*, as opposed to the *science fork*, of the LIDA methodology. MAX's goal is to replicate human diagnostic reasoning in a computational model as a technological application of LIDA's cognitive theories.

8 LIDA Framework

The LIDA Framework is a software framework written in the Java programming language that simplifies the process of developing LIDA agents. The framework implements much of the low-level functionality that is needed by most, if not all, LIDA agents including initialization, asynchronous and concurrent task management, and object creation. The framework also provides default implementations for many of the LIDA modules (see Table 1 below for a list). As a result, simple LIDA agents can often be created with a modest level of effort by leveraging "out of the box" functionality.

The framework contains a set of configuration files that specify global and module-specific parameters. By externalizing an agent's parameters in the framework's configuration files, developers can modify an agent's behavior without modifying its code. This has a number of advantages including improved maintainability and code reuse. It can also be useful for experimentation and parameter optimization. Included in the configurable parameters are the fully-qualified names of classes that implement the LIDA modules. These classes are instantiated by the framework during initialization using the Java Reflection API. Developers that require functionality not available in the default classes can replace the default class names in the configuration files with the names of their own classes. In this way, developers are empowered with the ability to easily extend or override default module behavior with custom modules and module initializers.

The framework also implements a multithreading engine, the task scheduler, that executes the operations required by the different modules. We called the basic operations "tasks". For example, attention and structure building codelets are implemented as tasks. Each task has a relative duration (compared with the duration of other tasks), and the task scheduler is responsible for these durations. This mechanism makes it possible to implement most simulations of behavioral or neuroscience experiments such as those described in the Section 7.

Historically, the LIDA Framework has used a data structure called the NodeStructure to represent much of an agent's transient and long-term knowledge. The NodeStructure is based on a graph-theoretical approach to knowledge representation in which entities are represented as nodes and associations are represented as links. NodeStructures are appealing because they are easily visualized (for node structures of moderate size) and associations between entities are trivial to create. Unfortunately, they suffer from several disadvantages.

Comparing NodeStructures can be computationally expensive. This presents a significant challenge because calculating the similarity between NodeStructures is a fundamental and ubiquitous operation. NodeStructures also do not work well with many of the state of the art learning strategies such as deep neural networks, which generally produce high-dimensional vectors as outputs. These and other limitations of NodeStructures have inspired the design of alternate framework implementations that utilize different common data structures. The vector framework (Snaider & Franklin, 2014b), which is based on MCR vectors (Snaider & Franklin, 2014a) is one promising alternative that is currently being developed. An abstract framework is also being developed that uses data structure agnostic LIDA module interfaces and core classes in order to maximize developer flexibility at the expense of limited opportunities for default module implementations.

Java remains one of the most popular general purpose programming languages because of its portability, support for object-oriented design, built-in memory management and concurrency support, and the proliferation of high-quality Java software libraries. By implementing the LIDA Framework in Java, we hope to make our framework, and hence our model, accessible to a large audience. The current version of the LIDA framework is available for download from the CCRG website. A detailed introduction to the LIDA Framework is available in (Snaider, McCall, & Franklin, 2011).

9 Current work and future directions

Work on the LIDA conceptual model and on its computational implementation continues. On the conceptual side, exploration of some of the so many and varied roles of structure building codelets (see Section 5.3.2) is of particular interest. Effort is continuing to specify the role of early perception, the relationship between Sensory Memory (see Section 5.1) and Perceptual Associative Memory (see Section 5.2.1). Further extensions of Sensory Motor Memory and Motor Plan Execution (see Section 5.8) so as to accommodate the effects of priming are proving necessary. Continued work on Medical Agent X (see Section 7.5) is beginning to lead us to think about aspects of deliberative (multi-cyclic) problem solving (see Section 6.2).

The computational instantiation of conceptual LIDA is still underway. In addition to progress leading to the implementation of Medical Agent X, there is also work on a LIDA based agent simulating a human priming experiment (Schmidt, 2002).

Table 1 contains a list of LIDA's major modules and mechanisms (cf. Figure 2 above). It also indicates their implementation status in the LIDA framework and in LIDA agents (some mechanisms exist in individual specialized agents but have not yet been transferred to the much more general computational framework), as well as references to papers describing them.

Planned future work on LIDA is computational in nature. The conceptual view of structures in LIDA is graph theoretical, based on nodes and links (see Section 5.2.1). Plans are afoot for a LIDA Framework (see Section 8) with structures based instead on vector representations (Snaider & Franklin, 2014b). A second plan involves the design and implementation of a LIDA based simulated robot in an artificial environment (Koenig & Howard, 2004) that will go through a developmental period in which it will learn to recognize entities and activities in its environment, and to respond appropriately to events.

Mechanism / Module	**LIDA Frame-work**	**LIDA agent(s)**	**References**
Sensory Memory	P	P	(Agrawal & Franklin, 2014; Franklin, et al., 2014; McCall, Snaider, & Franklin, 2010)
Perceptual Associative Memory	P	P	(Franklin, et al., 2014; McCall, Franklin, et al.,
Structure Building Codelets	P	P	(Franklin & Baars, 2010)
Conscious Contents Queue	F	F	(Snaider, McCall, & Franklin, 2012)
Workspace	F	F	(Franklin & Baars, 2010)
Spatial Memory	N	P	(Madl, et al., to appear; Madl, et al., 2013)
Transient Episodic Memory	F	F	(Franklin, et al., 2005)
Declarative Memory	F	F	(Franklin, et al., 2005)
Attention Codelets	P	P	(Faghihi, et al., 2012; Madl & Franklin, 2012)
Global Workspace	F	F	(Baars, et al., 2013; Franklin, et al., 2013; Franklin, et al., 2012)
Procedural Memory	F	F	(Franklin, et al., 2005)

Action Selection	F	F	(Negatu, D'Mello, & Franklin, 2007; Negatu & Franklin, 2002), Sections 5.7, 6
Sensory Motor Memory	P	P	(Dong, 2014; Dong & Franklin, 2015b)
Motor Plan Execution	P	P	(Dong, 2014; Dong & Franklin, 2015b)
"Embodiment", interface to robot	N	P	(Madl, et al., to appear).
Emotions, Appraisal	N	N	(Franklin, et al., 2014; Franklin & Ramamurthy, 2006)
Learning	P	P	(Faghihi, et al., 2012; Franklin, et al., 2005; Franklin & Ramamurthy, 2006)
Alarms	N	N	(Sloman, 1998; Sloman, 2001)
Volitional Decision Making	N	N	(Franklin, 2000; Kondadadi & Franklin, 2001)
Moral decision making	N	P	(Madl & Franklin, 2015; Wallach, Franklin, & Allen, 2010)

Table 1. LIDA's modules and mechanisms and their implementation status (F – fully implemented, P – partially implemented, N – not yet implemented).

10 References

Agrawal, P., & Franklin, S. (2014). *Multi-Layer Cortical Learning Algorithms*. Paper presented at the IEEE Symposium Series on Computational Intelligence (SSCI).

Allport, D. A. (1968). Phenomenal simultaneity and the perceptual moment hypothesis. *British Journal of Psychology, 59*, 395-406.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington, DC: Winston &. Sons.

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.
Baddeley, A. (1992). Consciousness and Working Memory. *Consciousness and Cognition, 1*, 3–6.
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. *Annu. Rev. Psychol., 59*, 617-645.
Bastos, A. M., Vezoli, J., & Fries, P. (2015). Communication through coherence with inter-areal delays. *Current opinion in neurobiology, 31*, 173-180.
Becker, T., Fabro, J. A., Oliveira, A. S. d., & Reis, L. P. (2015, 8-10 April 2015). *Adding Conscious Aspects in Virtual Robot Navigation through Baars-Franklin's Cognitive Architecture.* Paper presented at the Autonomous Robot Systems and Competitions (ICARSC), 2015 IEEE International Conference on.
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
Broadbent, N. J., Squire, L. R., & Clark, R. E. (2004). Spatial memory, recognition memory, and the hippocampus. *Proceedings of the National Academy of Sciences of the United States of America, 101*(40), 14515-14520. doi: 10.1073/pnas.0406344101
Brooks, R. A. (1991). Inteligence without represention. *Artificial Intelligence, 47*, 139-159.
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.
Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.
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*, 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
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.

Chandler, C. C. (1993). Accessing related events increases retroactive interference in a matching recognition test. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 967–974.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covaria-tion. *Psychological Science, 10*, 360-365.

Clark, A. (1999). An embodied cognitive science? *Trends in cognitive sciences, 3*(9), 345-351.

Clayton, M. S., Yeung, N., & Kadosh, R. C. (2015). The roles of cortical oscillations in sustained attention. *Trends in cognitive sciences, 19*(4), 188-195.

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

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

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

Daoyun, J., & Wilson, Matthew A. (2006). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience, 10*, 100–107.

de Spinoza, B. (1883). *Ethics Part III. On the Origin and Nature of the Emotions*: Library of Alexandria.

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

Dewey, J. (1896). The reflex arc concept in psychology. *Psychological review, 3*(4), 357.

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

Doesburg, S. M., Green, J. J., McDonald, J. J., & Ward, L. M. (2012). Theta modulation of inter-regional gamma synchronization during auditory attention control. *Brain research, 1431*, 77-85.

Dong, D. (2014). *Sensory Motor System: Modeling the Process of Action Execution*. Paper presented at the CogSci 2014, Quebec City, Canada.

Dong, D., & Franklin, S. (2015a). Modeling sensorimotor learning in LIDA using a dynamic learning rate. *Biologically Inspired Cognitive Architectures, 14*, 1-9.

Dong, D., & Franklin, S. (2015b). A New Action Execution Module for the Learning Intelligent Distribution Agent (LIDA): The Sensory Motor System. *Cognitive Computation*. doi: 10.1007/s12559-015-9322-3.

Doran, J. E., Norman, T. J., Franklin, S., & Jennings, N. R. (1997). On Cooperation in Multi-Agent Systems. *The Knowledge Engineering Review, 12*, 309–314.

Drescher, Gary L. (1991). *Made-Up Minds: A Constructivist Approach to Artificial Intelligence*. Cambridge, MA: MIT Press.
Drescher, Gary L. (1998). Learning from Experience Without Prior Knowledge in a Complicated World *Proceedings of the AAAI Symposium on Parallel Models*: AAAI Press.
Dreyfus, H. L. (2009). How representational cognitivism failed and is being replaced by body/world coupling *After Cognitivism* (pp. 39-73): Springer.
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*, 211–245.
Erlhagen, W., & Schoner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review, 109*(3), 545-572. doi: Doi 10.1037//0033-295x.109.3.545
Estes, W. K. (1993). Classification and cognition. *Oxford: Oxford University Press*.
Faghihi, U., Estey, C., McCall, R., & Franklin, S. (2015). A Cognitive Model Fleshes Out Kahneman's Fast and Slow Systems. *Biologically Inspired Cognitive Architectures, 11*, 38-52.
Faghihi, U., McCall, R., & Franklin, S. (2012). A Computational Model of Attentional Learning in a Cognitive Agent. *Biologically Inspired Cognitive Architectures, 2*, 25-36.
Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex, 1*(1), 1-47.
Fillmore, C. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic Theory* (pp. 1–90). New York: Holt, Rinehart and Wilson.
Franklin, S. (1995). *Artificial Minds*. Cambridge MA: MIT Press.
Franklin, S. (2000). Deliberation and Voluntary Action in 'Conscious' Software Agents. *Neural Network World, 10*, 505–521
Franklin, S. (2003). IDA: A Conscious Artifact? *Journal of Consciousness Studies, 10*, 47–66.
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., & Ferkin, Michael H. (2006). An Ontology for Comparative Cognition: a Functional Approach. *Comparative Cognition & Behavior Reviews, 1*, 36–52.
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., Madl, T., D'Mello, Sidney K., & Snaider, J. (2014). LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning. *IEEE Transactions on Autonomous Mental Development, 6*(1), 19-41.

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., McCall, R., & Baars, B. (2013). Conceptual Commitments of the LIDA Model of Cognition. *Journal of Artificial General Intelligence, 4*(2), 1-22. doi: 10.2478/jagi-2013-0002

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, Walter J. (1999). *How Brains Make Up Their Minds*. London: Weidenfeld & Nicolson General.

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.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences, 9*, 474-480.

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

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.

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

Gallagher, S. (2009). Philosophical antecendents to situated cognition.

Gelman, R. S. (1969). Conservation acquisition: a problem of learning to attend to relevant attributes. *Journal of Experimental Child Psychology, 7*, 167–187.

Gibson, James J. (1979). *The Ecological Approach to Visual Perception*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci., 15*(1))20-25.

Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature, 338*(6213), 334-337.

Haist, F., Gore, J. B., & Mao, H. (2001). .Consolidation of human memory over decades revealed by functional magnetic resonance imaging. *Nature neuroscience, 4*(11), 1139-1145.

Häkkinen, M. (2010). *Why Alarms Fail: A Cognitive Explanatory Model.* Ph.D Dissertation, University of Jyväskylä.

Hawkins, J., Ahmad, S., & Dubinsky, D. (2011). Hierarchical Temporal Memory including HTM Cortical Learning Algorithms (Vol. 32), from http://numenta.org/

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *arXiv preprint arXiv:1502.01852*.

Hernes, M. (2014). A Cognitive Integrated Management Support System for Enterprises *Computational Collective Intelligence. Technologies and Applications* (pp. 252-261): Springer.

Hintzman, D. L. (2011). Research Strategy in the Study of Memory: Fads, Fallacies, and the Search for the "Coordinates of Truth". *Perspectives on Psychological Science, 6*(3), 253-271. doi: 10.1177/1745691611406924

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.

Hohman, T. J., Peynirciofülu, Z. F., & Beason-Held, L. L. (2012). Flexibility of event boundaries in autobiographical memory. *Memory*, 1-12. doi: 10.1080/09658211.2012.725737

Holz, E. M., Glennon, M., Prendergast, K., & Sauseng, P. (2010). Theta-Gamma Phase Synchronization during Memory Matching in VisualWorking Memory. *NeuroImage*. doi: 10.1016/j.neuroimage.2010.04.003

Jackson, J. V. (1987). Idea for a Mind. *Siggart, Newsletter, 181*, 23–26.

James, W. (1890). *The Principles of Psychology*. Cambridge, MA: Harvard University Press.

Janson, J., & Kranczioch, C. (2011). Good vibrations, bad vibrations: Oscillatory brain activity in the attentional blink. *Advances in Cognitive Psychology, 7*, 92-107.

Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *TRENDS in Cognitive Sciences, 11*(7), 267-269.

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

Johnston, Victor 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.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kaur, S. (2011). *WHEN TO PAY ATTENTION?: ASYNCHRONY REQUIRES A TRIGGER.* The University of Memphis.

Kelso, J. A. S. (1995). *Dynamic Patterns: The Self Organization of Brain and Behavior*. Cambridge, MA: MIT Press.

Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep Supervised, but Not Unsupervised, Models May Explain IT Cortical Representation. *PLoS Computational Biology*(10(11): e1003915).

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci, 27*(12), 712-719. doi: 10.1016/j.tins.2004.10.007

Koenig, N., & Howard, A. (2004). *Design and use paradigms for gazebo, an open-source multi-robot simulator.* Paper presented at the Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on.

Kondadadi, R., & Franklin, S. (2001). A Framework of Deliberative Decision Making in "Conscious" Software Agents *Proceedings Of Sixth International Symposium on Artificial Life and Robotics (AROB-01)* (pp. 280–283).

Kruschke, J. K. (2010). Models of attentional learning. *In: E. M. Pothos and A. Wills (eds.), Models of Learning and Categorization, Cambridge University Press*.

Landau, A. N., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current biology, 22*(11), 1000-1004.

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

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23*, 155-184.

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

Liddell, B. J., Brown, K. J., Kemp, A. H., Barton, M. J., Das, P., Peduto, A., . . . Williams, L. M. (2005). A direct brainstem–amygdala–cortical dalarmT system for subliminal signals of fear. *NeuroImage 24*, 235- 243.

Lisman, J., & (2005), H. (2005). The Theta/Gamma Discrete Phase Code Occuring During the Hippocampal Phase Precession May be a More General Brain Coding Scheme. *HIPPOCAMPUS, 15*, 913–922.

Lisman, J., & Buzsaki, G. (2008). A neural coding scheme formed by the combined function of gamma and theta oscillations. [Research Support, N.I.H., Extramural

Review]. *Schizophrenia bulletin, 34*(5), 974-980. doi: 10.1093/schbul/sbn060

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron, 77*(6), 1002-1016.

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.

Madl, T., & Franklin, S. (2015). Constrained Incrementalist Moral Decision Making for a Biologically Inspired Cognitive Architecture. In R. Trappl (Ed.), *A Construction Manual for Robots' Ethical Systems*: Springer.

Madl, T., Franklin, S., Chen, K., Montaldi, D., & Trappl, R. (2014). Bayesian Integration of Information in Hippocampal Place Cells. *PLoS ONE*(9(3) e89762).

Madl, T., Franklin, S., Chen, K., Montaldi, D., & Trappl, R. (to appear). Towards real-world capable spatial memory in the LIDA cognitive architecture. *Biologically Inspired Cognitive Architectures*.

Madl, T., Franklin, S., Chen, K., & Trappl, R. (2013). Spatial Working Memory in the LIDA Cognitive Architecture. In R. West & T. Stewart (Eds.), *Proceedings of* *the 12th International Conference on Cognitive Modelling* (pp. 384-390). Ottawa, Canada: Carleton University.

Madl, T., Franklin, S., Snaider, J., & Faghihi, U. (2015). Continuity and the Flow of Time—A Cognitive Science Perspective. In B. Mölder, V. Arstila & P. Øhrstrøm (Eds.), *Philosophy and Psychology of Time*: Springer.

Marsella, S., Gratch, J., & others. (2010). *Computational Models of Emotion. Blueprint for Affective Computing. KR Scherer, T. Banziger and E. Roesch*: Oxford University Press.

Mateo, J. M. (2010). Alarm calls elicit predator-specific physiological responses. *Biol. Lett, 6*, 623–625.

Mayes, Andrew R., & Roberts, N. (2002). Theories of episodic memory. In A. Baddeley, M. Conway & J. Aggleton (Eds.), *Episodic Memory:New directions in research* (pp. 86–109). Oxford: Oxford University Press.

McCall, R. (2014). *Fundamental Motivation and Perception for a Systems-Level Cognitive Architecture.* PhD Thesis, University of Memphis, Memphis, TN USA.

McCall, R., Franklin, S., & Friedlander, D. (2010). *Grounded Event-Based and Modal Representations for Objects, Relations, Beliefs, Etc.* Paper presented at the FLAIRS-23, Daytona Beach, FL.

McCall, R., Snaider, J., & Franklin, S. (2010). Sensory and Perceptual Scene Representation. *Journal of Cognitive Systems Research*.

McCauley, L., & Franklin, S. (2002). A large-scale multi-agent system for navy personnel distribution. *Connection Science, 14*(4), 371-385.

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

Miller, S. M., & Fu, W.-T. (2007). The role of temporal sequence learning in guiding visual attention allocation. *Proceedings of the human factors and ergonomics society 51st annual meeting*, 1368-1372.

Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia, 46*(3), 774-785.

Minsky, M. (1985). *The Society of Mind*. New York: Simon and Schuster.

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

Nakatani, C., Raffone, A., & van Leeuwen, C. (2014). Efficiency of conscious access improves with coupling of slow and fast neural oscillations. *Journal of cognitive neuroscience, 26*(5), 1168-1179.

Negatu, A., D'Mello, Sidney K., & Franklin, S. (2007). Cognitively Inspired Anticipation and Anticipatory Learning Mechanisms for Autonomous Agents. In M. V. Butz, O. Sigaud, G. Pezzulo & G. O. Baldassarre (Eds.), *Proceedings of the Third Workshop on Anticipatory Behavior in Adaptive Learning Systems (ABiALS 2006)* (pp. 108-127). Rome, Italy: Springer Verlag.

Negatu, A., & Franklin, S. (2002). An action selection mechanism for 'conscious' software agents. *Cognitive Science Quarterly, 2*(special issue on "Desires, goals, intentions, and values: Computational architectures." Guest editors Maria Miceli and Cristiano Castelfranchi.), 363–386.

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: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.
Ogawa, H., & Yagi, A. (2002). The Implicit Processing in Multiple Object Tracking. *Technical Report on Attention and Cognition, 10*, 1-4.
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.
Oyama, S. (2000). The ontogeny of information: Developmental systems and evolution (science and cultural theory).
Panksepp, J. (2005). Affective consciousness: Core emotional feelings in animals and humans. *Consciousness and Cognition, 14*, 30–80.
Posner, M. I., & Fan, J. (2004). ATTENTION AS AN ORGAN SYSTEM. In J. R. Pomerantz & M. C. Crair (Eds.), *Topics in Integrative Neuroscience: From Cells to Cognition*. Cambridge UK: Cambridge University Press.
Quillian, M. R. (1966). *Semantic Memory*. Cambridge, Mass.: Bolt Beranek and Newman, 1966.
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.
Rolls, E. T. (2000). Neurophysiology and functions of the primate amygdala, and the neural basis of emotion. In The Amygdala: a Functional Analysis (ed. Aggleton, J. P.). *Oxford Univ. Press, Oxford, UK*, 447-478.
Rosenbaum, D. A. (2014). *It's a Jungle in There: How Competition and Cooperation in the Brain Shape the Mind*: Oxford University Press.
Rugg, Michael D., & Yonelinas, Andrew P. (2003). Human recognition memory: a cognitive neuroscience perspective. *Trends in Cognitive Science, 7*, 313–319.
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.
Schacter, D. L., & Tulving, E. (1994). *Memory systems*. Cambridge, MA: MIT Press.
Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. *Appraisal processes in emotion: Theory, methods, research, 92*, 120-120.
Schmidt, T. (2002). The finger in flight: Real-time motor control by visually masked color stimuli. *Psychological Science, 13*(2), 112-118.
Schöner, G. (2008). Dynamical systems approaches to cognition. In J. P. Spencer, M. S. Thomas & J. L. McClelland (Eds.), *Toward a Unified Theory of Development: Connectionism and Dynamic Systems Theory Re-Considered*. New York: Oxford University Press.
Sloman, A. (1998). Damasio, Descartes, Alarms and Meta-management *Proceedings Symposiumon Cognitive Agents: Modeling Human Cognition*. San Diego: IEEE.
Sloman, A. (1999). What Sort of Architecture is Required for a Human-like Agent? In M. Wooldridge & A. S. Rao (Eds.), *Foundations of Rational Agency* (pp. 35–52). Dordrecht, Netherlands: Kluwer Academic Publishers.
Sloman, A. (2001). Beyond shallow models of emotion. *Cognitive Processing, 2*(1), 177-198.

Sloman, A., & Chrisley, R. (2003). Virtual Machines and Consciousness. *Journal of Consciousness Studies, 10*, 133–172

Snaider, J., & Franklin, S. (2011). *Extended Sparse Distributed Memory.* Paper presented at the Biological Inspired Cognitive Architectures 2011, Washington D.C. USA.

Snaider, J., & Franklin, S. (2014a). Modular composite representation. *Cognitive Computation, 6*(3), 510-527.

Snaider, J., & Franklin, S. (2014b). Vector LIDA. *Procedia Computer Science, 41*, 188-203.

Snaider, J., McCall, R., & Franklin, S. (2010). *The Immediate Present Train Model Time Production and Representation for Cognitive Agents*. Paper presented at the AAAI Spring Symposium on "It's All In the Timing", Palo Alto, CA.

Snaider, J., McCall, R., & Franklin, S. (2011). *The LIDA Framework as a General Tool for AGI*. Paper presented at the Artificial General Intelligence (AGI-11), Mountain View, CA.

Snaider, J., McCall, R., & Franklin, S. (2012). Time production and representation in a conceptual and computational cognitive model. *Cognitive Systems Research, 13*(1), 59-71.

Sowa, J. F. (1991). Toward the expressive power of natural language. *Principles of Semantic Networks*, 157-189.

Sowa, J. F. (2014). *Principles of Semantic Networks: Explorations in the representation of knowledge*: Morgan Kaufmann.

Squire, L. R., & Kandel, E. R. (2000). *Memory: From Mind to Molecules*: W. H. Freeman.

Steriade, M. (2006). Grouping of brain rhythms in corticothalamic systems. *Neuroscience, 137*(4), 1087-1106. doi: 10.1016/j.neuroscience.2005.10.029

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

Strain, S., & Franklin, S. (2011). Modeling medical diagnosis using a comprehensive cognitive architecture. *Journal of Healthcare Engineering, 2*(2), 241-257.

Strain, S., Kugele, S., & Franklin, S. (2014). *The Learning Intelligent Distribution Agent (LIDA) and Medical Agent X (MAX): Computational Intelligence for Medical Diagnosis*. Paper presented at the IEEE Symposium Series on Computational Intelligence (SSCI), Symposium on Computational Intelligence for Human-like Intelligence (CIHLI).

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

Strogatz, S. H. (2014). *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*: Westview press.

Taatgen, N. A., Juvina, I., Schipper, M., Borst, J. P., & Martens, S. (2009). Too much control can hurt: A threaded cognition model of the attentional blink. *Cogn Psychol*. doi: S0010-0285(09)00003-6 [pii]

10.1016/j.cogpsych.2008.12.002

Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *[Frontiers in Bioscience, 14*, 321-332.

Thelen, E., & Smith, L. (1994). *A Dynamical Systems Approach to the Development of Cognition and Action*. Cambridge MA: MITPress.

Tulving, E. (1983). *Elements of episodic memory*. Oxford: Clarendon Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1-25.

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: Role of the hippocampus. *HIPPOCAMPUS, 8*, 198–204.

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 Gelder, T. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences, 21*(5), 615-628.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.

Vidnyánszky, Z., & Sohn, W. (2003). Attentional learning: learning to bias sensory competition. *Journal of Vision, 3*(174a).

Von Uexküll, J., & Mackinnon, D. L. (1926). *Theoretical biology*: K. Paul, Trench, Trubner & Company Limited.

Voytek, B., D'Esposito, M., Crone, N., & Knight, R. T. (2013). A method for event-related phase/amplitude coupling. *Neuroimage, 64*, 416-424.

Voytek, B., Kayser, A. S., Badre, D., Fegen, D., Chang, E. F., Crone, N. E., . . . D'Esposito, M. (2015). Oscillatory dynamics coordinating human frontal networks in support of goal maintenance. *Nature neuroscience*.

Walker, M. A. T. T. H. E. W. P., Brakefield, T. I. F. F. A. N. Y., Hobson, J. A. L. L. A. N., & Stickgold, R. O. B. E. R. T. (2003). Dissociable stages of human memory consolidation and reconsolidation. *Nature, 425*, 616–620.

Wallach, W., Franklin, S., & Allen, C. (2010). A Conceptual and Computational Model of Moral Decision Making in Human and Artificial Agents. In W. Wallach & S. Franklin (Eds.), *Topics in Cognitive Science, special issue on Cognitive Based Theories of Moral Decision Making* (pp. 454-485): Cognitive Science Society.

Watson, R. A., Mills, R., & Buckley, C. L. (2011). Global Adaptation in Networks of Selfish Components: Emergent Associative Memory at the System Scale. *Artificial Life*. doi: 10.1162/artl_a_00029

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.

Yoshida, H., & Smith, L. B. (2003). Known and novel noun extensions: Attention at two levels of abstraction. *Child Development, 76*, 564-577.

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction Error Associated with the Perceptual Segmentation of Naturalistic Events. *Journal of Cognitive Neuroscience*. doi: 10.1162/jocn_a_00078

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event Perception: A Mind–Brain Perspective. *Psychological Bulletin, 133*(2), 273–293.

February 22, 2016

Dear Alexei,

This is the LIDA Tutorial article we've been corresponding about. I wish to submit it for publication in the BICA Journal.

Very truly yours,

Stan Franklin