Science &Consciousness Review

http://psych.pomona.edu/scr

Editorial Review

A computer-based model of Crick and Koch's Framework for Consciousness

By Stan Franklin

[SCR, 2003 March, No. 1]

The Framework Recently Crick and Koch offered a "Framework for Consciousness" (2003). Pradeep Mutalik's review of that article in SCR (Mutalik 2003) asserts that "Crick and Koch describe ten aspects of a framework that they believe offers a coherent scheme for explaining the neural correlates of consciousness."

A framework must not be confused with a set of hypotheses

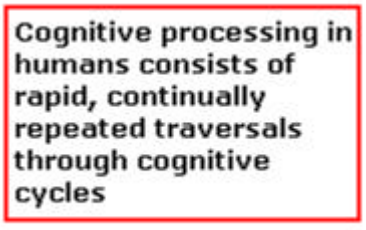
Crick and Koch explain that a framework must not be confused with a set of hypotheses. Rather a framework for consciousness offers a point of view from which to address the problems of consciousness. It's intended to guide research. A good framework should fit within current scientific knowledge reasonably well and should be roughly correct. It needn't be correct in all its details, but rather should guide research to fill in and correct it details. Such frameworks have proved useful in Biology and Physics. This one can be expected to be useful in consciousness studies.

Frameworks, by their very nature contain large gaps that can be filled with the results of later research. They can also be filled, temporarily, by hypotheses derived from cognitive models to give more specific guidance to research than does the framework itself. In this case, hypotheses from the IDA model of consciousness and cognition can serve to clarify some aspects of this framework and to flesh out others.

The IDA Model IDA is a conceptual (and computational) model of cognition (Franklin & Graesser, A. 2001, Franklin 2002) implemented as a software agent (Franklin & Graesser, A. C. 1997). IDA "lives" on a computer system with connections to the Internet and various databases and does personnel work for the US Navy, performing all the specific personnel tasks of a human (Franklin 2001). In particular, IDA negotiates with sailors in natural language, deliberates, and makes voluntary action selections. She has been the subject of a prior SCR article, **"IDA on Will: It's no Illusion"**.

Based on Baars' Global Workspace Theory of consciousness (Baars 1988, 2002) the IDA model is also consistent with several other psychological theories. The IDA model fleshes out Global Workspace theory so as to yield a fine-grained functional account of the steps involved in perception, several kinds of memory, consciousness, context setting, and action selection. We refer to this particular sequence of steps as a *cognitive cycle.* as described in some detail in the attached text box (Baars & Franklin 2003 in press). The reader is encouraged to read through, or at least skim, the cognitive cycle text box now, and to refer back to it frequently while reading the rest of this article. More details on the various steps in the cycle can be found in the references, and in the previous SCR article mentioned above. Codelets are described just below

The Cognitive Cycle The IDA model hypothesizes that cognitive processing in humans, as in IDA, consists of rapid, continually repeated traversals through these cognitive cycles. The cycles will overlap so as to act somewhat in parallel, but will carefully maintain the seriality of consciousness. A full cognitive cycle may take on the order of 200ms to run. But due to overlapping and to automaticity, which shortens the cycle, as many as a dozen cycles may be active during any given second. This cognitive cycling hypotheses implies that conscious contents come in discrete steps but rapidly enough to give the illusion of continuity.



Several of the steps in the cognitive cycle refer to codelets. Computationally a codelet is a small piece of code, a little program that performs one specialized, simple task. Codelets often play the role of demons waiting for a particular type of situation to occur and then acting as per their specialization. Codelets in the IDA model implement the processors postulated by global workspace theory. They may also correspond to Edeleman's neuronal groups (1987), Minsky's agents (1985) Jackson's demons (1987) and to Ornstein's small minds (1986). The notion of cognition being organized as this kind of a multi-agent system is widespread.

The Framework and the IDA Model In what follows the implications of IDA's cognitive cycling hypotheses for each of Crick and Kock's framework aspects will be discussed in turn. Quotes not otherwise identified will be from that very fine review and will be italicized. All in all, the IDA cognitive cycling hypotheses fits remarkably well into the Crick and Koch framework for consciousness, and serves to flesh it out with detailed functionality.

Keep in mind that the Crick and Koch framework is meant to guide research into the neural correlates of consciousness, and that it was developed with primate visual consciousness in mind. They also restrict their attention to processes occupying "…time periods on the order of a few hundred milliseconds, or at the most, several seconds…" This translates into one to a few tens of cognitive cycles.

The IDA cognitive cycling hypotheses fits remarkably well into the Crick and Koch framework for consciousness

The homunculus: *"…it is useful to think of the "front" or higher/executive part of the cortex … as looking at and interacting with the "back", or sensory part."* The IDA model would hypothesize that such looking and interacting happens over several consecutive cognitive cycles. Each individual cycle is largely, but by no means completely feedforward. Crick and Koch propose that humans are only conscious of sensory representations of their thought, not of the thoughts themselves. The IDA model concurs for the most part. For IDA, as we see in steps 4 and 5 of the cognitive cycle (CC4-consciousness-competition & CC5-broadcast, see text box), the contents of consciousness consist of the information carried by the member codelets of the winning coalition. These are for the most part sensory representations arising from the preconscious working memory buffer (CC2-percept-to-WM) or from the local associations (CC3-local-associations).

Zombie modes: *"These are response to sensory inputs that are rapid, transient, stereotyped and unconscious, and could be thought of as cortical reflexes…"* This mode is described in step

1 of the cognitive cycle (CC1-perception). *"…conscious modes deal with broader, more elaborate and less stereotyped actions."* These conscious modes are implemented in CC5-broadcast and CC6-recruitment in each cognitive cycle.

Coalitions of neurons: *"…conscious experiences are sustained by shifting coalitions of neurons."* CC4-consciousness-competition describes the formation of such coalitions and the resulting competition for consciousness. Where Crick and Koch talk of competition among coalitions of neurons, the IDA model imposes more structure, coalitions of processors (codelets, cell assemblies). *"…at any moment, the winning coalition is somewhat sustained, and embodies what we are conscious of."* This occurs during CC5-broadcast.

Explicit representations: *"Every explicit conscious representation of a particular visual scene has a group of neurons responsible for it. If they are all lost by brain damage, then the subject is unable to consciously perceive that aspect directly…"* CC4-consciousness-competition describes the role of attention codelets and information codelets in the organization of coalitions of codelets representing some particular conscious experience. The loss of an attention codelet responsible for organizing some particular type of conscious event would result in no further such events being experienced. The loss of information codelets would have more particular consequences. At an earlier step in the cognitive cycle, CC1b-perception describes the slipnet whose nodes extract meaning from incoming sensation. A missing node would result in certain types of stimuli not being recognized. Missing perception codelets from CC1a-perception would again have more particular consequences. Thus the IDA model suggests that there are two distinct ways for losses of sensory consciousness such as loss of color perception to occur. These would correspond to Crick and Koch's notion of the "back" and the "front" of the brain. Once again, there's good agreement between the framework and the IDA model, with the later providing more detail.

The higher level first: *"Neural activity first travels rapidly and unconsciously up the sensory hierarchy to a high level, and then starts to move backward down the hierarchy so that the first stages to reach consciousness are at the higher levels (e.g. showing the gist of the scene). This activity then moves further backward to give the details of the scene."* In IDA this describes the passing of activation among the nodes of the slipnet described in CC1b-perception.

Driving and modulating inputs: *"Inputs to a cortical neuron fall roughly into two broad classes: driving and modulating inputs. Driving inputs largely contact the basal dendrites, the*

classical input branches of neurons, whereas modulatory inputs include back-projections and are largely to the apical dendrites." There is nothing in the IDA model that sheds any additional light on this phenomenon.

Snapshots: *"Crick and Koch propose that conscious awareness is a series of static snapshots of shifting neuronal coalitions."* The IDA model makes precisely the same proposal with the contents of consciousness in CC5-broadcast of each cognitive cycle being one "snapshot."

"The coalitions need to be above a certain threshold to be conscious, but the nature of that threshold is completely unknown. We do not even know whether it is quantitative (firing rates, firing in bursts) or qualitative (pyramidal cells projecting to the front of the brain)." In the IDA model only a single coalition can win the competition for consciousness described in CC4-consciousness-competition. The winner is the most relevant, important, urgent coalition, that is, that coalition with the highest average activation among its constituent codelets. Perhaps, one day, this will be a testable hypothesis.

Attention and binding: *"Attention can be volitional and top-down … or saliency controlled and bottom-up."* The IDA model implements volitional decision making via James' ideomotor theory using one particular behavior stream (goal context hierarchy) in the behavior net operating over several cognitive cycles (see CC7-context-setting) (Franklin 2000). This includes voluntary decisions about what to attend to next. Saliency controlled attention (bringing to consciousness) occurs in single cognitive cycles.

"In either case, attention can bias competition among developing coalitions." Attention codelets do just that in CC4-consciousness-competition.

"The idea of coalitions effectively finesses the "binding problem"—how different attributes of the same object are brought together—by stating that the 'binding' of the features of a single object/event is simply the co-membership in a particular coalition." This is equally true as stated in the IDA model in CC4-consciousness-competition.

Styles of firing: "Synchronized firing may help a nascent coalition in its competition with other (nascent) coalitions. Such firing may not be needed once a successful coalition has reached consciousness." The IDA model has nothing to say about this aspect of the framework.

Penumbra and Meaning: *"The firing of NCC neurons will influence many neurons that are not part of the NCC. Crick and Koch call this sphere of influence the 'penumbra'. This penumbra includes past associations, expected consequences, possible plans for movement associated with NCC neurons, and other aspects that serve to give meaning to the percept."* The IDA model fleshes out this aspect of the framework. Past associations are implemented in long-term working memory (CC3-local-associations). Expectation is implemented by expectation codelets, a variety of attention codelet, described in CC6-recruitment and CC9-acting. Plans for movement are implemented as behavior streams (goal context hierarchies) as described in CC7-context-setting and CC8-action-selection. Not only are all these features of this aspect of the Crick and Koch framework present in the IDA model, but relationships between them are explored.

"Crick and Koch speculate that the penumbra neurons may be the site of unconscious priming." The role in IDA of the preconscious buffer and of long-term working memory in priming is described in CC4-consciousness-competition.

"Other authors have called this 'fringe' consciousness, after William James." The IDA model implements 'fringe' consciousness using particular 'fringe' attention codelets in CC4-consciousness-competition.

Conclusion As our trip through the ten aspects of the Crick and Koch framework has shown, there seems to be a remarkable concurrence between the aspects and the various steps in IDA's cognitive cycle. Though the IDA model adds nothing to two of the aspects, in each of the other cases it provides relatively detailed hypotheses about the functionality of the aspect and often about its structure. This concurrence between the framework and the model lends credence to both.

The IDA Cognitive Cycle

1. PERCEPTION. Sensory stimuli, external or internal, are received and interpreted by perception creating meaning. Note that this stage is unconscious.

a. EARLY PERCEPTION: Input arrives through senses. Specialized perception codelets descend on the input. Those that find features relevant to their specialty activate appropriate nodes in the slipnet (a semantic net with activation).

b. CHUNK PERCEPTION: Activation passes from node to node in the slipnet. The slipnet stabilizes bringing about the convergence of streams from different senses and chunking bits of meaning into larger chunks. These larger chunks, represented by meaning nodes in the slipnet, constitute the percept.

2. PERCEPT TO PRECONSCIOUS BUFFER. The percept, including some of the data plus the meaning, is stored in preconscious buffers of IDA's working memory. These buffers may involve visuospatial, phonological, and other kinds of information.

3. LOCAL ASSOCIATIONS. Using the incoming percept and the residual contents of the preconscious buffers as cues, local associations are automatically retrieved from transient episodic memory and from long-term associative memory. The contents of the preconscious buffers together with the retrieved local associations form transient episodic memory and long term associative memory. Together these roughly correspond to Ericsson and Kintsch's long-term working memory (Ericsson & Kintsch 1995) and Baddeley's episodic buffer (Baddeley 1993).

4. COMPETITION FOR CONSCIOUSNESS. Attention codelets, whose job it is to bring relevant, urgent, or insistent events to consciousness, view long-term working memory. Some of them gather information, form coalitions and actively compete for access to consciousness. The competition may also include attention codelets from a recent previous cycle.

The activation of unsuccessful attention codelets decays, making it more difficult for them to compete with newer arrivals. However, the contents of unsuccessful coalitions remain in the preconscious buffer and can serve to prime ambiguous future incoming percepts. The same is true of contents of long-term working memory that aren't picked up by any attention codelet.

5. CONSCIOUS BROADCAST. A coalition of codelets, typically an attention codelet and its covey of related information codelets carrying content, gains access to the global workspace and has its contents broadcast. This broadcast is hypothesized to correspond to phenomenal consciousness.

The current contents of consciousness are also stored in transient episodic memory. At recurring times not part of a cognitive cycle, the contents of transient episodic memory are consolidated into long-term associative memory (Shastri 2002). Transient episodic memory is an associative memory with a relatively fast decay rate (Conway 2001). It is to be distinguished from autobiographical memory, a part of long-term associative memory.

6. RECRUITMENT OF RESOURCES. Relevant behavior codelets respond to the conscious broadcast. These are typically codelets whose variables can be bound from information in the conscious broadcast. If the successful attention codelet was an expectation codelet calling attention to an unexpected result from a previous action, the responding codelets may be those that can help to rectify the unexpected situation. Thus consciousness solves the relevancy problem in recruiting resources.

7. SETTING GOAL CONTEXT HIERARCHY. Some responding behavior codelets instantiate an appropriate behavior stream, if a suitable one is not already in place. They also bind variables, and send activation to behaviors. Here we assume that there is such a behavior codelet and behavior stream. If not, then non-routine problem solving using additional mechanisms is called for.

8. ACTION CHOSEN. The behavior net chooses a single behavior (goal context) and executes it. This choice may come from the just instantiated behavior stream or from a previously active stream. The choice is affected by internal motivation (activation from goals), by the current situation, external and internal conditions, by the relationship between the behaviors, and by the activation values of various behaviors.

9. ACTION TAKEN. The execution of a behavior (goal context) results in the behavior codelets performing their specialized tasks, which may have external or internal consequences. This is IDA taking an action. The acting codelets also include an expectation codelet (see Step 6) whose task it is to monitor the action and to try and bring to consciousness any failure in the expected results.

References

Baars, B. J. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.

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

Baars, B. J., and S. Franklin. 2003 in press. How conscious events may interact with Working Memory. *Trends in Cognitive Science*.

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

Conway, M. A. 2001. Sensory-perceptual episodic memory and its context: autobiographical memory. In *Episodic Memory*, ed. A. Baddeley, M. Conway, and J. Aggleton. Oxford: Oxford University Press.

Crick, F., and C. Koch. 2003. A framework for consciousness. *Nature Neuroscience* 6:119–126.

Edelman, G. M. 1987. *Neural Darwinism*. New York: Basic Books.

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

Franklin, S. 2000. Deliberation and Voluntary Action in 'Conscious' Software Agents. *Neural Network World* 10:505–521.

Franklin, S. 2001. Automating Human Information Agents. In *Practical Applications of Intelligent Agents*, ed. Z. Chen, and L. C. Jain. Berlin: Springer-Verlag.

Franklin, S. 2002; ed. IDA on Will: It's no Illusion. Science and Consciousness Review. 2002. http://psych.pomona.edu/scr/ED_Sep02_IDAFreeWill.htm.

Franklin, S., and A. Graesser. 2001. Modeling Cognition with Software Agents. In *CogSci2001*: *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*, ed. J. D. Moore, and K. Stenning. Mahwah, NJ: Lawrence Erlbaum Associates; August 1-4, 2001.

Franklin, S., and A. C. Graesser. 1997. Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Intelligent Agents III*. Berlin: Springer Verlag.

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

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

Mutalik, P. Crick & Koch's a new "Framework for Consciousness". Science and Consciousness Review, February, No. 1. http://psych.pomona.edu/scr/editorials/20030201.html.

Ornstein, R. 1986. *Multimind*. Boston: Houghton Mifflin.

Shastri, L. 2002. Episodic memory and cortico-hippocampal interactions. *Trends in Cognitive Sciences* 6:162–168.