Intelligent Behavior in Humans and Machines

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 AI was closely linked to the empirical study of high-level cognition in humans.

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

In contrast, neuroscience has told us comparatively little except about low-level perception and action.

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

Re-establishing AI's links to cognitive psychology would aid its progress toward this challenging objective.

What is Artificial Intelligence?

Artificial intelligence is the *computational study of structures and processes that support intelligent behavior*.

The name reflects a focus on creating computational *artifacts*.

- Early researchers hoped to create systems with broad, general mental abilities that applied in many settings.
- In recent years, most AI practitioners have adopted narrower goals with shorter-term payoffs.

However, these approaches share many assumptions about the principles that underlie intelligence.

Many of these principles come originally *from studies of high-level cognition in humans*.

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 cognitive psychology were tightly intertwined during this critical period.

Lessons from Cognitive Psychology

Early Links Between AI and Psychology

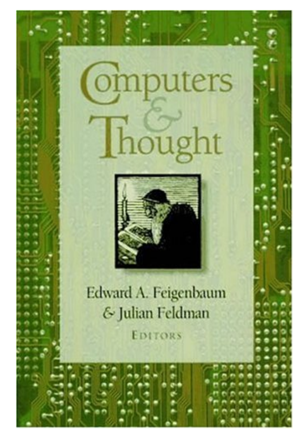
As AI emerged in the 1950s, a few researchers realized that computers might reproduce high-level cognition.

Some 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 number crunchers; they are *general symbol manipulators*.

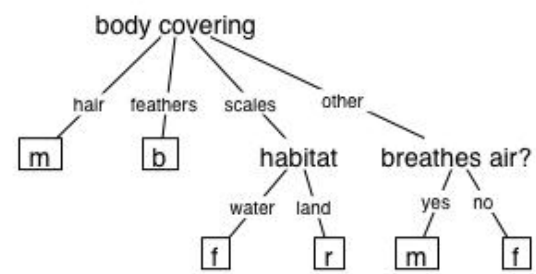
- This requires ways to *represent* symbol structures, to *interpret* such structures, and to *manipulate* them;
- These often take the form of *list structures* that can encode logic or logic-like relations;
- The insight came partly from detailed studies of human thinking (e.g., Newell & Simon, 1976).
- 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.

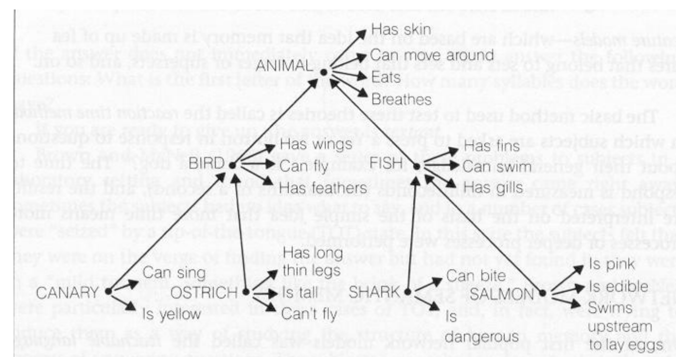
Research on Knowledge Representation

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 many problem spaces exhaustively.

2

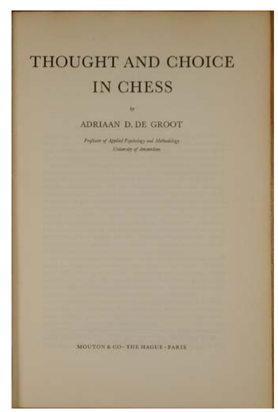
Research on Problem Solving

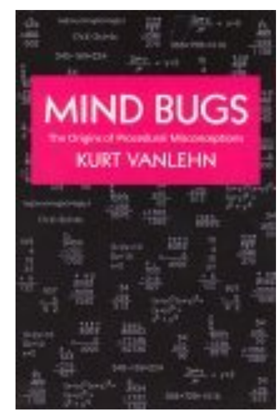
Studies of human problem solving have had a major influence on AI research:

- Newell, Shaw, and Simon’s (1958) Logic Theorist
- Newell, Shaw, and Simon’s (1961) General Problem Solver
- de Groot’s (1965) on progressive deepening in chess
- 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
- *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

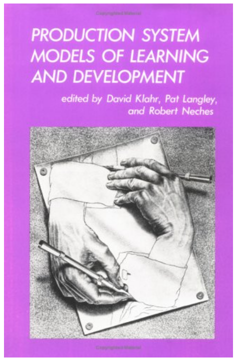
4

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*
- 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).



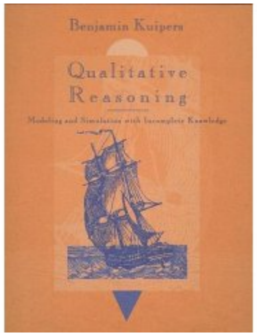
3 4

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 Unbalanced State of Modern AI

In recent decades, AI has moved away from modeling human cognition and become unfamiliar with results in psychology.

Despite the historical benefits, many AI researchers now believe psychology has little to offer it.

Similarly, few psychologists believe that results from AI have relevance to modeling human behavior.

This shift has taken place in many subareas, and it has occurred for a number of reasons.

The Shift and Its Causes

Why have many AI researchers abandoned the insights of the cognitive revolution?

- 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 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.

Distinguishing Characteristics

The cognitive systems paradigm (Langley, 2012) differs from most recent AI work in that it:

- Focuses on *high-level* cognitive tasks
- Uses *structured representations of knowledge*
- Adopts a *systems perspective* on intelligence
- Draws inspiration from results on *human cognition*
- Relies on *heuristic methods* and *satisficing*

The subfield also encourages *exploratory research* on novel problems, in the same spirit as early AI research.

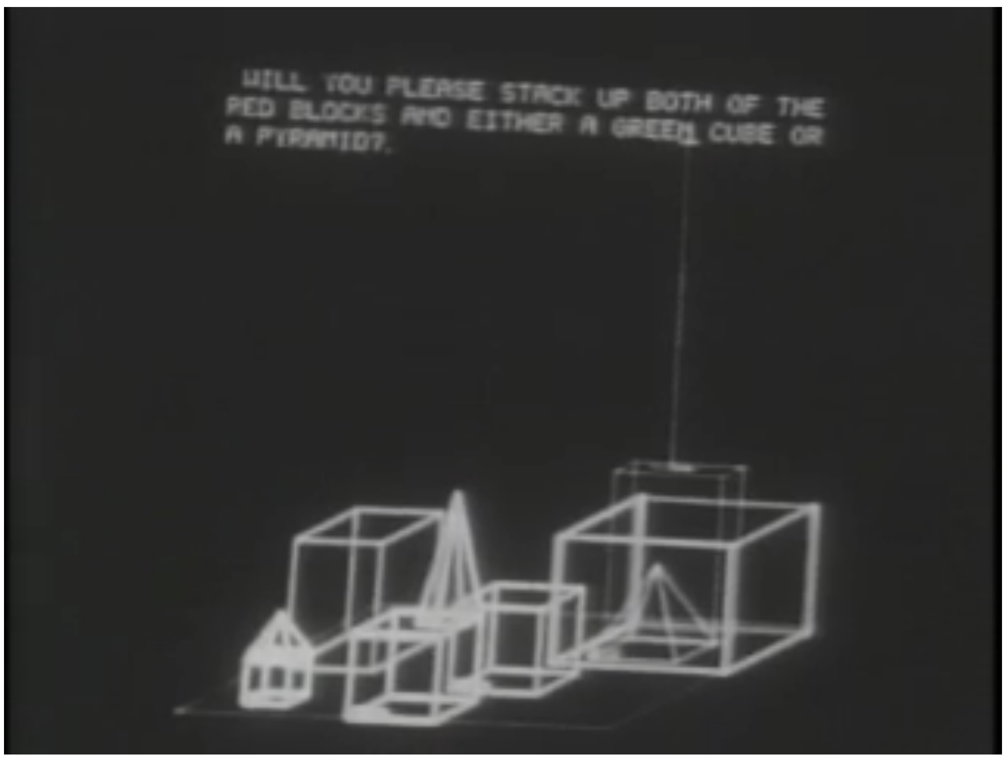
See *Advances in Cognitive Systems* (http://www.cogsys.org/).

Compelling Cognitive Systems

SHRDLU (1970)

SHRDLU (1970) was an early AI system that answered questions in English and executed complex commands.

Although limited in scope, it reproduced many aspects of human dialogue.



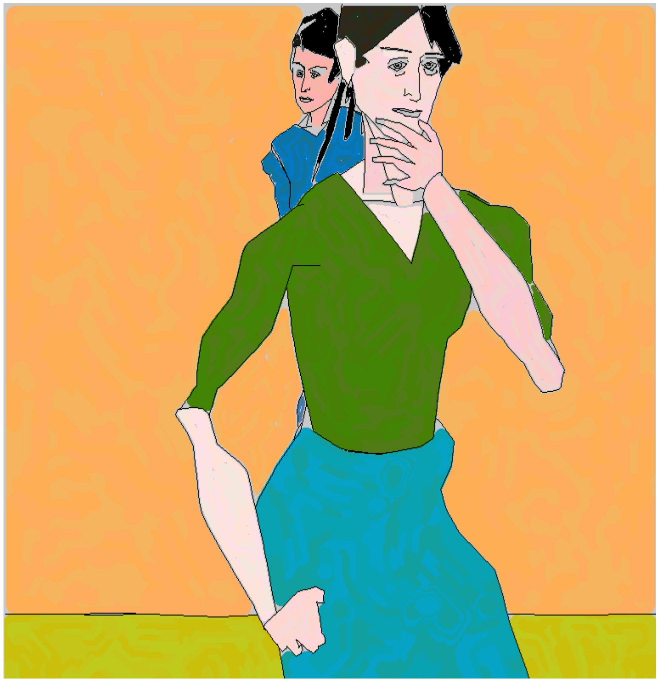
Person: Pick up a big red block.
Computer: *OK.*
Person: Grasp the pyramid.
Computer: *I don't understand which pyramid you mean.*
Person (changing his mind): Find a block which is taller than the one you are holding and put it into the box.
Computer: *OK.*
Person: What does the box contain?
Computer: *The blue pyramid and the blue block.*
Person: What is the pyramid supported by?
Computer: *The box.*
Person: How many blocks are not in the box?
Computer: *Four of them.*
Person: Is at least one of them narrower than the one which I told you to pick up?
Computer: *Yes, the red cube.*

Aaron (1973–2016)

The Aaron system composes and physically paints novel art work.

In some sense, it is rule-based expert system that operates in an area we usually associate with creativity.

But it integrates many different facets of artistic composition and has even used a robot arm to implement its designs.



Deep Thought / Blue (1985–1997)

Deep Blue – originally Deep Thought – took a cognitive systems approach to support chess play at the grandmaster level.

- Despite common beliefs, the program did not rely on ‘brute force’ to compete in this challenging game.
- Rather, the system combined substantial domain knowledge with heuristics to guide search through a large space.

In 1997, Deep Blue won a match against Gary Kasparov, then the world champion, taking 3.5 to 2.5 games.

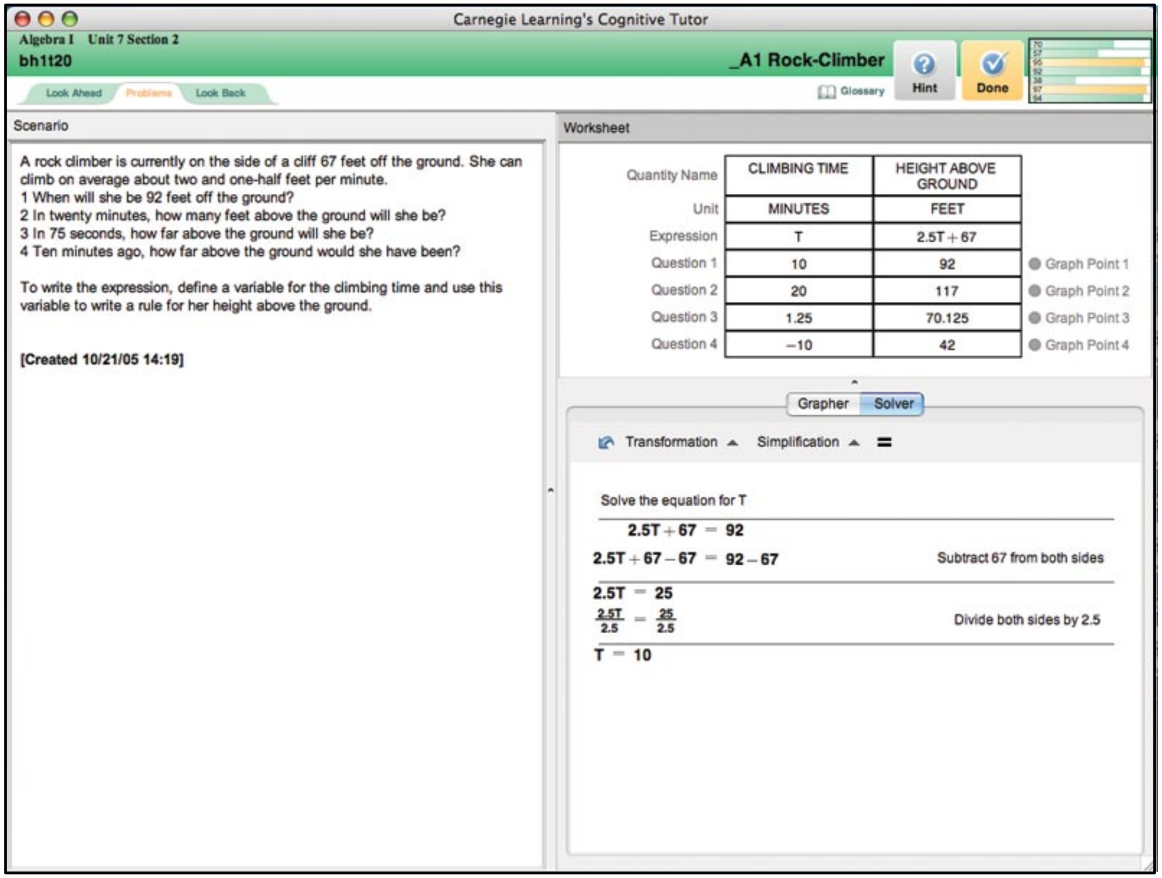
Over the past two decades, similar advances have occurred for checkers, backgammon, and other common games.

Carnegie Learning's Algebra Tutor (1999–present)

This tutor encodes knowledge about algebra as production rules, infers models of students' knowledge, and provides personalized instruction to them.

The system has been adopted by hundreds of US middle schools.

Studies have shown that it improves student learning in this domain by 75 percent.

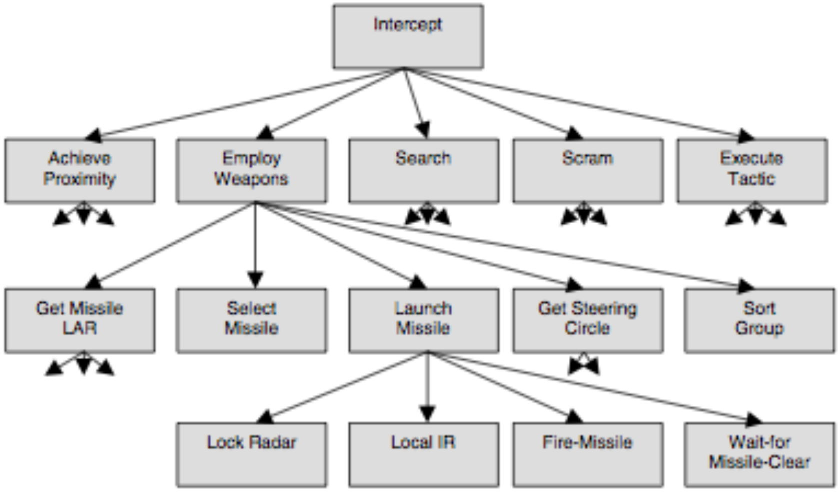


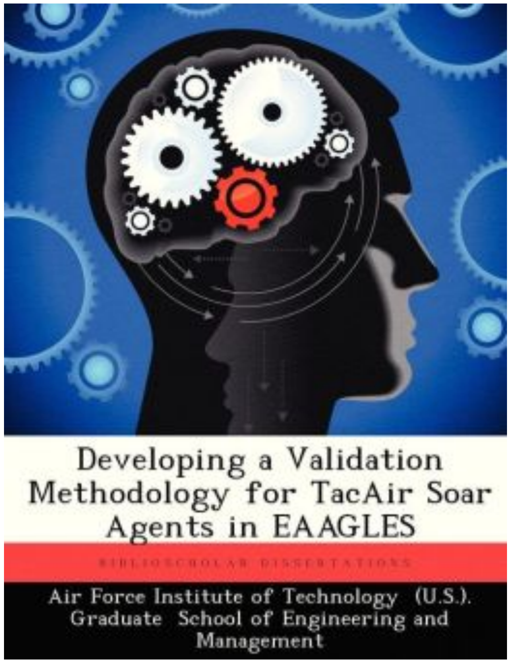
TacAir-Soar (1992–1997)

The TacAir-Soar system reproduces pilot behavior in tactical air combat.

It combines abilities for spatio-temporal reasoning, plan generation / recognition, language, and coordination.

The system flew 722 missions during the STOW-97 simulated training exercise.





Façade (2003–2007)

Mateas and Stern's Façade is a graphical environment in which characters interact with the user and each other.

The agents understand and generate sentences, control gaze and expression, and exhibit distinct personalities.

Façade characters use a rich knowledge base to produce inferences, carry out physical activities, and engage socially.

Some Other Examples

This does not exhaust research of impressive cognitive systems; other examples include:

- TRAINS, an interactive aid that helps users create plans through mixed-initiative spoken dialogue (Allen et al., 1996)
- COLLAGEN (Rich et al. 2001), which helps users in operating complex devices, asking questions and giving advice as needed
- Tutorial dialogue systems (Graesser et al., 2001) that converse in spoken language, giving personalized instruction.
- The Artificial Receptionist (Bohus & Horvitz, 2009), which welcomes and interacts with visitors in spoken dialogue.
- The Robot Scientist (King et al., 2009) combines experiment design and execution with model revision in cell biology.

These diverse systems show the range of possible applications.

Benefits of Renewed Interchange

Benefit: Understanding Human Cognition

One reason for renewed interchange is to better understand the nature of the human mind.

- This would have important societal applications in education, interface design, and other areas;
- Human intelligence comprises an important set of phenomena that demand scientific explanation.

This remains an open and challenging problem, and AI systems remain a promising way to tackle it.

Benefit: Source of Challenging Tasks

Another reason is that human abilities can serve as a source of challenges, such as:

- Understanding language at a deeper level than current systems
- Interleaving planning with execution in pursuit of many goals
- Learning complex knowledge structures from few experiences
- Carrying out creative activities in the arts and sciences

Most AI sets its sights too low by focusing on tasks that require limited intelligence.

Psychological studies reveal the impressive abilities of humans and pose new problems for AI research.

Benefit: Constraints on Intelligent Artifacts

To develop intelligent systems, we must constrain their design, and findings about human behavior can suggest:

- How the system can represent and organize knowledge
- How the system can use that knowledge in performance
- How the system can acquire knowledge from experience

We can use psychological ideas as *design heuristics*, including abilities *not* needed (e.g., carrying out extensive search).

Humans are still our only example of general intelligence, and insights into their operation merit serious consideration.

Beyond Component Algorithms

5

Newell (1973) argued that "*You can't play twenty questions with nature and win*", proposing instead that we:

- Move beyond isolated phenomena and capabilities to develop complete models of intelligent behavior;
- Develop cognitive systems that make strong theoretical claims about the nature of the mind;
- Combine ideas from cognitive psychology with rigorous AI methods to implement these accounts.

Newell claimed that a successful framework would provide a *unified* theory of intelligent behavior.

He attached these aims to the idea of a *cognitive architecture*.

Assumptions of Cognitive Architectures

5

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 *recognize-act cycles*
- Cognition involves *dynamic composition* of mental structures
- Learning is *monotonic* and *interleaved with performance*

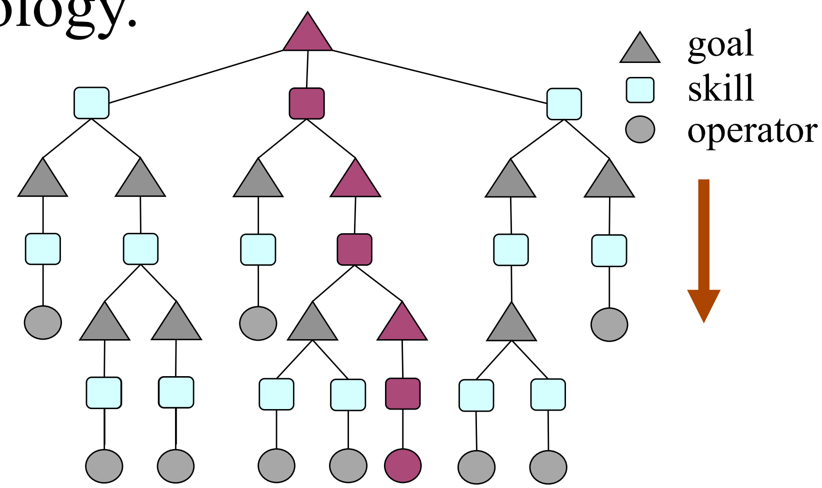
These claims are shared by a variety of frameworks, including ACT-R, Soar, Prodigy, and ICARUS.

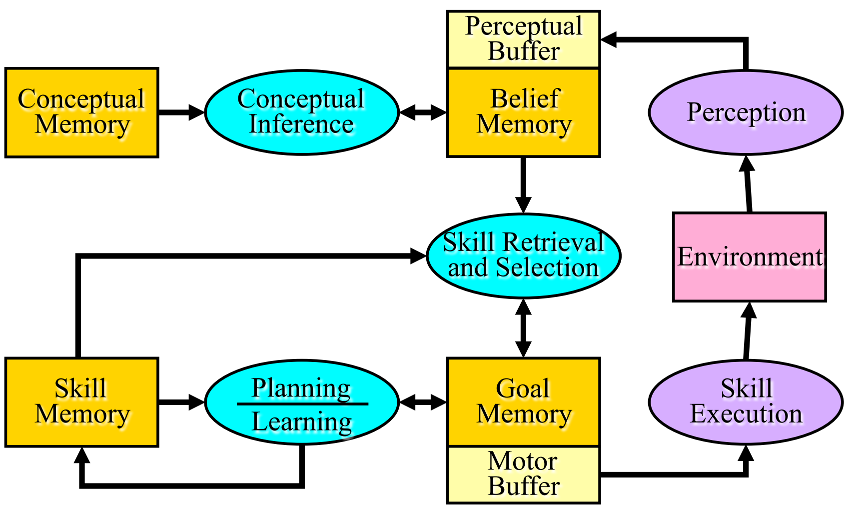
The ICARUS Cognitive Architecture

ICARUS (Langley et al., 2009) is a unified theory of intelligent behavior that borrows ideas from psychology.

Theoretical postulates:

1. Cognition grounded in perception/action
2. Categories and skills in distinct memories
3. Short-term elements instances of long-term
4. Knowledge organized in hierarchies



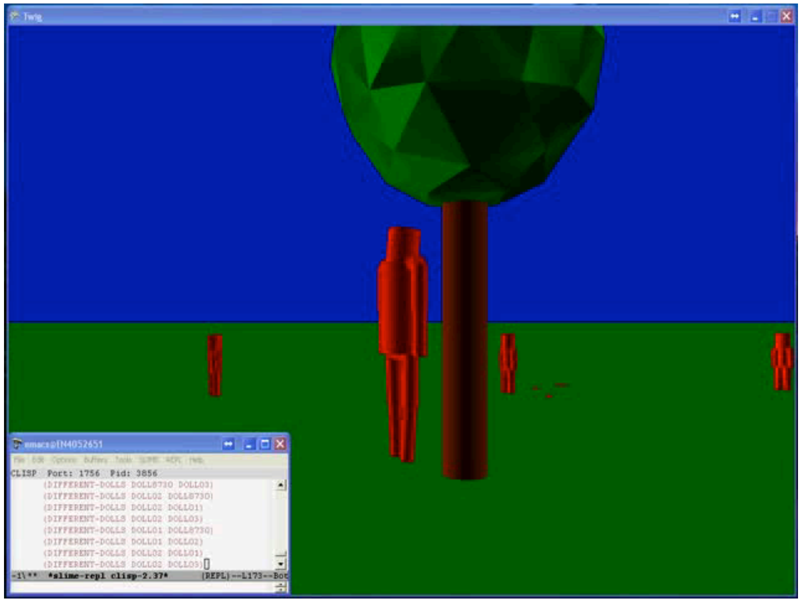



```
((driving-well-in-rightmost-lane ?self ?line1 ?line2)
 :percepts     ((self ?self) (segment ?seg) (line ?line1 segment ?seg))
                (line ?line2 segment ?seg))
 :start        ((not (lane-to-right ?line1 ?line2 ?anyline))
 :subgoals     ((driving-well-in-segment ?self ?seg ?line1 ?line2)))

((driving-well-in-segment ?self ?seg ?line1 ?line2)
 :percepts     ((self ?self) (segment ?seg) (line ?line1 segment ?seg))
                (line ?line2 segment ?seg))
 :start        ((steering-wheel-straight ?self))
 :subgoals     ((in-segment ?self ?seg)
                (aligned-and-centered-in-lane ?self ?line1 ?line2)
                (steering-wheel-straight ?self)))
```

Synthetic Agents in ICARUS

Twig



Urban Combat

Urban Driving

Rush 2008

AI and Other Disciplines

AI and Robotics

AI is often linked to robotics, which develops embodied artifacts that operate in the physical world, but:

- Robotics is concerned with sensori-motor behavior, which humans share with other animals;
- Many robotic systems exhibit little or no intelligence in the sense discussed earlier;
- Most applications (vacuum cleaners, walking robots, self-driving cars) focus on low-level control tasks.

There remains potential for interactions, but AI and robotics have remained largely separate.

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.