A Standard Model of Mind:
AAAI Technical Report FS-17-05

ICARUS' Implications for the Standard Model of Mind

Dongkyu Choi
Department of Aerospace Engineering
University of Kansas
1530 W 15th Street
Lawrence, Kansas 66045
dongkyuc@ku.edu

Abstract

ICARUS is a cognitive architecture that shares many features with other theoretical frameworks, including a recent proposal for a standard model of human-like minds, but it also makes some distinctive assumptions. In this paper, we review ICARUS, discuss how it relates to this proposal, and propose extensions based on its novel features.

Introduction

Researchers from different background have proposed a variety of accounts for intelligence. The AI community typically promotes functional abilities but gives less consideration to producing human-like behavior. In contrast, cognitive scientists aim to develop computational models of human behavior, using specific tasks for evaluation purposes. Neuroscience focuses instead on the physical infrastructure, while robotics studies embodied agents, often focusing on perception and action.

These differing viewpoints suggest it will be impractical to produce a truly standard model of the mind. However, the organizers of this meeting have found that, surprisingly, there appears to be a loose but notable consensus toward a common set of assumptions. This suggests that the relevant fields are now mature enough to propose organized accounts of human-like minds both to scientists in other fields and to society in general.

However, there remain many aspects of this emerging model that are open to debate. As the organizers have suggested, this may be desirable, with initial accounts acting mainly to invoke discussions, debates, and comparisons, rather than serving as specifications or constraints on existing or future systems that attempt to explain or reproduce intelligent behavior.

In this paper, we review ICARUS, a cognitive architecture that shares a number of assumptions with other theories but that also makes distinctive claims. We also discuss how the architecture relates to elements of the proposed standard model, finding that it agrees on more features than otherwise. In closing, we return to ICARUS' distinguishing characteristics, proposing that they should be considered for inclusion in the emerging consensus.

Architectural Commitments of ICARUS

ICARUS is an architecture for intelligent agents that has been under development, in its current form, for over ten years (Choi & Langley, in press). The framework shares some basic aspects with other cognitive architectures like ACT-R and Soar. These include assumptions that:

- Short-term memories, which contain information that changes rapidly, are distinct from long-term stores, which contain elements that are static or that change gradually.
- Both types of memories comprise collections of symbolic structures that encode relational content through shared symbols, typically cast as list structures.
- Retrieval of long-term structures, and interpretation of short-term elements, arises through relational pattern matching accesses of the former against the latter.
- Cognitive processing alternates between accessing stable structures through pattern matching, selecting some of the structures, and using them to alter dynamic memories in a recognize-act cycle.
- High-level cognition arises from the dynamic composition of mental structures during performance, where changes to short-term memory on cycle enable changes on later ones, and learning, which combines existing long-term structures into new ones.

The specific ways in which ICARUS incorporates these ideas differ somewhat from traditional production-system architectures. For instance, the architecture posits distinct short-term memories for the agent's beliefs and goals, rather than a single working memory. Nevertheless, these similarities are strong enough that it is clearly a member of the same theoretical family.

However, ICARUS also makes additional assumptions that differentiate it from other architectures. These include claims that:

- *Cognition is grounded in perception and action*. Every symbol is linked to results of sensors, calls to effectors, or their imagined analogs.
- *Categories and skills are distinct cognitive structures*. The first are used to describe states and goals, whereas the second describe ways to achieve goals by altering states.

- *Long-term knowledge is organized in a hierarchical manner*. Complex concepts are defined in terms of simpler ones, where high-level skills are made up of lower-level ones.
- *Short-term elements are instances of long-term structures*. Beliefs and goals must be instantiated versions of defined concepts, whereas intentions must be instances of known skills.
- *Inference has primacy over execution, with both underlying problem solving*. Higher-level mechanisms operate over the results that lower-level processes desposit in short-term memories.

We believe these theoretical tenets are plausible candidates for addition to the standard model, but we will delay this topic until later.

First, we should discuss our position on elements that Laird et al. (in press) have proposed for inclusion. For each one, we summarize the claim and note ICARUS' stance on the issue:

- Long-term memories contain symbol structures and associated metadata. ICARUS encodes both conceptual rules and skills as such structures and can include information like recency.
- Factual knowledge is provided by declarative long-term memory. Most versions of the architecture do *not* have a declarative long-term memory of facts, although they can store 'static beliefs' that are stable over time. Some variants (Stracuzzi et al., 2009; Menager & Choi, 2016) also include a form of episodic memory.
- Processing is based on a small number of domain-independent modules. ICARUS includes modules for conceptual inference, skill execution, problem solving, and skill acquisition.
- Complex behavior arises from a sequence of independent cognitive cycles that operate in local context. ICARUS operates in cognitive cycles for conceptual inference, skill execution, and problem solving.
- Behavior is driven by a cognitive cycle with a duration comparable to humans' $\sim 50ms$. This holds for ICARUS' execution and problem solving, but conceptual inference operates on a faster cycle that supports them.
- Global communication occurs through a short-term working memory. ICARUS includes short-term memories for beliefs, goals, and intentions that inference, execution, and problem solving inspect and modify.
- The architectural processing supports bounded rationality, not optimality. The problem-solving module carries out heuristic search to solve novel problems and execution selects among skills whose conditions match; neither come with optimality guarantees.
- Perception yields symbol structures in specific working memory buffers. ICARUS assumes this occurs but does not model the perceptual process.

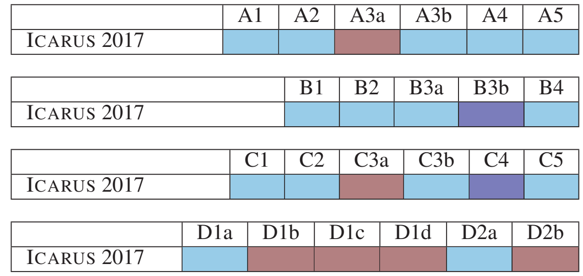

Figure 1: Analysis of ICARUS with respect to the standard model proposed in Laird et al. (in press). Light blue blocks indicate supported features and dark blue blocks mark partially or indirectly supported ones. Red blocks show currently unsupported aspects.

- Motor control converts symbolic structures in its buffers into external actions. The architecture also makes this assumption but does not address low-level control.
- All forms of long-term memory content can be learned in an online and incremental manner. ICARUS supports incremental acquisition of hierarchical skills from problem solving. Some variants can learn skills from demonstrations (Li et al., 2009), whereas others can acquire new conceptual predicates (Li et al., 2012).

Like many other frameworks, ICARUS framework also assumes parallel matching of conditions on conceptual rules and skills, and decision making during execution and problem solving results from interactions between knowledge and architecture-level processing. It does place more emphasis on the latter than Soar, in that problem solving relies on means-ends analysis rather than knowledge-based strategies. The main architecture does not have a facility for reinforcement learning, although the earliest version (Shapiro & Langley, 2002) did include it. Figure 1 summarizes how ICARUS maps onto the standard model with color-coded blocks. Lack of current support for a feature does not necessarily imply that the ICARUS theory rejects it in principle.

Proposed Extensions to Standard Model

Earlier we listed a number of ICARUS' theoretical commitments that do not appear in the standard model. In this section we discuss whether the latter should be extended to include these aspects of human cognition. We do not discuss the issue of grounded representation, as there appears to be an emerging consensus about its importance.

Concepts and Skills

The ICARUS theory maintains that concepts and skills are distinct types of cognitive structures that serve different functions. Concepts let the agent describe classes of situations and support conceptual inference, whereas skills let the agent describe activities in support of reactive execution and problem solving. One can encode concepts and skills in a uniform notation like production rules, but we believe this camouflages their important differences and their distinction should become part of the standard model.

Defined Predicates

As we have noted, ICARUS assumes that all predicates appearing in long-term and short-term memory refer to defined concepts or skills. Conceptual knowledge discretizes the agent's perceptions and serves as the vocabulary for states and goals, whereas skills aggregate actions into higher-level activities. Agents cannot specify beliefs, goals, or intentions in short-term memory without referring to such defined predicates. This is loosely analogous to ACT's claim that working memory is the active portion of declarative long-term memory, but it takes a quite different form that we believe should become part of the standard model.

Hierarchical Organization of Knowledge

Both concepts and skills in ICARUS are organized in a hierarchy that defines aggregate predicates in terms of simpler ones, letting it describe states, goals, and intentions at multiple levels of abstraction. This idea is similar to the notion that *chunks* (in the original sense, not that in Soar) serve to organize the agent's experience. Other architectures support such hierarchies, but they do not make the same strong theoretical commitment and we maintain it should become part of the standard model.

Goal Reasoning

The ICARUS architecture also has a commitment to goal reasoning. This includes a theoretical claim that agents have long-term structures that determine which top-level goals should be active in given situations. The goal-reasoning process matches and instantiates this knowledge to nominate, retract, and prioritize goals under relevant conditions. This lets agents adapt to situations by deciding which goals to pursue, thus supporting extended autonomy. Humans clearly exhibit the ability to set their own goals, making this another candidate for inclusion in the standard model.

Conclusions

In this paper, we reviewed the ICARUS architecture and examined its relation to the standard model of cognition. We discussed some high-level features that it shares with other architectures, some ways in which it differs, and its position on elements of a recent proposal for the standard model. Despite its differences, we saw that ICARUS shares many assumptions with other theories of the cognitive architecture.

However, we also argued that four postulates – the distinction between concepts and skills, the use of defined predicates in short-term memories, the hierarchical organization of knowledge, and the importance of goal reasoning – figure centrally in ICARUS but not in other frameworks. We believe that they play important roles in human cognition and, as such, that they merit serious consideration for inclusion in the standard model.

References

Choi, D., & Langley, P. in press. Evolution of the ICARUS architecture. *Cognitive Systems Research*.

Laird, J. E., Lebiere, C., and Rosenbloom, P. S. in press. A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine*.

Li, N., Stracuzzi, D. J., Langley, P., and Nejati, N. 2009. Learning hierarchical skills from problem solutions using means-ends analysis. *Proceedings of the Thirty-First Annual Meeting of the Cognitive Science Society*. Amsterdam.

Li, N., Stracuzzi, D. J., and Langley, P. 2012. Improving acquisition of teleoreactive logic programs through representation extension. *Advances in Cognitive Systems*, *1*, 109–126.

Menager, D. H., and Choi, D. 201). A robust implementation of episodic memory in a cognitive architecture. *Proceedings of the Thirty-Eighth Annual Meeting of the Cognitive Science Society*.

Shapiro, D., and Langley, P. 2002. Separating skills from preference: Using learning to program by reward. *Proceedings of the Nineteenth International Conference on Machine Learning* (pp. 570–577). Sydney: Morgan Kaufmann.

Stracuzzi, D. J., Li, N., Cleveland, G., and Langley, P. 2009. Representing and reasoning over time in a cognitive architecture. *Proceedings of the Thirty-First Annual Meeting of the Cognitive Science Society*. Amsterdam.