Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 20 (2007) 955–961
Neural Networks
www.elsevier.com/locate/neunet
2007 Special Issue

An architectural model of conscious and unconscious brain functions: Global Workspace Theory and IDA

Bernard J. Baars[a,b,*], Stan Franklin[c]

[a] The Neurosciences Institute, San Diego, CA 92121, USA
[b] 3616 Chestnut St., Apt. 3, Lafayette, CA 94549, USA
[c] Computer Science Department & Institute for Intelligent Systems, University of Memphis, Memphis, TN, USA

Abstract

While neural net models have been developed to a high degree of sophistication, they have some drawbacks at a more integrative, "architectural" level of analysis. We describe a "hybrid" cognitive architecture that is implementable in neuronal nets, and which has uniform brainlike features, including activation-passing and highly distributed "codelets," implementable as small-scale neural nets. Empirically, this cognitive architecture accounts qualitatively for the data described by Baars' Global Workspace Theory (GWT), and Franklin's LIDA architecture, including state-of-the-art models of conscious contents in action-planning, Baddeley-style Working Memory, and working models of episodic and semantic longterm memory. These terms are defined both conceptually and empirically for the current theoretical domain. The resulting architecture meets four desirable goals for a unified theory of cognition: practical workability, autonomous agency, a plausible role for conscious cognition, and translatability into plausible neural terms. It also generates testable predictions, both empirical and computational.

Keywords: Cognitive architecture; Global workspace theory; Conscious cognition; LIDA architecture

1. Introduction

While neural net models have been developed to a high degree of sophistication, they have some drawbacks at a more integrative, "architectural" level of analysis. We describe a "hybrid" cognitive architecture that is implementable in neuronal nets, and which has uniform brainlike features, including activation-passing and highly distributed "codelets," implementable as small-scale neural nets. This cognitive architecture integrates Global Workspace Theory, as developed by Baars (1988, 2002), and Franklin et al.'s LIDA model (Ramamurthy, Baars, D'Mello, & Franklin, 2006). Together, they have theoretical and empirical implications for our understanding of both conscious and unconscious human cognition.

It is not possible to test all parameters of cognitive architectures, which perform real cognitive activities that humans perform; actual working models are far too complex to be exhaustively tested by simple experiments. Models designed to exclusively reproduce experimental data are often too limited to accomplish real-world tasks. Thus cognitive modellers are caught between cognitive architectures that perform real tasks but which cannot be fully tested experimentally, and experiment-based models that cannot fully perform the tasks human beings routinely do. We suggest therefore that *workability* must be combined with experimental evidence as desirable constraints on cognitive models.

1.1. Autonomous agency

To accomplish real-world tasks, any cognitive architecture must be capable of "living" in an environment, that is, capable of sensing that environment and acting on it in meaningful ways, so as to accomplish significant life goals. A working model of cognition (WMC) must be an *autonomous agent,* in that it must sense and act in pursuit of its own agenda. The IDA implementation of Global Workspace Theory, which we discuss here, is such an autonomous agent. An *autonomous agent* is a system situated within and part of an environment

* Corresponding author at: The Neurosciences Institute, San Diego, CA 92121, USA.
E-mail address: baars@nsi.edu (B.J. Baars).
URL: http://www.nsi.edu (B.J. Baars).

0893-6080/$ - see front matter © 2007 Published by Elsevier Ltd
doi:10.1016/j.neunet.2007.09.013

that senses its environment and acts on it, over time, in pursuit of its own agenda. Its actions must affect what it will sense in the future, in much the way that turning one's head will change one's visual experience (Franklin & Graesser, 1997). Thus the agent is structurally coupled to its environment in the sense of Maturana and Varela (Maturana, 1975; Maturana & Varela, 1980; Varela, Thompson, & Rosch, 1991). Biological examples include humans and animals in their evolutionary niches. Artificial examples include autonomous robots and autonomous software agents, as well as computer viruses. The computational IDA model is an autonomous software agent as defined above (Franklin, 2000).

The basic motivations, the *agendas*, of biological agents, are "built-in" by evolution and modified by development. In the case of artificial agents, basic motivations are built in. Once an artificial autonomous agent is let loose in its environment, its agenda is its own and no longer influenced by the designer.

Based on two decades of development we believe that the issue of conscious cognition, as an essential component of the cognitive architecture, can be integrated into such working autonomous agents.

1.2. Consciousness as a explanandum for cognitive theory

The empirical evidence for a central role for conscious cognition in the human brain is now difficult to dispute (e.g. Baars (2002), Dehaene (2002)). While we have some viable conceptions of consciousness as a scientific construct, we do not at this time know necessary and sufficient conditions for conscious cognition in the brain. Could such conditions as cholinergic neuromodulation, spike coding and the like, be essential? The answer is simply unknown at this time. Therefore, we define here an intermediate concept of a *functionally conscious agent* as an autonomous agent which implements a Global Workspace Theory (GWT) for the role of consciousness in a WMC, to facilitate the performance of real-world tasks.

Note that we make no claim that a functionally conscious agent such as IDA is phenomenally conscious, that it has subjective experience (Franklin, 2003). Rather, we claim that the non-phenomenal consciousness mechanism in a functionally conscious agent can perform the known functions of consciousness in humans. Within a single cognitive cycle (Fig. 4) consciousness functions to filter the attention paid to the agent's internal model of its world, and to select contents to be learned. Higher-level cognitive processes requiring multiple cycles employ consciousness for a variety of other functions including volitional decision making, analogy forming, self-monitoring, etc. (Baars, 1988, p. 349).

1.3. Neural translatability

Neural translatability is a major goal. Since biological minds are implemented in brains, we must be able to plausibly translate an architecture, such as IDA's, that claims to model minds, into neural terms. IDA's computational components operate by activation-passing, a biologically plausible mechanism. In addition, one might argue that the nodes and links (described below) of IDA components like Perceptual Associative Memory may be translatable, with PAM nodes corresponding to cell assemblies and their links to neuronal interactions between them. This type of translation may be too simplistic. Rather, following Freeman (1999), we suggest that nodes in perceptual memory may correspond to attractors in the state spaces of cell assemblies, and their links to the dynamics operating over the interactions between them. To use the Skarda and Freeman (1987) example, a rabbit recognizes the odour of a fox when the trajectory of its olfactory bulb activity state following a sniff falls into the "fox attractor". The fox node in PAM would correspond to the fox attractor. This suggests a general approach to neural translatability between hybrid models like IDA and brain dynamics.

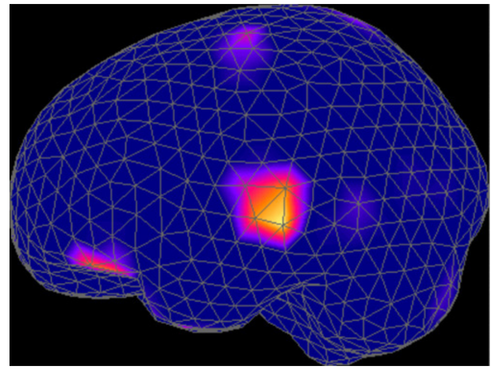

Fig. 1. Rapid "broadcasting" of activation for a single auditory word. This MEG recording of a single, unexpected auditory word shows widespread brain activity beginning 150 ms post-stimulus. In less than a second, activity is visible in the auditory region of the superior temporal lobe (A1), but also in orbitofrontal regions, which are not required for auditory analysis. In addition, activation can be seen in the sensorimotor strip and even in visual regions (occipital). Note that the MEG signal is projected onto a mathematically inflated left hemisphere. Bright yellow signals highest MEG activation; light blue is lowest. From Pulvermueller, Shtyrov, and Ilmoniemi (2003).

1.4. Brain and behavioural evidence

But how is direct brain evidence brought to bear on such complex models? Fortunately we can point to some distinctive predictions that have been supported. Both GWT and IDA make testable predictions, as listed below.

1. The Global Access Hypothesis: conscious contents evoke widespread brain activation, as proposed by Global Workspace Theory;

There is now a sizable body of evidence to support the GWT hypothesis that conscious, but not unconscious, brain events evoke widespread cortical activity related to the reportable content[1] (Baars, 2002; Dehaene, 2002) (See Fig. 1). This

[1] "Accurate reportability" of conscious contents is a standard operational definition for conscious events in empirical studies of human conscious cognition. Unconscious comparison conditions are those brain processes that are believed to exist but which are not accurately reportable. Animal analogues of "accurate report" are often used in macaque monkeys, for example, by having the animal match stimuli from memory in a "match to sample" task.

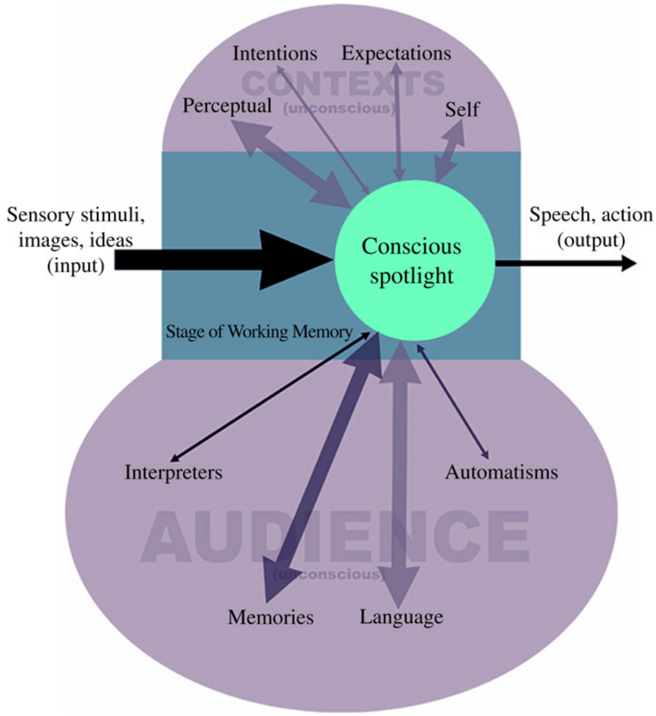

Fig. 2. Global Workspace Theory (GWT). A theatre metaphor for Global Workspace Theory.

prediction was initially made in 1983, and is not suggested by any other theory of which we know.[2] There is quite extensive current debate about the evidence regarding this hypothesis in the cognitive neuroscience literature (Tse, Martinez-Conde, Schlegel, & Macknik, 2005).

2. The Working Memory Hypothesis (conscious contents recruit unconscious WM functions needed for verbal rehearsal, visual semantics, and executive functions) (Figs. 1 and 2);

GWT makes other novel predictions. For example, it suggests that classical Working Memory (WM) may involve distributed specialized systems, including language components, long-term memory, visuospatial knowledge and the like, which are recruited by the conscious components of WM tasks. Current brain evidence strongly suggests that the specialized components of WM are highly distributed in the cortex and subcortical structures like the basal ganglia. Most of these functions are unconscious in their details, but they generally have briefly conscious components. It is noteworthy, therefore, that all the classical "boxes" of Alan Baddeley's WM models have a conscious component—including conscious perception of input, conscious access to verbal rehearsal, and conscious decisions regarding verbal report. The most recent version of Baddeley's WM has a new conscious component, called the Episodic Buffer (Baddeley, 2000). However, it does not have a central role in recruiting linguistic, visuospatial and executive functions; the current concept of the Episodic Buffer is only the front end of long-term episodic memory. GWT suggests a more active view of the conscious aspects of human cognition. It is the consciously evoked "broadcast" that serves to mobilize and guide the many unconscious knowledge domains that enable Working Memory functions like inner speech, visual problem solving and executive control (Fig. 2).

3. The Conscious Learning Hypothesis (all significant learning is evoked by conscious contents, but the learning process itself and its outcomes may be unconscious).

The theoretical reason for this claim is that learning novel information requires a novel integration of existing knowledge with unpredictable input. Thus GWT provides a principled prediction for the role of consciousness in learning. It is noteworthy, in this respect, that after five decades of attempts to prove learning without consciousness, most findings show typically small effect sizes, at very brief time intervals, using highly predictable stimuli such as emotional facial expressions (Snodgrass & Shevrin, 2006). More demanding learning tasks almost always have a clear conscious component,[3] and there is a clear "dose-response" function between the degree of conscious exposure and the amount of learning that results.[4] This is indeed what was historically called the Law of Effect, which should perhaps be called the Law of *Conscious* Effect. The "conscious" aspect of learning, which was taken for granted before the behavioristic revolution, has now become largely forgotten. Nevertheless, the evidence continues to show a clear monotonic relationship between conscious study time and learning.

We now describe these two theoretical domains, Global Workspace Theory and LIDA.

1.5. Global workspace theory

Global workspace theory aims to specify the role of conscious brain events in cognition (Baars, 1983, 1988, 1997).

A theatre metaphor for GWT is a useful first approximation. Unconscious processors in the theatre audience receive broadcasts from a conscious "bright spot" on the stage. Control of the bright spot corresponds to selective attention. Backstage, unconscious contextual systems operate to shape and direct conscious contents. GWT is a rigorously developed set of testable hypotheses, and the theatre metaphor is only a convenient reminder of its basic features (Baars, 1988, 2002).

GWT was developed based on robust evidence regarding conscious processes, combined with the artificial intelligence

[2] GWT also converges well with the work of Chein and Schneider (2005), whose "net of nets" architecture is based on experimental studies of skills that are novel vs. practiced (and therefore less conscious). Practiced, predictable skills show a marked reduction in cortical activity (Schneider & Shiffrin, 1977). It is interesting that the resulting network architecture bears a striking resemblance to GWT.

[3] Implicit learning allows behavior that can be described as rule-directed to be learned from conscious experience without the subject being able to articulate the rule. However, all studies of implicit learning make use of conscious events to evoke implicit learning processes.

[4] Recent evidence indicates more robust learning effects for emotional stimuli, such as emotional facial expressions. Such biologically relevant inputs can be treated as single chunks in GWT, which do not require the recruitment of novel knowledge sources that require consciousness to be integrated.

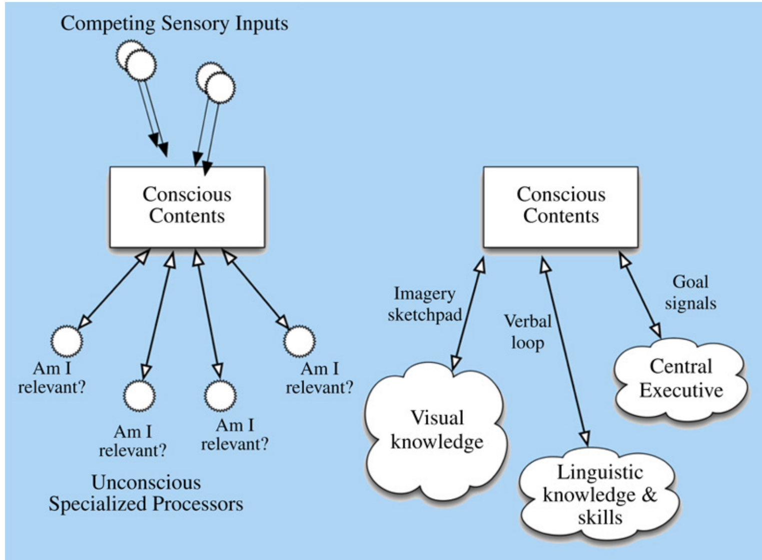

Fig. 3. Global Workspace Theory and working memory. Global Workspace Theory integrates conscious contents with unconscious distributed expertise in the brain. Notice the radically distributed nature of the architecture, with the exception of functions supported by consciousness, including action planning. On the left is a simplified GWT model from Shanahan (2006). Notice that the small white circles, representing unconscious processors, constantly search conscious for (globally distributed) messages that are relevant to them, somewhat like humans listening for air flight announcements. The right side shows how GWT suggests that cognitive Working Memory may be mobilized by brief conscious access to perceptual input, rehearsed words or digits, output decisions, and other conscious events. Notice that WM components like verbal rehearsal have both conscious and unconscious aspects. The details of language, perceptual processing and storage are handled "offline" by unconscious distributed processors. Only the contents of the GW need to be conscious, as assessed by accurate reportability (Baars, 1988, 2002). (With thanks to Murray Shanahan.)

concept of a "blackboard architecture" that combined multiple sources of knowledge in order to identify an acoustical signal in a complex, noisy, and ambiguous environment (Hayes-Roth & Lesser, 1977). Such noisy and ambiguous signals are routine in human perception, thought, and motor planning and control. Based on a large body of experiments comparing conscious and unconscious events, empirical generalizations were proposed (Baars, 1988). These indicated that conscious events were strongly associated with (a) limited-capacity processes[5] showing (b) internal consistency of conscious contents, and (c) low computational efficiency. Thus skilled speakers cannot consciously label the syntax of a sentence, even though they constantly use the results of unconscious syntactic analysis. In contrast to conscious contents, unconscious events showed (a') much larger capacity limits,[6] (b') with no internal consistency constraint, and (c') often with great computational efficiency. This is a puzzling pattern if one thinks of consciousness in terms of algorithms or neural network outputs, where efficiency is at a premium. Rather, these aspects of conscious cognition would seem to endanger survival in a world in which fast, accurate decision-making is essential. This puzzle makes sense, however, if consciousness is viewed as "fame in the brain" (Dennett, 2005)—the ability to recruit numerous unconscious knowledge sources, which can respond in a distributed fashion to focal conscious events (Figs. 1 and 3).

The first global workspace (GW) architecture was developed by Alan Newell and coworkers to identify spoken words in a noisy acoustical space (Nii, 1986). This is a very difficult challenge, because rooms add echoes and background noises to an already underspecified vocal signal, rendering standard algorithmic pattern recognition largely ineffective. The GW architecture of Newell et al., called *Hearsay*, helped to resolve this challenge, just as humans must in order to survive and reproduce. GW systems therefore showed a major functional advantage, to compensate for the puzzling drawbacks listed above.

GWT postulates that human cognition is implemented by a multitude of relatively small, special purpose processes, almost always unconscious. Although that may seem commonplace today, the idea of widely distributed specialized processing in the brain was highly controversial at the time it was proposed. Processing coalitions compete for access to a global workspace (and subjectively into consciousness, assessed behaviourally by accurate reports). This limited capacity global workspace serves to broadcast the message from the winning coalition to all the unconscious processors, in order to recruit resources to join in handling novel and high-priority input, and in solving current problems. Consciousness in this view allows us to deal with novel or challenging situations that cannot be dealt with efficiently, or at all, by local, routine unconscious processes. Conscious cognition solves the "relevance problem" encountered in artificial intelligence and robotics, by enabling access to unpredictable but necessary knowledge sources. As a default, consciousness serves a lookout function to spot *potential* dangers or opportunities, so that there is a particularly close relationship between conscious content and

[5] There is empirical evidence for capacity limits in conscious perception, selective attention, immediate memory, and voluntary control (Cowan, 2001).

[6] For example, long-term declarative memory, or the vocabulary of natural language.

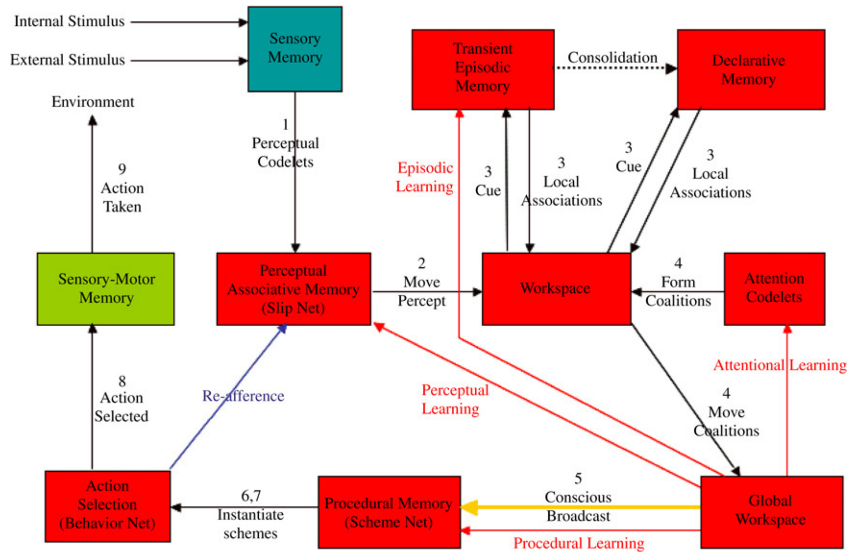

Fig. 4. The IDA cognitive cycle.

significant sensory input. The external senses can be simulated endogenously by way of conscious visual imagery, inner speech, and inner sensorimotor practice. These endogenous "senses" have been shown to mobilize regions of cortex and subcortex that become active with similar exogenous events.

Conscious *contents* are always guided and constrained by unconscious *contexts*: goal contexts, perceptual contexts, conceptual contexts and shared cultural contexts. Each context is itself a coalition of processes. Though contexts are unconscious, they shape conscious processes. For example, unconscious spatial knowledge is required to interpret the orientation of conscious visual objects. In GWT, widespread learning activity is evoked by conscious contents. Implicit learning occurs when unconscious processors are cued by conscious contents to perform problem-solving without reportable rule generation. Tasks like language learning are largely implicit, but they are primarily evoked by conscious input. (See Baars (1988, 2002), for further details.)

2. The IDA cognitive architecture

IDA (Intelligent Distribution Agent) is an intelligent software agent (Franklin & Graesser, 1997) developed for the US Navy to perform tasks that previously required trained human experts (Franklin, 2001). At the end of each sailor's tour of duty, he or she is assigned to a new billet in a way that takes into account the sailor's preferences, the requirements of the Navy, and a host of regulations. This assignment process is called *distribution*. The Navy employs some 300 people, called detailers, on a full-time basis to make these new assignments. IDA completely automates the role of the human detailer. The IDA software agent is up and running, and has matched the performance of Navy detailers.

The IDA model includes a computational component, i.e. one that is the currently implemented and running, and a broader conceptual IDA, some of which is yet to be implemented. In addition to incorporating GWT as a functional equivalent of conscious cognition, IDA implements a number of other psychological theories (Baddeley, 1993; Barsalou, 1999; Conway, 2001; Ericsson & Kintsch, 1995; Glenberg, 1997; Gray, 2002; Sloman, 1999). Autonomous agents in complex, dynamic environments frequently and iteratively sample their environments and act on them. Much of human cognition operates by rapid interactions between conscious contents, the various memory systems and decision-making. We call these rapid routines *cognitive cycles* (Baars & Franklin, 2003; Franklin, Baars, Ramamurthy, & Ventura, 2005). Findings consistent with such cognitive cycles have been reported by neuroscientists (Freeman, 2003; Fuster, Bodner, & Kroger, 2000; Halgren, Boujon, Clarke, Wang, & Chauvel, 2002; Lehmann, Strik, Henggeler, Koenig, & Koukkou, 1998). While cognitive cycles can overlap, producing parallel actions, their conscious components can only occur serially. The IDA model suggests therefore that conscious cognition occurs as a sequence of discrete, coherent episodes separated by quite short periods of no conscious content (see also VanRullen and Koch (2003)).

The IDA model is a major step toward finer-grained computational modelling of GWT, including the vital question of how to specify the specialized knowledge sources that GWT must treat merely as black boxes. A domain-independent version of IDA is now being developed to handle a wider range of tasks. IDA aims to be a true autonomous agent with a capacity corresponding to conscious contents in humans.

The IDA cognitive cycle can be divided into nine steps (see Fig. 4).

1. Incoming sensory stimuli is filtered through preconscious perception, where meaning is added and a percept produced.
2. The current percept moves to preconscious working memory,[7] where it participates, along with un-decayed

[7] Note that the lower-case spelling of "working memory" refers to unconscious memory buffers unrelated to the standard Working Memory model of cognition proposed by Baddeley and colleagues.

percepts from previous cycles, in the structure building of higher-level perception.
3. The current structure from working memory cues transient episodic memory and declarative memory, producing *local associations,* which are stored in long-term working memory.
4. Coalitions of the contents of long-term working memory compete for consciousness, thus making it possible to recruit system resources for the most relevant, urgent, and important task components.
5. The conscious broadcast (a la Gobal Workspace Theory) occurs, enabling various forms of learning and the recruitment of internal resources. The broadcast is hypothesized to require approximately 100 ms, based on a number of empirical sources.
6. Procedural memory responds to the contents of the conscious broadcast.
7. Other responding (unconscious) schemes instantiate copies of themselves in the action selection mechanism, bind variables, and pass activation.
8. The action selection mechanism chooses an action for this cognitive cycle.
9. IDA then acts on her internal and/or external environment.

2.1. *The cognitive cycle hypothesis*

Most higher-level cognitive processes are built on multiple perception-action cycles as iterative "atoms". The IDA cognitive cycle predicts that frequent (5–10 Hz) selection takes place in the following domains: perceptual associative memory, preconscious working memory, transient episodic memory, procedural memory, action selection—all by way of consciousness. Sensory data, past events, and possible responses are all selected for importance, urgency, insistence, and relevancy. The IDA cognitive cycle can be thought of as a rapidly iterating sequence of selective operations designed to turn incoming sensory data into actions on the world. Recent evidence indicates that episodic retrieval in the human brain involves theta-wave synchrony between hippocampus and neocortex; such oscillations are in the correct temporal domain for 5–10 Hz cycles (Jensen & Tesche, 2002). This rapid iteration of cognitive cycles allows for timesharing to produce multitasking such as visually tracking the movements of several objects at once. The IDA model suggests additional testable predictions. Among these is the selective character of cognition.

While IDA has a number of symbolic features, it is brainlike in using activation-passing throughout.[8] Thus, it can be thought of as a hybrid symbolic/connectionist system. In particular, the implementation of perceptual associative memory (PAM) is modelled after the slipnet in the Copycat architecture (Hofstadter & Mitchell, 1995). Nodes and links of PAM that are over threshold are selected as part of the percept (please see Step 2 above of the IDA cognitive cycle), and serve as a common representational currency throughout the rest of the agent's architecture.

The original IDA software agent was entirely hand crafted, and was not designed to learn. An extension of IDA to a software agent LIDA ("Learning IDA") that does learn is now in progress (Franklin & Patterson, 2006; Ramamurthy et al., 2006). The LIDA model fleshes out the GWT prediction about learning by postulating four types of learning during each rapid cognitive cycle (perceptual learning, episodic learning, procedural learning and attentional learning). Each of these procedures is initiated by the contents of consciousness during the cycle, and is entirely automatic (unconscious) from that point onward. As a consequence, to attend (consciously) is to learn. Volitional learning and more complex problem solving, which occur over multiple cycles, are therefore a matter of directing the automatic learning processes.

3. Summary and conclusion

In recent decades progress has been made in modelling human cognition. Most current models do not, however, generally meet four desirable standards: They would not perform functions attributed to them in the challenging environment of the real world; they are not autonomous agents, like humans and animals living in their natural environments; they rarely deal explicitly with the challenge of conscious cognition; they are not readily translatable into plausible neural terms. The combined approach of Global Workspace Theory and the IDA (LIDA) symbolic/connectionist model aims to address these challenges, while generating distinctive and testable hypotheses. One major prediction that has found experimental support is the very widespread cortical activation observed for matched conscious vs. unconscious brain events. This result has now been found in a large number of studies (e.g. Baars (2002), Dehaene (2002)). We also see promising links with the strong empirical tradition of Working Memory research (Baars & Franklin, 2003), and with renewed interest in brain rhythms in the near-10 Hz range, postulated to be the iteration time of the IDA cognitive cycle (Freeman, 2003; Halgren et al., 2002; Lehmann et al., 1998). In addition, the GWT/LIDA model makes the testable claim that conscious cognition is necessary for non-trivial learning to occur. We believe this approach has been fruitful and will lead to further progress.

References

Baars, B. J. (1983). Conscious contents provide the nervous system with coherent, global information. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness & self-regulation*. New York: Plenum Press.

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.

Baars, B. J. (1997). *In the theater of consciousness*. Oxford: Oxford University Press.

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

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

[8] Though there's frequent activation passing, there are no neural nets as such since no transfer function is applied to a node's activation.

Baddeley, A. D. (1993). Working memory and conscious awareness. In A. Collins, S. Gathercole, M. Conway, & P. Morris (Eds.), *Theories of memory*. Howe: Erlbaum.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science*, *4*, 417–423.

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

Chein, J. M., & Schneider, W. (2005). Neuroimaging of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*, *25*, 607–623.

Conway, M. A. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. In A. Baddeley, M. Conway, & J. Aggleton (Eds.), *Episodic memory*. Oxford: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–185.

Dehaene, S. (Ed.) (2002). *The cognitive neuroscience of consciousness*. Cambridge, MA: MIT Press.

Dennett, D. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT Press.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*, 21–245.

Franklin, S. (2000). Modeling consciousness and cognition in software agents. In Neils Taatgen (Ed.), *Proceedings of the third international conference on cognitive modeling* (pp. 100–109). Veenendal, NL: Universal Press.

Franklin, S. (2001). Automating human information agents. In Z. Chen, & L. C. Jain (Eds.), *Practical applications of intelligent agents*. Berlin: Springer-Verlag.

Franklin, S. (2003). IDA: A conscious artifact? *Journal of Consciousness Studies*, *10*, 47–66.

Franklin, S., Baars, B. J., Ramamurthy, U., & Ventura, M. (2005). The role of consciousness in memory. *Brains, Minds and Media*, *1*, 1–38, pdf.

Franklin, S., & Graesser, A. C. (1997). Is it an agent, or just a program? A taxonomy for autonomous agents. In *Intelligent agents III*. Berlin: Springer Verlag.

Franklin, S., & Patterson, F. G., Jr. (2006). The lida architecture: Adding new modes of learning to an intelligent, autonomous, software agent. In *Idpt-2006 proceedings (integrated design and process technology)*. Society for Design and Process Science.

Freeman, W. J. (1999). *How brains make up their minds*. London: Weidenfeld & Nicolson General.

Freeman, W. J. (2003). Evidence from human scalp electroencephalograms of global chaotic itinerancy. *Chaos*, *13*, 1067–1077.

Fuster, J. M., Bodner, M., & Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, *405*, 347–351.

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

Gray, J. (2002). The sound of one hand clapping. *Psyche*, *8*, 8–11.

Halgren, E., Boujon, C., Clarke, J., Wang, C., & Chauvel, P. (2002). Rapid distributed fronto-parieto-occipital processing stages during working memory in humans. *Cerebral Cortex*, *12*, 710–728.

Hayes-Roth, F., & Lesser, V. R.(1977). Focus of attention in the Hearsay-II system. In *Proc. 5th int. jt. conf artifictal intelltgence*.

Hofstadter, D. R., & Mitchell, M. (1995). The copycat project: A model of mental fluidity and analogy-making. In K. J. Holyoak (Ed.), *Logical connections*: *Vol. 2*. *Advances in connectionist and neural computation theory* (pp. 205–267). Norwood, NJ: Ablex (J A O. Role of Editor eds Barnden).

Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*, 1395–1399.

Lehmann, D., Strik, W. K., Henggeler, B., Koenig, T., & Koukkou, M. (1998). Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. *International Journal of Psychophysiology*, *29*, 1–11.

Maturana, H. R. (1975). The organization of the living: A theory of the living organization. *International Journal of Man-Machine Studies*, *7*, 313–332.

Maturana, R. H., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. Dordrecht, Netherlands: Reidel.

Nii, H. P. (1986). The blackboard model of problem solving and the evolution of blackboard architectures. *The AI Magazine*, (Summer), 38–53.

Pulvermueller, F., Shtyrov, Y., & Ilmoniemi, R. (2003). Spatiotemporal dynamics using neural language processing: An MEG study using minimum-norm current estimates. *NeuroImage*, *20*, 1020–2025.

Ramamurthy, U., Baars, B. J., D'Mello, S. K., & Franklin, S. (2006). Lida: A working model of cognition. In Danilo Fum, Fabio Del Missier, & Andrea Stocco (Eds.), *Proceedings of the 7th international conference on cognitive modeling* (pp. 244–249). Trieste: Edizioni Goliardiche.

Skarda, C., & Freeman, W. J. (1987). How brains make chaos in order to make sense of the world. *Behavioral and Brain Sciences*, *10*, 161–195.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human informaton processing: I. detection, search, and attention. *Psychological Review*, *84*, 1–57.

Shanahan, M. (2006). A cognitive architecture that combines internal simulation with a global workspace. *Consciousness and Cognition*, *15*, 433–449.

Sloman, A. (1999). What sort of architecture is required for a human-like agent? In M. Wooldridge, & A. Rao (Eds.), *Foundations of rational agency*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Snodgrass, M., & Shevrin, H. (2006). Unconscious inhibition and facilitation at the objective detection threshold: Replicable and qualitatively different unconscious perceptual effects. *Cognition*, *101*, 43–79.

Tse, P. U., Martinez-Conde, S., Schlegel, A. A., & Macknik, S. L. (2005). Visibility, visual awareness, and visual masking of simple unattended targets are confined to areas in the occipital cortex beyond human V1/V2. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 17178-83.

VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Science*, *7*, 207–213.

Varela, F. J, Thompson, E., & Rosch, Eleanor (1991). *The embodied mind*. Cambridge, MA: MIT Press.