Varieties of Unconscious Processes

Stan Franklin and Bernard J. Baars

Introduction

Cognitive processes come in two varieties, conscious and unconscious, thought to be distinguished by the contents of conscious processes being broadly accessible throughout the nervous system (Baars, 2002). This conscious access hypothesis is consistant with a theory of conscious content being accessed throughout the thalamacortical core (Edelman & Tononi, 2000). A dynamical systems approach leads to the same hypothesis (W J Freeman, 2003). Philosophers have also discussed this hypothesis (Block, 2007). Finally, the conscious access hypothesis is the central tenet of Global Workspace Theory (Baars, 1988) (see below).

But unconscious process also have their varieties. These are fundamental to our understanding of conscious events, because they provide the necessary comparison conditions. Unconscious brain processes often resemble conscious ones very closely, and it is the *differences* between them that provide clues about the nature of consciousness.

The contents of *preconscious processes* may remain unconscious or may become part of the conscious contents, while we will introduce the term "Never-conscious" to refer to those brain events that never come to consciousness --- such as automatisms, neuronal processes in the cerebellum, and long-term procedural memories. Deheane et al (Dehaene et al, 2006) have suggested using *subliminal processing* for this purpose, but that term has been in regular usage for more than 50 years in the field of subliminal (subthreshold) sensory perception. We therefore prefer to use "Never-conscious," abbreviated as N-conscious.

The LIDA model of Global Workspace Theory (Franklin, 2006; Franklin, et al, 2005) (see below) enables us to catalog N-conscious vs. preconscious processes. The sections below will describe the LIDA model, catalog N-conscious and preconscious processes, and offer an explanation of the functionally of the distinction.

Global Workspace Theory

Global workspace theory (GWT) (Baars 1988) was originally conceived as a neuropsychological model of consciousness, but has come to be widely recognized as a high-level theory of human cognitive processing, which is well supported by empirical studies (Baars 2002). GWT views the nervous system as a distributed parallel system with many different specialized processes. Some coalitions of these processes enable the agent to make sense of the sensory data coming from the current environmental situation. Other coalitions incorporating the results of the processing of sensory data compete for attention. The winner occupies what Baars calls a global workspace, whose settled contents are broadcast to all other processes. These contents of the global workspace are presumed to be conscious. This conscious broadcast serves to recruit other processes to be used to select an action with which to deal with the current situation. GWT is a theory of how consciousness functions within cognition. Unconscious contexts influence this competition for consciousness.

GWT is constantly being tested against new bodies of evidence. For example, any adequate theory must account for both conscious and unconscious brain conditions. The unconscious comparison conditions include:

1. Large brain regions that are highly active but never result in direct conscious output, such as the cerebellum and basal ganglia. The cerebellum, for example, may have as many neurons as the cerebral cortex. But cortex has traditionally been regarded as the source of conscious contents, while these other regions are not. It is obvious therefore that the sheer number of neurons involved in brain structures does not necessarily give rise to conscious experiences.
2. Sizable regions of the cortex itself are not believed to give rise to conscious contents, notably the "dorsal visual stream," which includes most of the parietal lobes. The same point applies here. Some researchers maintain that cortex "as such" gives rise to consciousness. But that is not true for the dorsal stream.
3. The entire cortex is unconscious for about six hours of the 24-hour cycle. Some researchers maintain that true unconsciousness only during the so-called "DOWN" states of the delta wave during Slow-Wave Sleep. If that is true, such totally unconscious events only occur for a total of three hours per 24 hour cycle. Other evidence indicates that Slow-Wave Sleep becomes deeper through the night, and that more brain regions decrease their metabolic activity over the hours of sleep.
4. It has long been known that novel skills are more conscious than the identical behaviors after practice. Riding a bicycle is a good example. A great deal of evidence indicates that cortical activity declines rapidly with the loss of conscious access to repeated skills (e.g., Schneider, 2009). Automatic (habitual) processes are believed to be handled by the basal ganglia when they become highly predictable. Cortical synaptic changes are no doubt involved as well.
5. A number of experimental ways of manipulating conscious access have been carefully studied over half a century, including subliminal stimulation, visual backward masking, the attentional blink and selective attention. In many cases the unconscious comparison condition has been demonstrated to activate the same brain regions that are involved in the conscious condition (e.g., Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006).

6. A number of disease states lead to characteristic impairments of conscious states and experiences. These include certain epileptic seizures and epileptic "automatisms," loss of consciousness due to fainting, sleep disorders like sleep walking, brain impairments leading to vegetative states and minimally conscious states, analgesia and anesthesia, hallucinations and delusions, and so on. The same conditions can be drug-induced and hypnotically suggested in highly suggestible subjects.

The LIDA Model and its Architecture

The LIDA model is a comprehensive, conceptual and computational[1] model covering a large portion of human cognition (Stan Franklin & Patterson, 2006; Stan Franklin, et al., 2007). Besides GWT, the model implements and fleshes out a number of psychological and neuropsychological theories including situated cognition (Varela, Thompson, & Rosch, 1991) perceptual symbol systems (Barsalou, 1999), working memory (Baddeley & Hitch, 1974), memory by affordances[2] (Glenberg, 1997). long-term working memory (Ericsson & Kintsch, 1995), and the H-CogAff architecture (Sloman, 1999). The comprehensive LIDA model includes a broad array of cognitive modules and processes, a database of which, including known possible neural correlates can be found online at <ccrg.cs.memphis.edu/tutorial/correlates.html>.

The LIDA Cognitive Cycle

The LIDA model and its ensuing architecture are grounded in the LIDA cognitive cycle. Every autonomous agent (Stan Franklin & Graesser, 1997), human, animal, or artificial, must frequently sample (sense) its environment and select an appropriate response (action). Sophisticated agents such as humans process (make sense of) the input from such sampling in order to facilitate their decision making. Neuroscientists call this three part process the action-perception cycle (W J Freeman, 2002). The agent's "life" can be viewed as consisting of a continual sequence of these cognitive cycles. Each cycle constitutes a unit of sensing, processing and acting. A cognitive cycle can be thought of as a cognitive "moment." Higher-level cognitive processes are composed of many of these cognitive cycles, each a cognitive "atom."

---

[1] At this writing the LIDA model is only partially implemented. We claim it as a computational model since each of its modules and most of its processes have been designed for implementation.

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

Just as atoms have inner structure, the LIDA model hypothesizes a rich inner structure for its cognitive cycles (Baars & Franklin, 2003), (Stan Franklin, Baars, Ramamurthy, & Ventura, 2005). During each cognitive cycle the LIDA agent first makes sense of (see below) its current situation as best as it can by updating its representation of both external and internal features of its world. By a competitive process to be described below, it then decides what portion of the represented situation is most in need of attention. This portion is broadcast, making it the current contents of consciousness, and enabling the agent to choose an appropriate action and execute it.

Figure 1 shows the process in more detail. It starts in the upper left corner and proceeds roughly clockwise.

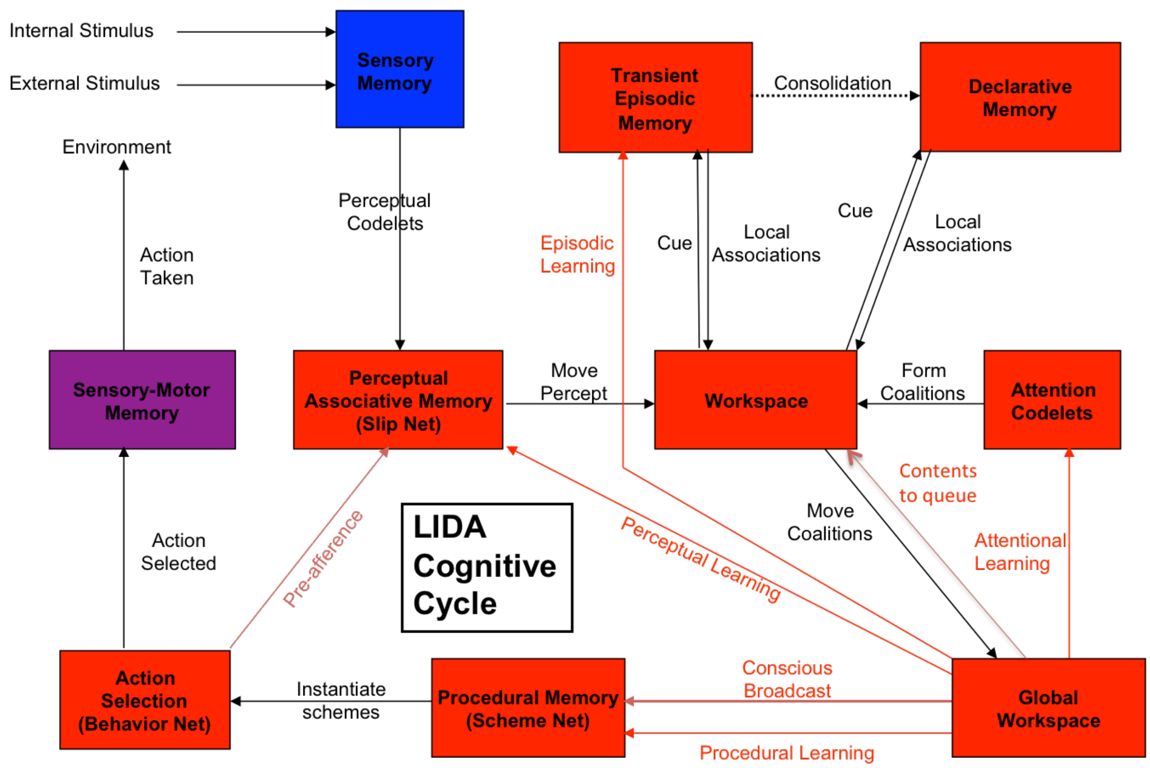

Figure I: LIDA Cognitive Cycle Diagram

The cycle begins with sensory stimuli from external and internal sources in the agent's environment. Low-level feature detectors in sensory memory begin the process of making sense of the incoming stimuli. These low-level features are passed on to perceptual memory where higher-level features, such as objects, categories, relations, situations, etc. are recognized. These entities, which have been recognized preconsciously, make up the percept that passes to the workspace, where a model of the agent's current situation is assembled. This percept serves as a cue to two forms of episodic memory, transient and declarative. Responses to the cue consist of local associations, that is, remembered events from these two memory systems that were associated with the various elements of the cue. In addition to the current percept, the workspace contains recent percepts and the models assembled from them that haven't yet decayed away.

A new model of the agent's current situation is assembled from the percepts, the associations, and the undecayed parts of the previous model. This assembling process will typically require structure-building codelets[3]. These structure-building codelets are small, special purpose processors, each of which has some particular type of structure it is designed to build. To fulfill their task these codelets may draw upon perceptual memory and even sensory memory, to enable the recognition of relations and situations. The newly assembled model constitutes the agent's understanding of its current situation within its world. It has made sense of the incoming stimuli.

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

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

The LIDA model hypothesizes that all human cognitive processing is via a continuing iteration of such cognitive cycles. These cycles occur asynchronously, with each cognitive cycle taking roughly 300 ms. These cycles

[3] The term codelet refers generally to any small, special purpose processor or running piece of computer code.

cascade, that is, several cycles may have different processes running simultaneously in parallel. This cascading must, however respect the serial nature of conscious processing necessary to maintain the stable, coherent image of the world it provides (Stan Franklin, 2005; Merker, 2005). This cascading, together with the asynchrony, allows a rate of cycling in humans of five to ten cycles per second. A cognitive "moment" is thus quite short! There is considerable empirical evidence from neuroscience suggestive of and consistent with such cognitive cycling in humans (Massimini, et al., 2005; Sigman & Dehaene, 2006; Uchida, Kepecs, & Mainen, 2006; Willis & Todorov, 2006). None of this evidence is conclusive, however.

Preconscious and N-conscious Processes

Figure 1 illustrates the various cognitive processes as hypothesized by the LIDA model, and the relationships between. Beginning at the upper left, Sensory Memory consists of the initial representations of sensory input together with the products of the primary level feature detectors (Koch, 2004). The contents of Sensory Memory are N-conscious, in that they can never become conscious. These contents are sent to Sensory-Motor Memory at a 40 hz rate to enable the execution of the action selected during each cognitive cycle. In the visual system, for example, these contents travel along the dorsal stream (Goodale & Milner, 2004)

Several subsequent processes in the figure, Perceptual Associative Memory, the Workspace, and the two episodic memories, are concerned with making sense of the current situation. Any of the contents of these processes may be included in the winning coalition in the global workspace. Thus all these contents are preconscious. The workings of the attention codelets and the global workspace (competition for consciousness and the conscious broadcast) never become conscious and are thus N-conscious. The same is true of the workings and the contents of Procedural Memory, Action Selection and Sensor-Motor Memory. All

References

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. 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.

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

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

Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences, 10*, 204–211.

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

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

Franklin, S. (2005). Evolutionary Pressures and a Stable World for Animals and Robots: A Commentary on Merker. *Consciousness and Cognition, 14*, 115–118.

Franklin, S., Baars, B. 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 *Intelligent Agents III* (pp. 21–35). Berlin: Springer Verlag.

Franklin, S., & Patterson, F. G. J. (2006). The LIDA Architecture: Adding New Modes of Learning to an Intelligent, Autonomous, Software Agent *IDPT-2006 Proceedings (Integrated Design and Process Technology)*: Society for Design and Process Science.

Franklin, S., Ramamurthy, U., D'Mello, S. K., McCauley, L., Negatu, A., L., R. S., et al. (2007, November 9-11). *LIDA: A Computational model of global workspace theory and developmental learning.* Paper presented at the AAAI Fall Symposium on AI and Consciousness: Theoretical Foundations and Current Approaches, Arlington, VA.

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

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

Goodale, Melvyn A., & Milner, D. (2004). *Sight Unseen*. Oxford: Oxford University Press.

Koch, C. (2004). *The Quest for Consciousness: A neurobiological approach*. Englewood, Colorado Roberts & Co.

Massimini, M., Ferrarelli, F., Huber, R., Esser, Steve K., Singh, H., & Tononi, G. (2005). Breakdown of Cortical Effective Connectivity During Sleep. *Science, 309*, 2228–2232.

Merker, B. (2005). The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition, 14*, 89–114.

Schneider, W. (2009). Automaticity and Consciousness. In W. Banks (Ed.), *Encyclopedia of Consciousnes* (pp. 83-92). St. Louis: Elsevier

Sigman, M., & Dehaene, S. (2006). Dynamics of the Central Bottleneck: Dual-Task and Task Uncertainty. *PLoS Biol., 4*.

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.

Uchida, N., Kepecs, A., & Mainen, Zachary F. (2006). Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision making. *Nature Reviews Neuroscience, 7*, 485–491.

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

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a Face. *Psychological Science, 17*, 592–599.