Information-Processing Psychology, Artificial Intelligence, and the Cognitive Systems Paradigm

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.

Introductory Remarks

Main Points

Early research on AI was closely linked to empirical studies of high-level cognition in humans.

This alliance produced many ideas that have been crucial to the field's long-term development.

In the past 30 years, the connection has faded, hurting our ability to build artifacts that exhibit human-like intelligence.

The *cognitive systems* movement is reestablishing these links to psychology to aid progress toward this goal.

What is Intelligence?

When we say humans are *intelligent*, we mean that they exhibit high-level cognitive abilities like:

- Carrying out complex reasoning
  - E.g., solving physics problems, proving theorems
- Drawing plausible inferences
  - E.g., diagnosing automobile faults, solving murder cases
- Using natural language
  - E.g., reading stories, engaging in extended conversations
- Solving novel, complex problems
  - E.g., completing puzzles, generating plans, designing artifacts

We do *not* mean that people can recognize familiar objects or execute motor skills, abilities they share with dogs and cats.

The Cognitive Revolution

During the 1950s and 1960s, the key breakthroughs in both AI and cognitive psychology (Miller, 2003) resulted from:

- Rejecting behaviorists' obsession with learning on simple tasks and information theory's focus on statistics;
- Studying problem solving, language understanding, and other tasks that involve *thinking* (i.e., *high-level cognition*);
- Emphasizing the central role of *mental structures and processes* in such complex behavior.

Artificial intelligence and 'information-processing' psychology were tightly intertwined during this critical period.

Lessons from
Information-Processing Psychology

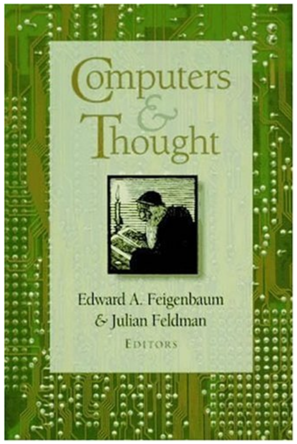
Early Links Between AI and Psychology

As AI emerged in the 1950s, a few researchers realized that computers might reproduce high-level cognition.

Some, like McCarthy and Minsky, took human intelligence as an inspiration without trying to model the details.

Others, like Herb Simon and Allen Newell, viewed themselves as psychologists aiming to explain human thought.

Carnegie Tech pursued this paradigm most vigorously, but it was also respected elsewhere.



This approach was represented in the edited volume *Computers and Thought* (Feigenbaum & Feldman, 1963).

Symbolic Structures and Processes

1

The insight behind AI was that computers (and people) are not mere number crunchers; they are *general symbol manipulators*.

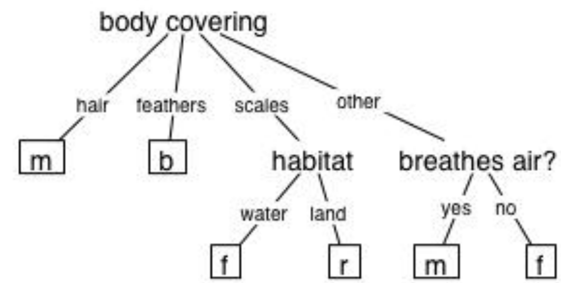
- This requires ways to *represent* symbol structures, to *interpret* such structures, and to *manipulate* them;
- These often take the form of *list structures* that can encode logic or logic-like relations;
- The insight came partly from detailed studies of human thinking (e.g., Newell & Simon, 1976).
- AI's six decades of progress has relied largely on advances in symbolic notations and mechanisms that operate on them.

Recent excitement about statistical techniques has not made this insight any less valid or important.

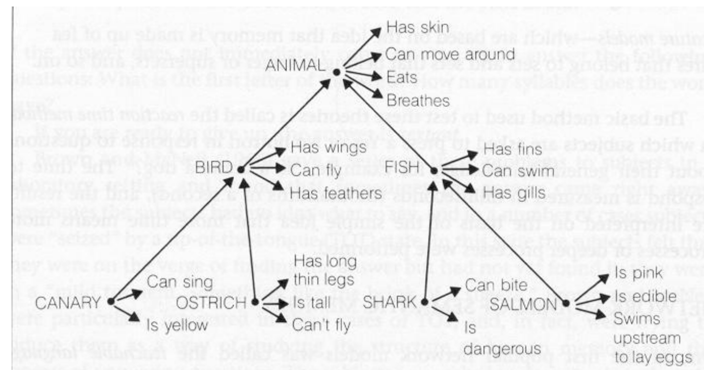
Research on Knowledge Representation

1

Early work on representation often dealt with the structure and organization of human knowledge:



- Hovland/Hunt's (1960) decision trees
- Feigenbaum's (1963) discrimination nets
- Quillian's (1968) semantic networks
- Minsky's (1975) frames
- Schank and Abelson's (1977) scripts



Not all research was motivated by psychological concerns, but it had a strong impact on the field.

Problem Solving as Heuristic Search

2

Human intelligence includes the ability to solve novel problems.

Newell and Simon's studies of think-aloud protocols led them to propose the *heuristic search* hypothesis:

- A problem solver represents states, actions, and solution paths as *symbol structures*;
- Problem solving involves a *search process* that generates and modifies these structures;
- The problem solver *evaluates* alternatives to determine whether they are desirable or acceptable.

This process is *heuristic* because, in practice, one cannot search large problem spaces exhaustively.

Research on Problem Solving

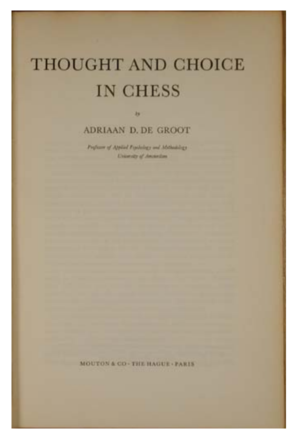
2

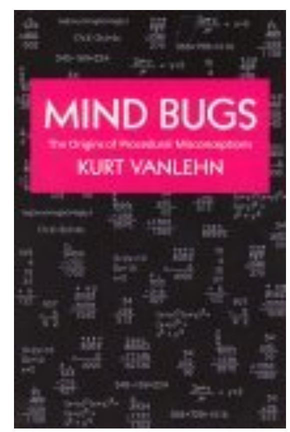
Studies of human problem solving have had a major influence on AI research:

- Newell, Shaw, and Simon’s (1958) Logic Theorist
- Newell, Shaw, and Simon’s (1961) General Problem Solver
- de Groot’s (1965) discovery of progressive deepening
- VanLehn’s (1980) analysis of impasse-driven errors

Psychological studies led to key insights about both state-space and goal-directed heuristic search.

These ideas are still widely used in AI planning and game playing.





Knowledge and Intelligence

3

Another key insight is that intelligence benefits from the ability to draw on substantial *knowledge* about:

- *Concepts* and *relations* that let one describe situations;
- *Procedures* and *skills* that let one achieve goals; and
- *Heuristics* and *constraints* that let one guide search.

This idea led to the first widespread application of AI technology in commerce and industry.

The movement was linked closely to psychological studies of *human expertise* (e.g., Chase & Simon, 1973).

Rule-Based Systems

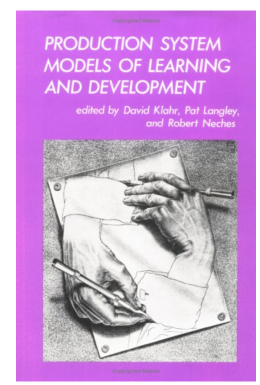
4

Many AI systems have been written in rule-based programming languages that:

- Specify behavior entirely in terms of *if-then rules*;
- Emphasize the *conditional* nature of behavior;
- Utilize *list structures* and relational *pattern matching*; and
- Support coding of highly *flexible* behaviors.

Rule-based formalisms have many practical applications and led to many successful AI systems.

One important framework – *production systems* – came directly from studies of human cognition (Newell, 1973).



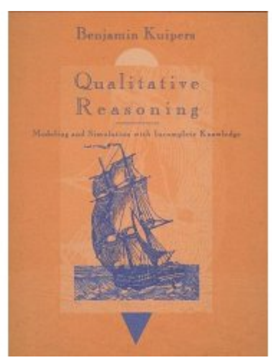
3 4

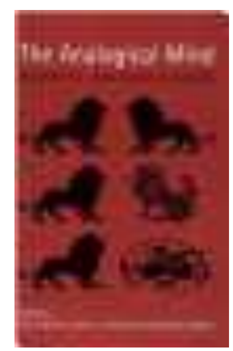
Knowledge-Based Systems

The 1980s saw multiple developments in knowledge-based reasoning that incorporated ideas from psychology:

- Expert systems (e.g., Waterman, 1986)
- Qualitative physics (e.g., Kuipers, 1984; Forbus, 1984)
- Model-based reasoning (e.g., Gentner & Stevens, 1983)
- Analogical reasoning (e.g., Gentner & Forbus, 1991)

Research on natural language also borrowed many ideas from studies of structural linguistics.





Learning and Discovery

Early machine learning systems also modeled human learning and discovery:

- Categorization (Hovland & Hunt, 1960; Fisher, 1987)
- Problem solving (Anzai & Simon, 1979; Anderson, 1981; Jones & VanLehn, 1994)
- Natural language (Reeker, 1976; Anderson, 1977; Berwick, 1979)
- Discovery in mathematics / science (Lenat, 1977; Langley, 1981)

These built on earlier insights about representation, knowledge, and heuristic search.

They were concerned with acquisition of cognitive *structures*, not with tuning statistical *parameters* (Langley, 2016).

The Cognitive Systems Paradigm

The Shift and Its Causes

Many AI researchers have now abandoned the insights of the cognitive revolution. Why did this happen?

- Commercial successes of ‘niche’ AI
  - Encouraging focus on narrow problems
- Faster processors and larger memories
  - Favoring blind search and statistical schemes
- Obsession with quantitative metrics
  - Encouraging mindless ‘bakeoffs’
- Formalist trends imported from computer science
  - Favoring simple tasks with optimality guarantees

Together, these have drawn many researchers’ attention away from AI’s original vision.

The Cognitive Systems Movement

However, the original problems remain and some researchers are committed to pursuing them.

Because "AI" now has such limited connotations, we will refer to *cognitive systems* as the paradigm that:

- *Designs, constructs, and studies computational artifacts that exhibit human-like intelligence*.

Brachman and Lemnios (2002) promoted this term for their DARPA funding initiative in the area.

See *Advances in Cognitive Systems* (http://www.cogsys.org/).

We can distinguish the cognitive systems movement from most current AI work by five characteristics.

Feature 1: Focus on High-Level Cognition

One key feature of the cognitive systems movement lies in its emphasis on *high-level cognition*.

People share with dogs and cats their abilities for perception, categorization, and empirical learning, but only humans can:

- Understand and generate language
- Solve novel and complex problems
- Design and use complex artifacts
- Reason about others’ mental states
- Think about their own thinking

Computational replication of these abilities is the key charge of cognitive systems research.

Feature 2: Structured Representations

Another aspect of cognitive systems research is its reliance on *structured representations* and *knowledge*.

The insight behind the 1950s AI revolution was that computers are not mere number crunchers.

Computers and humans are *general symbol manipulators* that:

- Encode content as list structures or similar formalisms
- Create, modify, and interpret this relational content
- Utilize numbers mainly as annotations on these structures

The paradigm assumes that representing, and reasoning over, rich symbolic structures is key to human-level cognition.

Feature 3: Influence of Human Cognition

Research on cognitive systems draws ideas and inspiration from information-processing psychology.

Many of AI's early insights came from studying human problem solving, reasoning, and language use, including:

- How people represent knowledge, goals, and beliefs
- How humans utilize knowledge to draw inferences
- How people acquire new knowledge from experience

We still have much to gain from this strategy, even when our artifacts differ in their operational details.

Human capabilities also offer *challenges* for cognitive systems researchers to pursue.

Feature 4: Heuristics and Satisficing

Another important assumption of cognitive systems work is that intelligence relies on *heuristic methods* that:

- Are not guaranteed to find the best or even *any* solution but
- Greatly reduce search and make problem solving tractable
- Apply to a broader range of tasks than methods with guarantees

They mimic high-level human cognition in that they *satisfice* by finding acceptable rather than optimal solutions.

Much of the flexibility in human intelligence comes from its use of heuristic methods.

Feature 5: Systems Perspective

Finally, the paradigm is distinctive in approaching intelligence from a *systems perspective*.

While most AI efforts idolize component algorithms, work on cognitive systems is concerned with:

- How different intellectual abilities fit together and interact
- Integrated intelligent agents that combine these capabilities

Such systems-level research provides an avenue to artifacts that exhibit the breadth and scope of human intelligence.

Otherwise, we will be limited to the *idiot savants* so popular in academia and industry.

Cognitive Architectures

Beyond Component Algorithms

5

A *cognitive architecture* (Newell, 1990) is a unified theory of mental abilities that:

- Moves beyond isolated abilities to support complete models of intelligent behavior;
- Specifies facets of cognition that are constant across different domains (memories / representations but *not* their content);
- Provides a programming language with a high-level syntax that reflects strong theoretical assumptions.

A cognitive architecture is all about *mutual constraints*, as it aims to provide a unified account of the mind.

Assumptions of Cognitive Architectures

5

Most cognitive architectures (e.g., ACT-R, Soar) 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 claims are widely adopted by researchers who study high-level cognition in humans.

The ICARUS Architecture

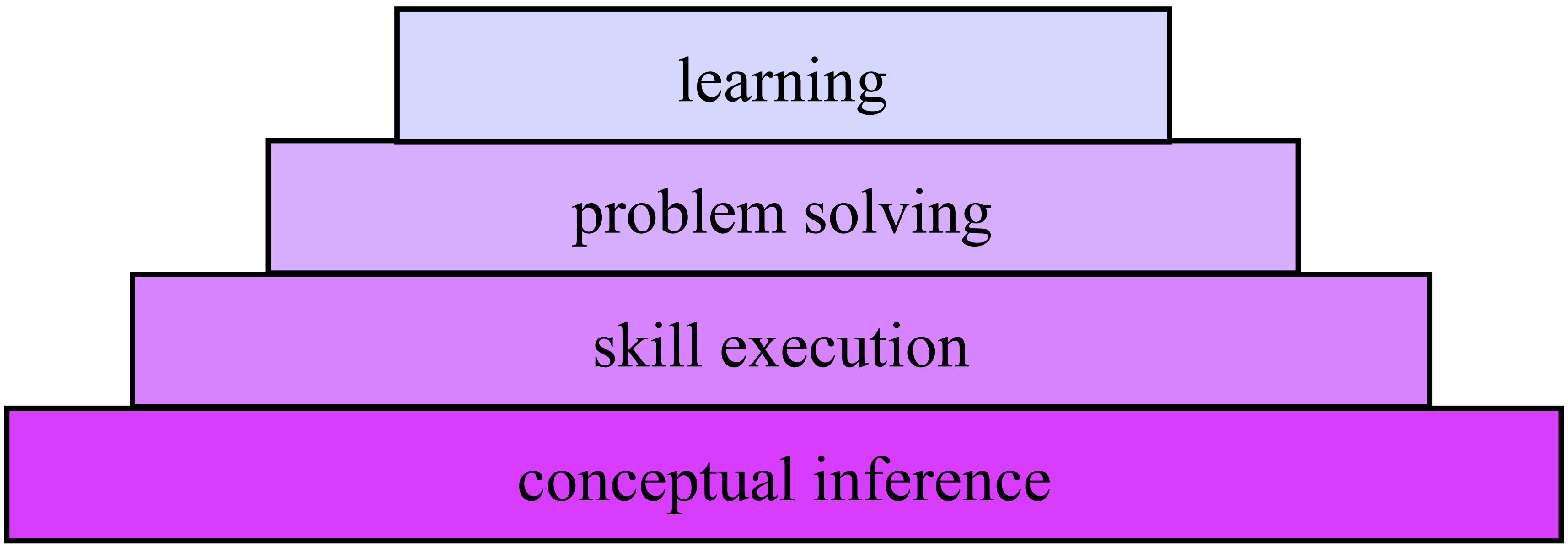
ICARUS (Langley, Choi, & Rogers, 2009) is a distinctive cognitive architecture that assumes:

1. Cognition is grounded in perception and action
2. Categories and skills are separate cognitive entities
3. Short-term elements are instances of long-term structures
4. Long-term knowledge is organized in a hierarchical manner
5. Inference and execution are more basic than problem solving

Some of these claims also appear elsewhere, but only ICARUS combines them into a unified cognitive theory.

Cascaded Integration in ICARUS

Like other unified cognitive architectures, ICARUS incorporates a number of distinct modules.



ICARUS adopts a *cascaded* approach to integration in which lower-level modules produce results for higher-level ones.

Theory of Conceptual Inference

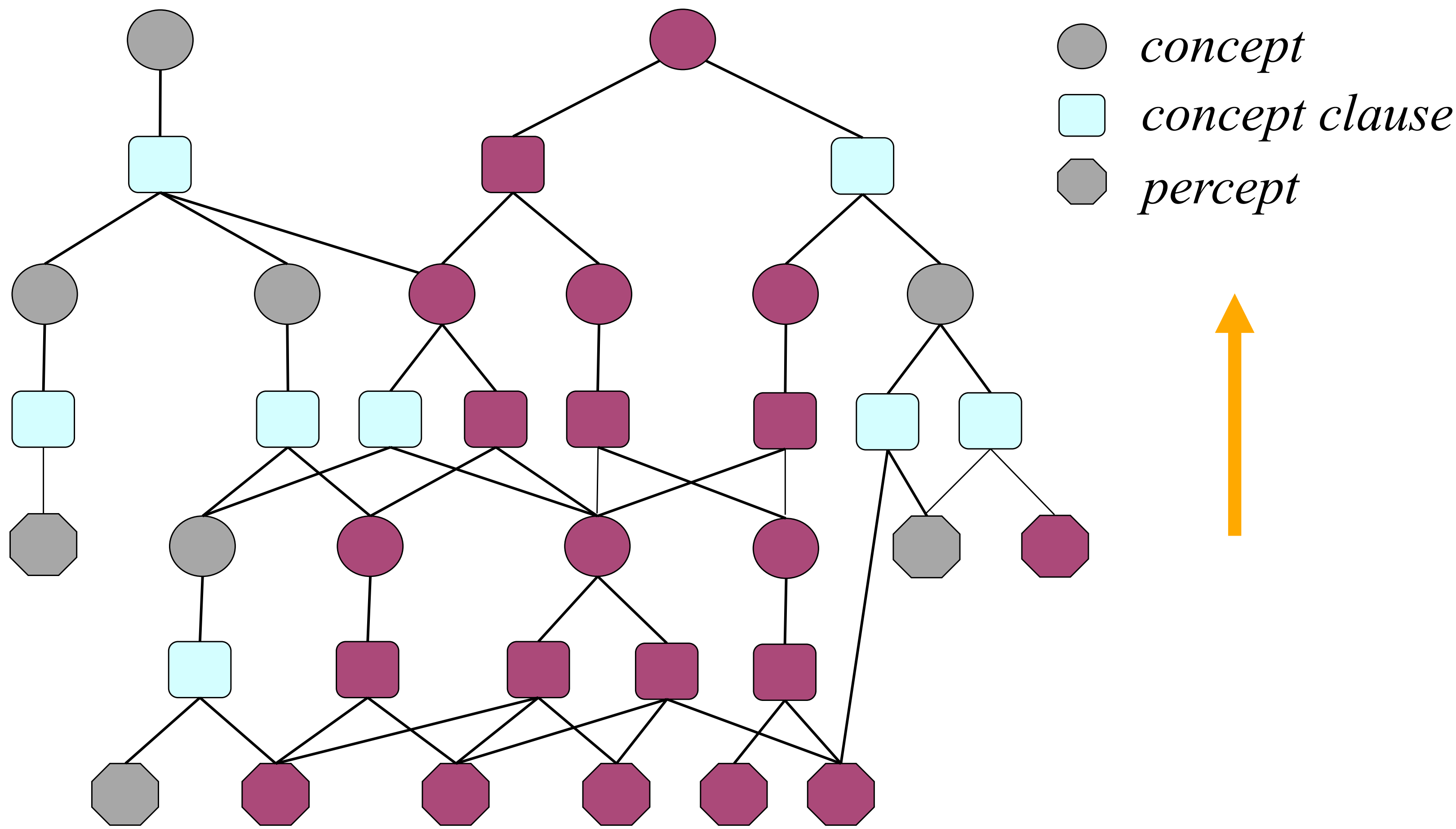
Concepts are distinct cognitive entities that humans use to describe their environment:

- Most categories are *grounded* in perception, in that they refer to the physical characteristics of objects or events.
- Many concepts are *relational*, in that they describe connections or interactions among objects or events.
- Concepts are organized in a *hierarchy*, with more complex categories defined in terms of simpler structures.
- Everyday conceptual inference is an automatic process that proceeds in a *bottom-up* manner.

ICARUS incorporates and instantiates these assumptions about conceptual structures and processing.

Conceptual Inference in ICARUS

Conceptual inference in ICARUS occurs from the bottom up.



Starting with observed percepts, this process produces high-level beliefs about the current state.

Theory of Skill Execution

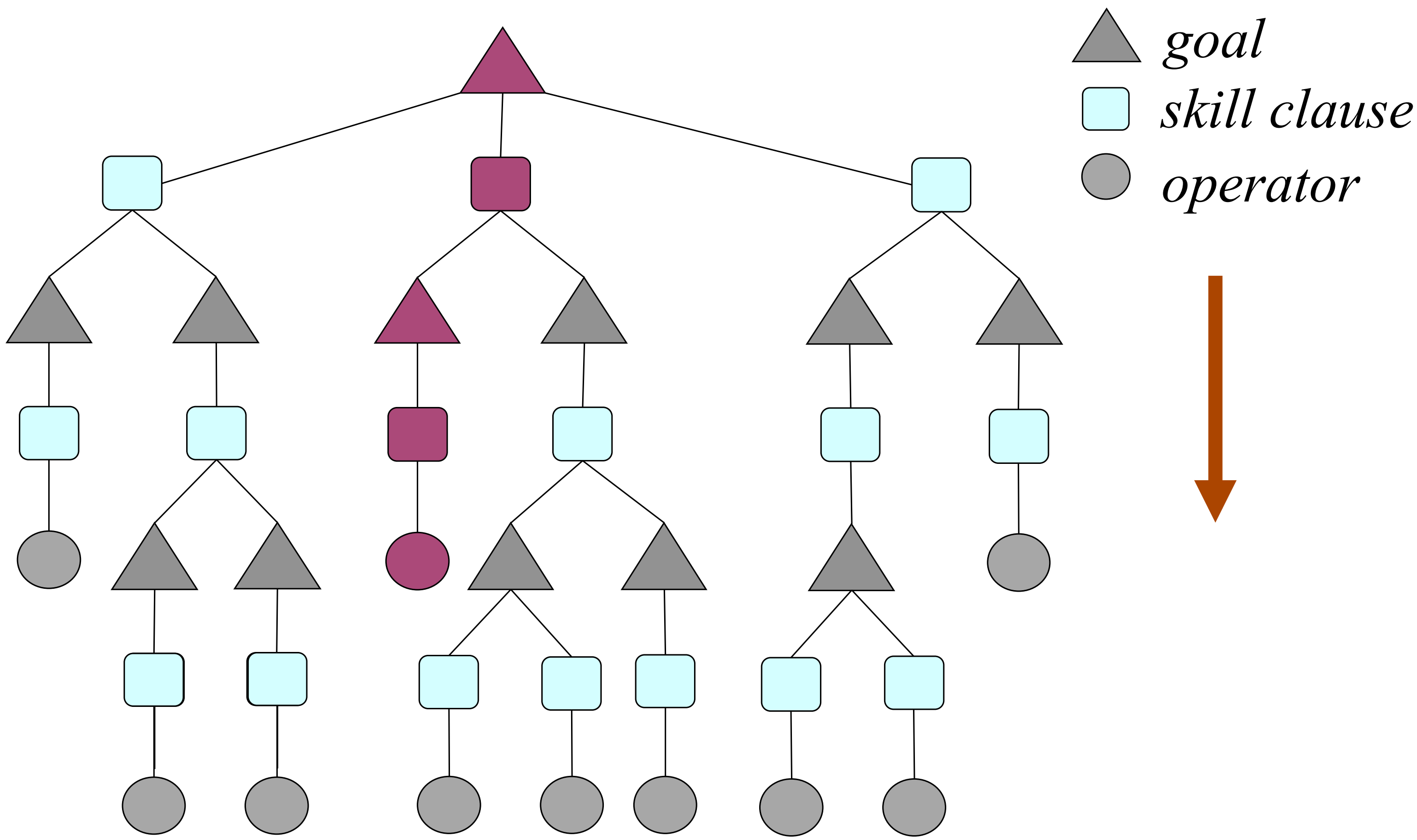
Skills are distinct cognitive structures that describe how one interacts with the environment:

- Most human skills are *grounded* in perception (indirectly through concepts) and in action.
- Skills are *relational* in that they describe changes in conceptual structures as a result of their execution.
- Memory for skills is organized as a *hierarchy*, with more complex activities decomposed into simpler ones.
- Skills are *indexed* by goals they achieve on successful execution in the environment.
- Execution is *teleoreactive*, i.e., guided by the agent's goals but sensitive to environmental factors.

ICARUS incorporates and instantiates these assumptions about skill representation and processing.

Skill Execution in ICARUS

ICARUS executes skills from the top down, starting from goals, to find applicable paths through the skill hierarchy.



When it cannot find an applicable path, it falls back on problem solving to generate a novel hierarchical plan.

Theory of Problem Solving

Problem solving lets humans achieve goals even on complex, unfamiliar tasks:

1. The search process relies on *means-ends analysis,* a mix of goal-directed backward chaining and state-driven forward chaining.
2. Problem solving remains *grounded* in perception and actions, yet often occurs at an *abstract* level of description.
3. Such problem solving typically *interleaves* mental processing with physical execution.
4. Learned skills, which are acquired from individual experiences, are *generalized traces* of successful means-ends analysis.

ICARUS adopts and utilizes these tenets about the components and operation of problem solving.

ICARUS Summary

ICARUS is a theory of the cognitive architecture that supports:

- conceptual inference over grounded relational categories
- goal-directed but reactive execution of hierarchical skills
- means-ends problem solving when routine execution fails
- acquisition of new skills from traces of problem solving

The theory is consistent with many findings about how humans represent, use, and learn knowledge.

We have also used ICARUS to develop synthetic game agents, simulated urban drivers, and robotic controllers.

Closing Remarks

Mind and Brain

Many people identify the *mind* with the *brain*, then assume that we cannot understand the former without the latter.

- But theories of the mind can be independent of the hardware or wetware on which they operate.
- The same computer program typically runs on entirely different computer architectures and operating systems.
- Quantum physics may underlie chemistry, yet chemists seldom use it in theory or practice.

These involve *different levels of description*. Reductionism may sound promising, but it is not a practical scientific strategy.

AI and Neuroscience

Neuroscience has made great strides in the past 50 years, but it still has little to say about how we:

- Represent beliefs, goals, or knowledge in mental structures;
- Use such structures for multi-step reasoning, problem solving, and language processing;
- Acquire these structures rapidly, from only a few experiences.

Most results have focused on perception and action, not on abilities that give us human-level intelligence.

Neuroscience may provide insights about the mind, but AI has made great progress without it, and this will continue.

AI and Cognitive Psychology

In contrast, much early AI research was inspired by, and gained insights from, studies of human thinking.

This link has produced many of the most powerful ideas about the computational character of the mind:

- *Symbol structures and processing*
- *Heuristic search in problem solving*
- *Knowledge, expertise, and rule-based systems*
- *Unified cognitive architectures*

Artificial intelligence can grow even stronger by drawing on its deep psychological roots.

References

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55–81.

Feigenbaum, E. A., & Feldman, J. (Eds.) (1963). *Computers and thought*. New York: McGraw-Hill.

Langley, P. (2012b). The cognitive systems paradigm. *Advances in Cognitive Systems*, *1*, 3–13.

Langley, P. (2016). The central role of cognition in learning. *Advances in Cognitive Systems*, *4*, 3–12.

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

Newell, A. (1973). You can't play 20 questions with nature and win. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Newell, A., and Simon, H. A. (1976). Computer science as empirical enquiry: Symbols and search. *Communications of the ACM*, *19*, 113–126.

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *TRENDS in Cognitive Sciences*, *7*, 141–144.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision*. New York: McGraw-Hill.

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.