2012

Meta Learning, Change of Internal Workings, and LIDA: A commentary on Thórisson and Helgasson's "Cognitive architectures and autonomy: A Comparative Review"

R. McCall

S. Franklin

Meta Learning, Change of Internal Workings, and LIDA

Ryan McCall RMcCall@Memphis.Edu
Fedex Institute of Technology 403h
The University of Memphis
Memphis, TN 38152, USA

Stan Franklin Franklin@Memphis.Edu
Fedex Institute of Technology 301
The University of Memphis
Memphis, TN 38152, USA

Abstract

Here we comment on issues of concern in the target article including meta learning and change of internal workings. We also address several misinterpretations of the LIDA model.

Keywords: meta learning, autonomy, self-modification, resource management, embodied cognition, processing time

1. Introduction

The authors review a number of well-known cognitive architectures, as well as several explicitly chosen for their integrated nature and broad scope. Recently there have been other comparisons of cognitive architectures including one by Duch et al. (2008) in addition to the BICA table[1].

2. Meta Learning and Change of Internal Workings

A novel aspect of this review is its focus on autonomy. The authors identify “meta learning” as a necessary sub-aspect of autonomy, and use it as a metric for comparing architectures. There are several issues with this process. First, no definition or defining references in the initial explanation of meta learning are given. However, the term is used in multiple contexts, e.g. metacognition, where it is called metacognitive regulation[2], and also in the context of machine learning (e.g. Schaul & Schmidhuber 2010). What aspects of “meta learning” can be accomplished by biological agents (e.g. humans)? What aspects of “meta learning” go beyond the capabilities of biological agents?

While there is ample evidence that the ability to monitor and regulate one’s learning is absolutely critical — at least in humans, it seems stronger to suggest that an agent might “change its own internal workings.” Does this include major changes to the agent’s architecture? There doesn't seem to be any justification for such drastic self-modification being necessary or even valuable. Has there been a demonstration that an agent changing its “own internal workings” in real-time has improved performance? It seems that if an agent changes its own architecture online, there could be a catastrophic failure resulting in the agent’s being inoperable. For example a change in “internal workings” of a human brain could produce epilepsy rendering that human dysfunctional.

---

[1] http://bicasociety.org/cogarch/architectures.htm
[2] http://en.wikipedia.org/wiki/Metacognition

Why is meta learning, as defined, a research priority when a working human level AGI system has not yet been produced? It could be argued that the best meta learning approach for the foreseeable future is for humans to continue researching AGI architectures, and generating and testing various designs. No AGI architecture seems remotely close to be competing with humans on AGI design capabilities at this time.

3. Clarifying Issues with the LIDA Architecture

This portion of this comment is dedicated to explicating misimpressions of the LIDA model.

Describing LIDA, page 19 reads "... availability of resources does not affect the processing of the system, with each operating cycle always selecting a single coalition of data for further processing." Though the second clause of the quoted sentence is correct, the first clause does not follow from it. Here are four examples: 1) Even over a single cognitive cycle, availability of resources would typically affect the determination of saliency (presence, importance, urgency, insistence, novelty, unexpectedness, loudness, brightness, motion, etc.), the currency of the competition for consciousness. For example a situation recognized as urgent (perhaps due to a limited time situation) will have increased saliency thus increasing its chance to be attended upon by the consciousness mechanism. 2) Over multiple cycles, LIDA has long implemented, at least conceptually, deliberation. Given a circumstance with multiple appropriate options (information overload) LIDA performs volitional decision making via James' ideomotor theory (1890). LIDA's implementation of volitional decision making employs a timekeeper whose time window determines when the volition must stop, and a decision be made (Franklin 2000). 3) Since schemes are chosen for instantiations largely by the intersection of their context with the current conscious broadcast, the availability of resources plays a critical role in deciding which schemes in Procedural Memory are instantiated into behaviors to be considered by the Action Selection module. 4) Finally, recent work on LIDA modeling the attentional blink phenomenon (Madl, Franklin, 2012) has employed a finite attention resource that may be temporarily "used up" by demanding percepts.

On the same page we find "...the many different types of learning supported by the architecture, both symbolic (e.g. declarative) and sub-symbolic (e.g. perceptual)." The LIDA model is in no sense symbolic. Rather it would come under the category of embodied (situated) cognition, as it implements Barsalou's perceptual symbol system in its Perceptual Associative Memory. One might classify LIDA as subsymbolic in that activation passing occurs throughout. However, there are no artificial neural networks, as such, implemented in LIDA. So, the term "subsymbolic" might also be a little misleading.

There are also issues with the assessment of LIDA along the "Realtime" dimension. Ticks are not "operating cycles", that is, they are not directly tied to the cognitive cycle. Ticks are a discretization of time for purpose of computer implementation of LIDA's processes, and are specific to the LIDA computational framework. One tick can take an arbitrarily short amount of real time given a sufficiently powerful computer. Ticks are not a part of the conceptual LIDA model, however the processes of the LIDA model do differ in their real-time length, which has consequences for LIDA's real-time performance. To further illustrate: to achieve a cognitive cycle of 300 ticks there may be several processes running at various frequencies e.g. 20, 100, 50 etc. Please see Snaider, McCall, & Franklin (2011) for an extended discussion of ticks and the larger LIDA computational framework. Madl, Baars, & Franklin, (2011) gives an example of an agent implementation using ticks.

Central to the LIDA model is the cognitive cycle. LIDA's cognitive cycles cascade, that is, they overlap. As a model of human cognition LIDA predicts that three different such sense-action cycles could be occurring concurrently (each in a different processing stage) (Madl, Baars, Franklin, 2011). The only bottleneck in the cognitive cycle comes from the consciousness mechanism. Here there is a minimum amount of time that must pass (~100ms in humans) between successive conscious broadcasts which functions to provide stability and to preserve seriality.

Time can be involved in behavior streams implementing higher order cognitive processes. For example, as mentioned earlier, LIDA performs volitional decision making via James' ideomotor theory (1890), using a timekeeper whose time window determines when volition must stop and a decision be made (Franklin 2000). This is one example of time as a resource that may affect LIDA's cognitive processing. Another example is in LIDA's consciousness mechanism where a time-based trigger may initiate a competition for consciousness if a certain amount of time has passed with no conscious broadcast occurring (see Madl, Baars, Franklin (2011)).

References

Duch W., Oentaryo R. J., Pasquier M., *Cognitive architectures: where do we go from here?* In: Frontiers in Artificial Intelligence and Applications, Vol. 171 (Ed. by Pei Wang, Ben Goertzel, and Stan Franklin), IOS Press, pp. 122-136, 2008.

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

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

Madl, T., Baars, B. J., & Franklin, S. (2011). The Timing of the Cognitive Cycle. *PLoS ONE, 6*(4), e14803.

Madl, T., & Franklin, S. (2012, April 13-15). *A LIDA-based Model of the Attentional Blink.* Paper presented at the 11th International Conference on Cognitive Modeling, Berlin.

Schaul, T. & Schmidhuber, J. (2010) Metalearning. *Scholarpedia, 5*(6):4650.

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.* The Proceedings of the Fourth Conference on Artificial General Intelligence (AGI-11).