Progress and Challenges in Research on Cognitive Architectures

Pat Langley

Institute for the Study of Learning and Expertise
Palo Alto, California, USA

http://www.isle.org/~langley/

Thanks to Herbert Simon, Allen Newell, John Anderson, David Nicholas, John Laird, Randy Jones, and many others for discussions that led to the ideas in this talk.

What is a Cognitive Architecture?

A *cognitive architecture* (Newell, 1990) is an infrastructure for intelligent systems that:

- Specifies those facets of cognition that remain constant across different domains;
- Including memories and representations of elements in those memories, but *not* their content, which changes over time;
- Comes with a programming language with a high-level syntax that reflects the theoretical assumptions.

A cognitive architecture moves beyond isolated capabilities, as it aims to provide a *unified* account of the mind.

Assumptions of Cognitive Architectures

Most cognitive architectures incorporate key postulates from psychological theories:

- *Short-term* memories are distinct from *long-term* stores
- Memories contain *modular* elements cast as *symbol structures*
- Long-term structures are accessed through *pattern matching*
- Cognitive processing occurs in *retrieval/selection/action cycles*
- Cognition involves *dynamic composition* of mental structures
- Learning is *monotonic* and *interleaved with performance*

These assumptions are shared by many frameworks, with some also including problem-space search as a core tenet.

Example Cognitive Architectures

Some well-known cognitive architectures that share these key assumptions include:

- ACT (Anderson, 1982, 1993)
- Soar (Laird et al., 1987; Laird, 2012)
- ICARUS (Langley, Choi, & Rogers 2009)

Other architectures that share some but not all assumptions are:

- Prodigy (Minton, 1988; Veloso et al. 1995)
- CAPS (Thibadeau, 1983)
- EPIC (Kieras & Meyer, 1997)
- CLARION (Sun & Zhang, 2004)

For additional details, see Langley, Laird, and Rogers (2009).

Progress: Hybrid Representations / Processing

Early production-system frameworks like PSG and OPS2 were almost entirely symbolic.

- This was consistent with general emphasis at the time, in both AI and cognitive psychology, on symbolic processing.

But not long after, architectures like ACT, CAPS, and PRISM introduced strengths and activations.

- Later, ACT-R interpreted these numbers in decision-theoretic terms, with ICARUS and Soar adopting similar ideas.

Many modern architectures are hybrid in character rather than purely symbolic.

Progress: Learning Procedural Knowledge

Cognitive architectures had their roots in accounts of problem solving and heuristic search.

- Early work had fixed rules, but *adaptive* production systems supported learning new rules.
- Many efforts on learning search-control knowledge adopted this framework.

Architectures that have incorporated this property include Soar, Prodigy, ICARUS, ACT-R, and CLARION.

Some mechanisms focus on compiling declarative knowledge into procedural, others on aiding problem solving.

Progress: Using Large-Scale Structures

Most cognitive architectures encode long-term knowledge as condition-action rules.

- Their supports modularity automated composition, flexibility, and ease of acquisition.
- But other paradigms for intelligent systems, like frames and scripts, instead propose larger-scale structures.

A few architectures have incorporated such structures into their framework and syntax (e.g., Prodigy, ICARUS).

Still, this approach remains uncommon in the paradigm and deserves more attention from researchers.

Progress: Embodied Agency

Early cognitive architectures focused on mental capacities and were effectively disembodied; this has led to work on:

- Agents that linked cognition to sensors and effectors:
  - Robo-Soar, which controlled a mobile robot
  - An ICARUS agent for simulated urban driving
  - Soar, ACT, and ICARUS agents for computer games
- Agents that interacted with humans:
  - TacAir-Soar, which flew simulated tactical air missions
  - ACT-R/E, which lets robots carry out joint tasks with humans

Extending architectures to include interaction − both physical and social − has moved them beyond pure cognition.

Progress: Declarative / Episodic Memories

Initial cognitive architectures encoded all long-term knowledge as production rules.

- Some efforts attempted to represent static facts as rules, but the results were awkward.
- ACT introduced a separate declarative store for facts, with working memory being the active portion.

More recently, multiple architectures (Prodigy, Eureka, Soar, ICARUS) have added episodic memories of agent experience.

These extensions offer a reasonable balance between procedural and declarative content.

Challenge: Understanding / Interpretation

Traditional cognitive architectures adopted an action metaphor; rules comprise a condition side and an action side.

- This emphasis came from merging theories of problem solving with behaviorist notions of stimulus-response pairs.

But understanding sequences of connected events has received little attention from architecture researchers.

- *John needed money. He got his gun. He drove to the pawn shop.*

This requires generating explanations via abductive inference; it does not lend itself to the action metaphor.

There has been research on such problems, but not within the cognitive architecture paradigm.

Challenge: Dynamic Memory

As noted earlier, cognitive architecture has long been concerned with procedural learning.

- But other knowledge involves *categories* and their organization, which are not primarily about action.

There have been some encouraging forays into this area:

- Schank's (1982) theory of dynamic memory focused on it, but did not special a complete architecture.
- Li et al. (2012) report a refinement of ICARUS that extends its conceptual memory by defining new terms.

But we need more work in the cognitive architecture paradigm on this important topic.

Challenge: Creative Problem Solving

One of the distinctive features of human cognition is *creativity*: solving novel problems in surprising ways.

- There has been AI work on this topic, but little cast within cognitive architectures.

The two primary exceptions to this trend have been:

- EUREKA (Jones & Langley, 2005), which joined problem-space search with spreading-activation retrieval;
- CLARION (Helie & Sun, 2010), which also used activation-based retrieval, but for soft constraint satisfaction.

An especially important, but understudied, topic is *reformulating* problems to make them more tractable.

Challenge: Emotions / Metacognition

Like most AI, architecture research has focused on *intellectual* activities like planning, reasoning, and language.

But people also experience *emotions* when playing a challenging opponent or reading a moving story.

- Some (Marsella et al. 2010) has developed models of emotion using existing architectures.
- But only a few efforts (Marinier & Laird, 2007) have added them as core architectural elements.

We need more work in this area, especially as emotions relate to metacognition (Cox, 2007) to modulate other mental activities.

Challenge: Personality / Goal Reasoning

Another topic widely studied in psychology, but not in cognitive architectures, is *personality*.

Trait theory is widely used for synthetic characters, but it makes little contact with other aspects of cognition.

- Rizzo et al. (1999) reported an extension to Prodigy that models personality in terms of *priorities* on abstract goals.
- This suggests a link to *goal reasoning* (Aha et al., 2013), with personal styles determined by goal-generating rules.

If so, then personalities, like emotions, play metacognitive roles that modulate intelligent behavior.

We need more research to explore this and other promising ideas.

Peripheral Topics

I have omitted three topics that may concern some listeners:

- Sensorimotor processing is necessary to interact with the world.
  - But rats, pigeons, and roaches do this quite well; this suggests it is less central to understanding intelligence.
- Statistical learning accounts for gradual change over time.
  - But such background processes are not distinctive to humans and do not explain their often rapid learning.
- Neuroscience studies the biological underpinnings of the mind.
  - But it offers little about high-level cognition, and we can model intelligence in more abstract terms.

These are legitimate areas of research, but they are not the most important for progress on cognitive architectures.

Summary Remarks

In this talk, I reviewed the notion of a cognitive architecture and some common themes in the area.

- Some tenets (e.g., symbolic matching), are shared with other parts of AI, but others (e.g., unified theories) are distinctive.

I also examined areas in which the paradigm has progressed:

- Hybrid representations, procedural learning, large-scale structures, embodied agents, and declarative memories.

However, I also identified some open challenges for research:

- Abductive understanding, dynamic memory, creative problem solving, emotions, and personality.

The cognitive architecture movement has been very successful, but does not yet have truly unified theories of the mind.

References

Anderson, J. R. 1982. Acquisition of cognitive skill. *Psychological Review*, *89*, 369–406.

Anderson, J. R. 1993. *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Kieras, D., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction*, *12*, 391–438.

Klahr, D., Langley, P., & Neches, R. (Eds). (1987). *Production system models of learning and development*. Cambridge, MA: MIT Press.

Laird, J. E. (2012). *The Soar cognitive architecture*. Cambridge, MA: MIT Press.

Langley, P., Choi, D., & Rogers, S. (2009). Acquisition of hierarchical reactive skills in a unified cognitive architecture. *Cognitive Systems Research*, *10*, 316–332.

Langley, P., Laird, J. E., & Rogers, S. (2009.) Cognitive architectures: Re-search issues and challenges. *Cognitive Systems Research*, *10*, 141–160.

Newell, A. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Sun, R., & Zhang, X. (2004). Top-down versus bottom-up learning in cognitive skill acquisition. *Cognitive Systems Research*, 5, 63–89.

Closing Dedication

I would like to dedicate this talk to two of AI's founding fathers:

Allen Newell (1927 – 1992)

Herbert Simon (1916 – 2001)

Both were interdisciplinary researchers who contributed not only to AI but to other disciplines, including psychology.

Allen Newell and Herb Simon were excellent role models who we should all aim to emulate.