Cognitive Architecture: Past, Present, Future

John Laird

28th Soar Workshop

University of Michigan

May 7, 2008

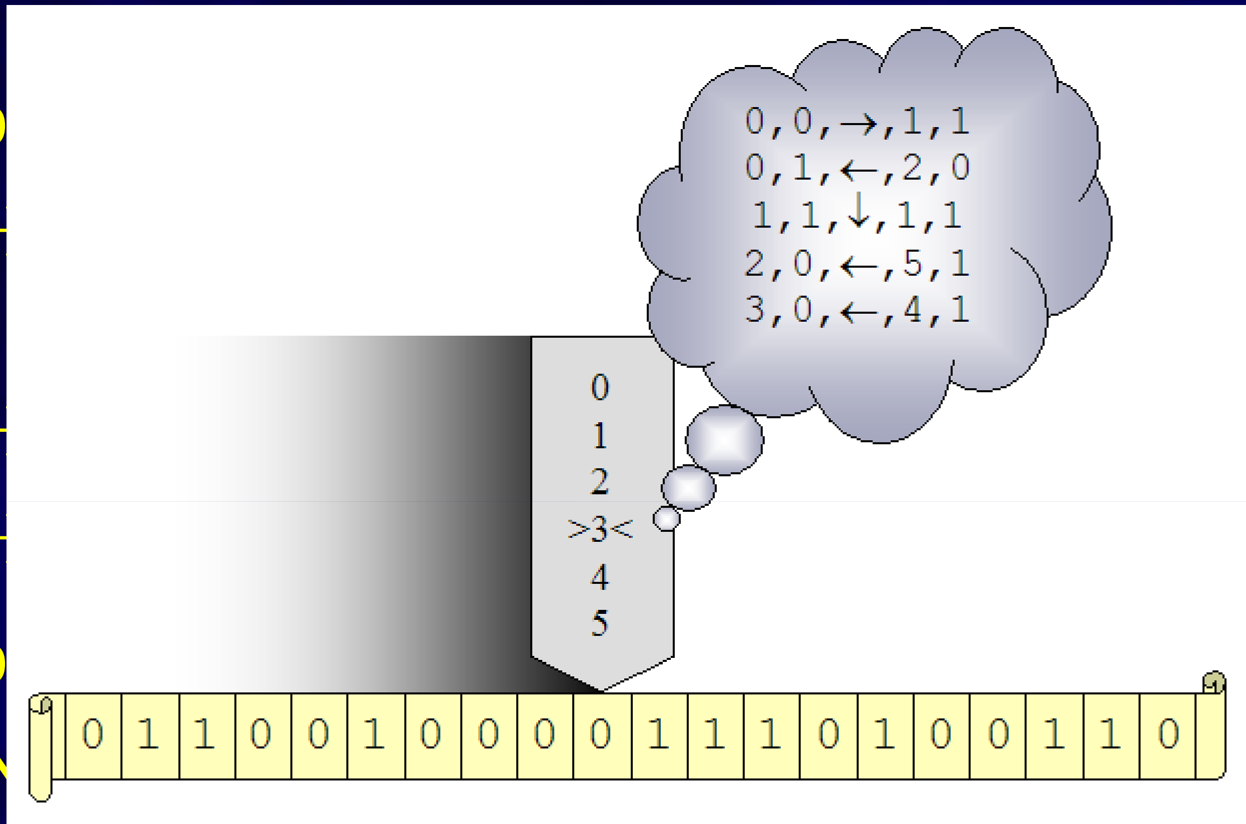
What is Cognitive Architecture?

Fixed mechanisms and structures that underlie cognition

- Processors that manipulate data
- Memories that hold knowledge
- Interfaces that interact with an environment

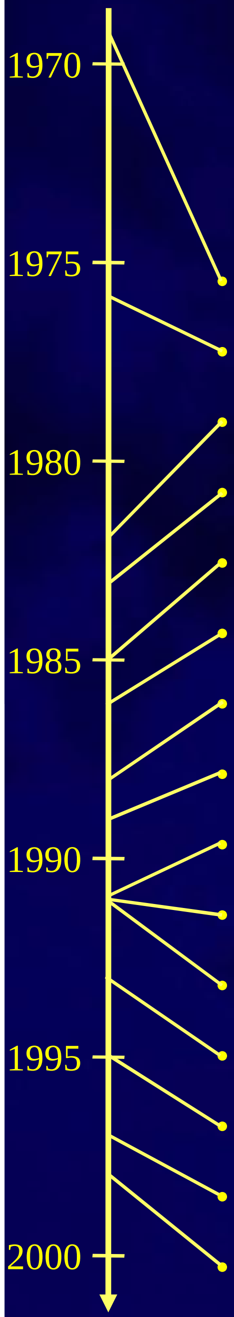
Why is Cognitive Architecture Important?

- Pro … lems
  - H …
    … g, language
  - H … l?
  - H … ned?
- Pro … ure
  - N … e



- Turing equivalence isn't sufficient
  - Architectures have different *complexity profiles*

History of Cognitive Architecture 1969-2000



- GPS (Ernst & Newell, 1969) Means-ends analysis, recursive subgoals
- ACT (Anderson, 1976) Human semantic memory
- CAPS (Thibadeau, Just, Carpenter) Production system for modeling reading
- Soar (Laird, & Newell, 1983) Multi-method problem solving, production systems, and problem spaces
- Theo (Mitchell et al., 1985) Frames, backward chaining, and EBL
- PRS (Georgeff & Lansky, 1986) Procedural reasoning & problem solving
- BB1/AIS (Hayes-Roth & Hewitt 1988) Blackboard architecture, meta-level control
- Prodigy (Minton et al., 1989) Means-ends analysis, planning and EBL
- MAX (Kuokka, 1991) Meta-level reasoning for planning and learning
- Icarus (Langley, McKusick, & Allen,1991) Concept learning, planning, and learning
- 3T (Gat, 1991) Integrated reactivity, deliberation, and planning
- CIRCA (Musliner, Durfee, & Shin, 1993) Real-time performance integrated with planning
- AIS (Hayes-Roth 1995) Blackboard architecture, dynamic environment
- EPIC (Kieras & Meyer, 1997) Models of human perception, action, and reasoning
- APEX (Freed et al., 1998) Model humans to support human computer designs

History of Cognitive Architecture 1969-2000

Psychological Modeling

1970 1975 1980 1985 1990 1995 2000

- GPS (Ernst & Newell, 1969) Means-ends analysis, recursive subgoals
- *ACT (Anderson, 1976) Human semantic memory*
- *CAPS (Thibadeau, Just, Carpenter, 1982) Production system for modeling reading*
- Soar (Laird, & Newell, 1983) Multi-method problem solving, production systems, and problem spaces
- Theo (Mitchell et al., 1985) Frames, backward chaining, and EBL
- PRS (Georgeff & Lansky, 1986) Procedural reasoning & problem solving
- BB1/AIS (Hayes-Roth & Hewitt 1988) Blackboard architecture, meta-level control
- Prodigy (Minton et al., 1989) Means-ends analysis, planning and EBL
- MAX (Kuokka, 1991) Meta-level reasoning for planning and learning
- Icarus (Langley, McKusick, & Allen,1991) Concept learning, planning, and learning
- 3T (Gat, 1991) Integrated reactivity, deliberation, and planning
- CIRCA (Musliner, Durfee, & Shin, 1993) Real-time performance integrated with planning
- AIS (Hayes-Roth 1995) Blackboard architecture, dynamic environment
- *EPIC (Kieras & Meyer, 1997) Models of human perception, action, and reasoning*
- *APEX (Freed et al., 1998) Model humans to support human computer designs*

History of Cognitive Architecture 1969-2000

Active Architectures

1970 · 1975 · 1980 · 1985 · 1990 · 1995 · 2000

- GPS (Ernst & Newell, 1969) Means-ends analysis, recursive subgoals
- *ACT (Anderson, 1976) Human semantic memory*
- CAPS (Thibadeau, Just, Carpenter, 1982) Production system for modeling reading
- *Soar (Laird, & Newell, 1983) Multi-method problem solving, production systems, and problem spaces*
- Theo (Mitchell et al., 1985) Frames, backward chaining, and EBL
- PRS (Georgeff & Lansky, 1986) Procedural reasoning & problem solving
- BB1/AIS (Hayes-Roth & Hewitt 1988) Blackboard architecture, meta-level control
- Prodigy (Minton et al., 1989) Means-ends analysis, planning and EBL
- MAX (Kuokka, 1991) Meta-level reasoning for planning and learning
- Icarus (Langley, McKusick, & Allen,1991) Concept learning, planning, and learning
- 3T (Gat, 1991) Integrated reactivity, deliberation, and planning
- CIRCA (Musliner, Durfee, & Shin, 1993) Real-time performance integrated with planning
- AIS (Hayes-Roth 1995) Blackboard architecture, dynamic environment
- *EPIC (Kieras & Meyer, 1997) Models of human perception, action, and reasoning*
- *APEX (Freed et al., 1998) Model humans to support human computer designs*

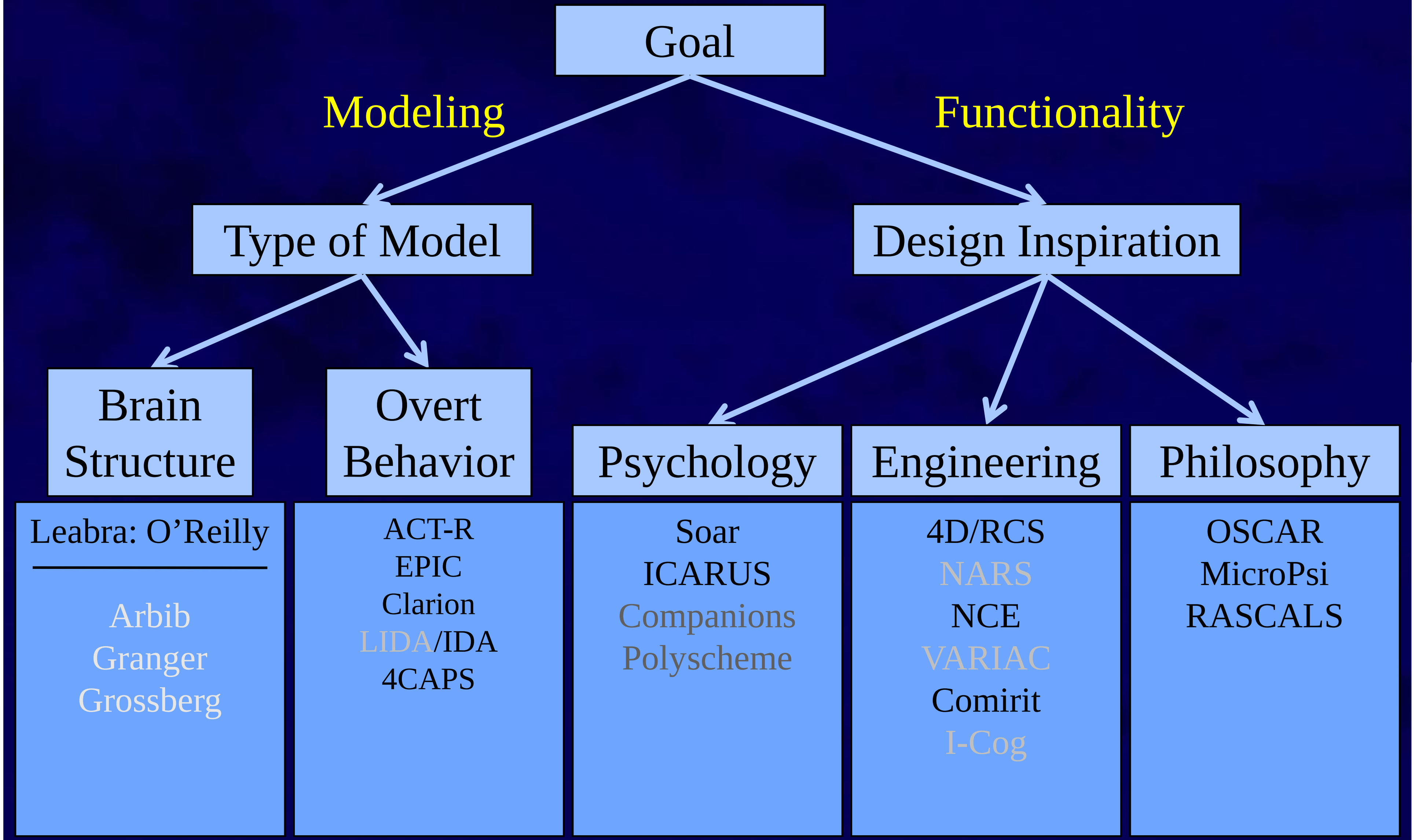
Current State of Cognitive Architecture

- Explosion of different architectures
  - Developed with different goals in mind
- Lots of different components
- But some significant commonalities

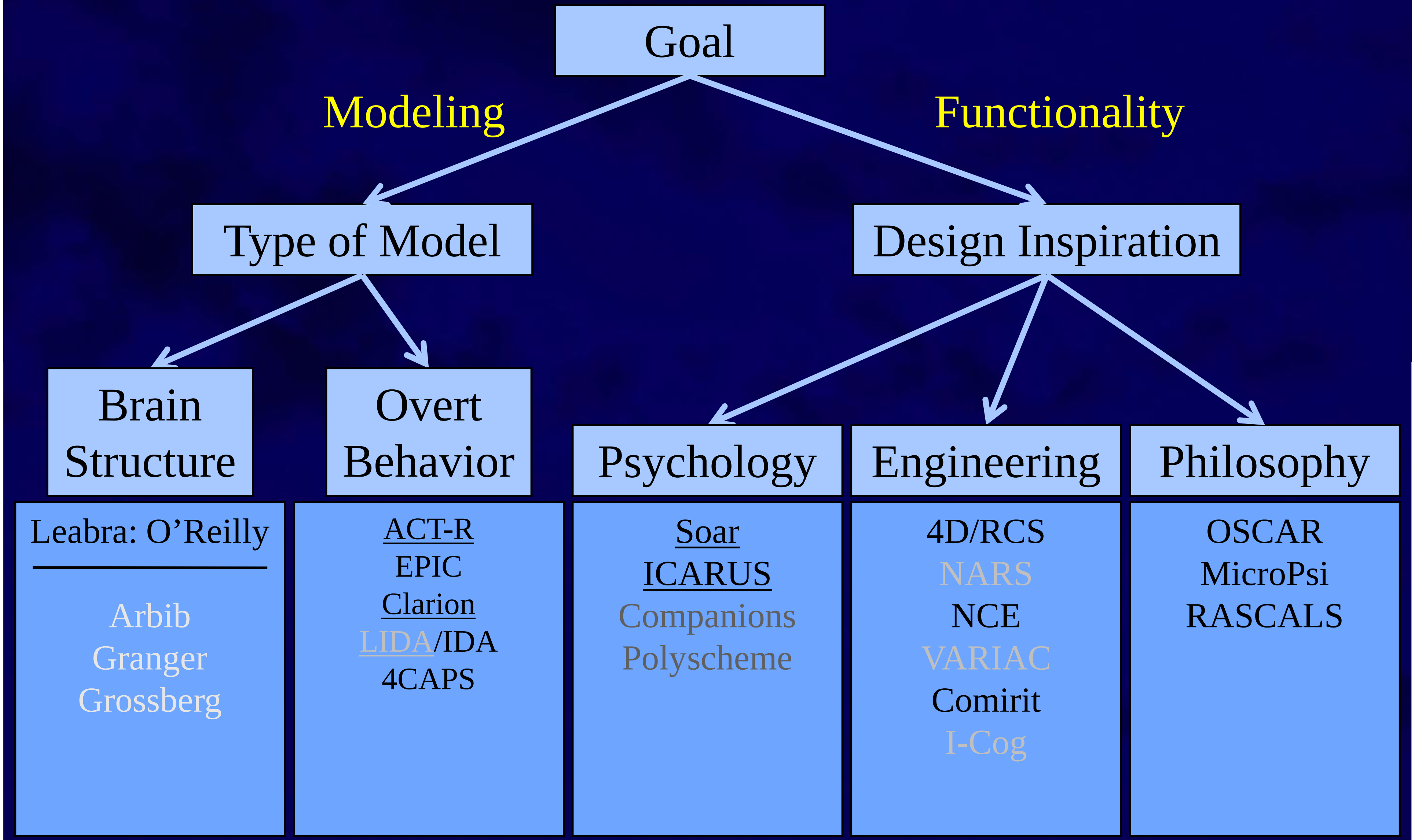
Classification of Active Architectures

- Goal
  - Modeling → Type of Model
    - Brain Structure
      - Leabra: O'Reilly
      - Arbib
      - Granger
      - Grossberg
    - Overt Behavior
      - ACT-R
      - EPIC
      - Clarion
      - LIDA/IDA
      - 4CAPS
  - Functionality → Design Inspiration
    - Psychology
      - Soar
      - ICARUS
      - Companions
      - Polyscheme
    - Engineering
      - 4D/RCS
      - NARS
      - NCE
      - VARIAC
      - Comirit
      - I-Cog
    - Philosophy
      - OSCAR
      - MicroPsi
      - RASCALS

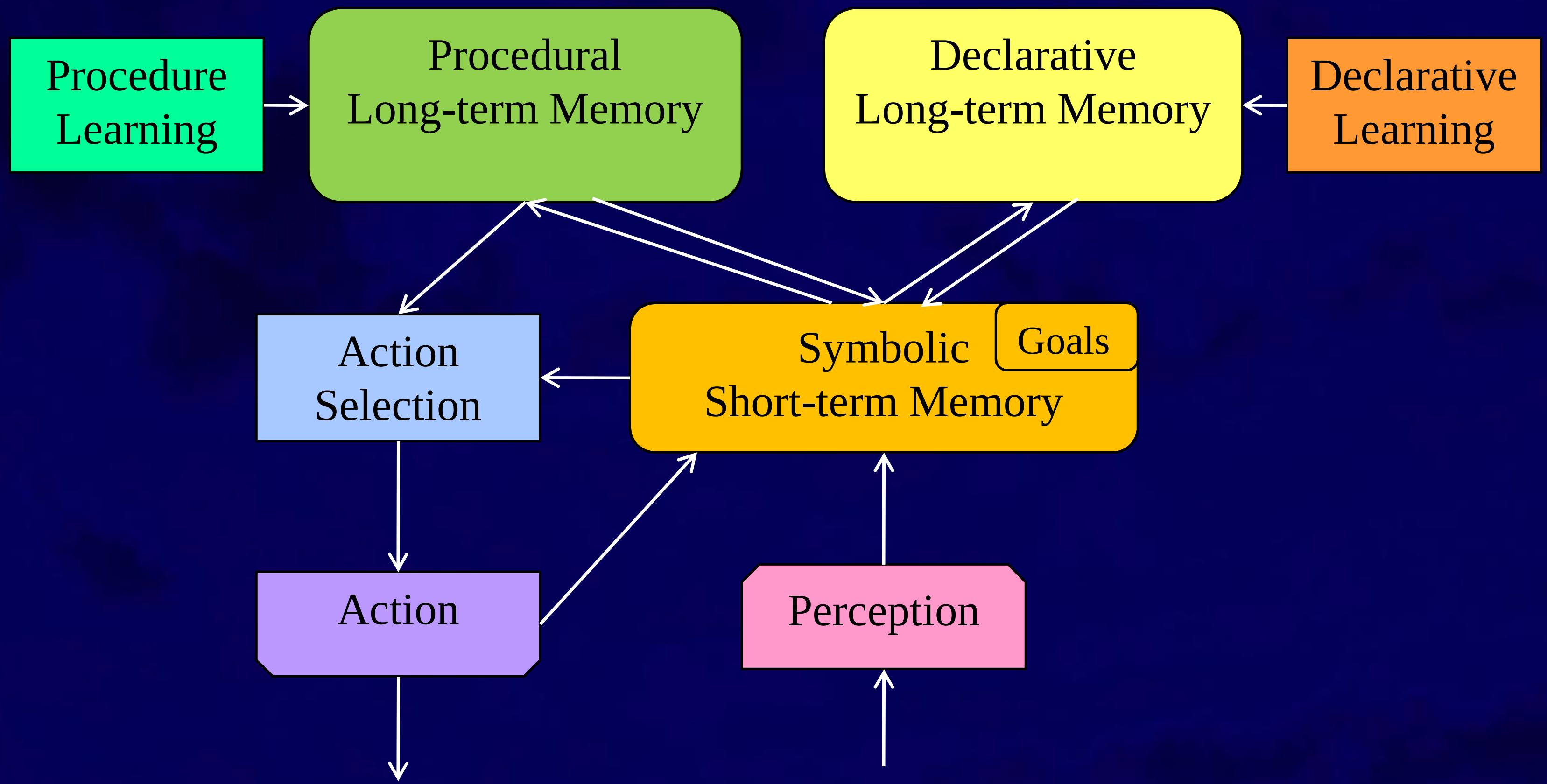
Classification of Active Architectures



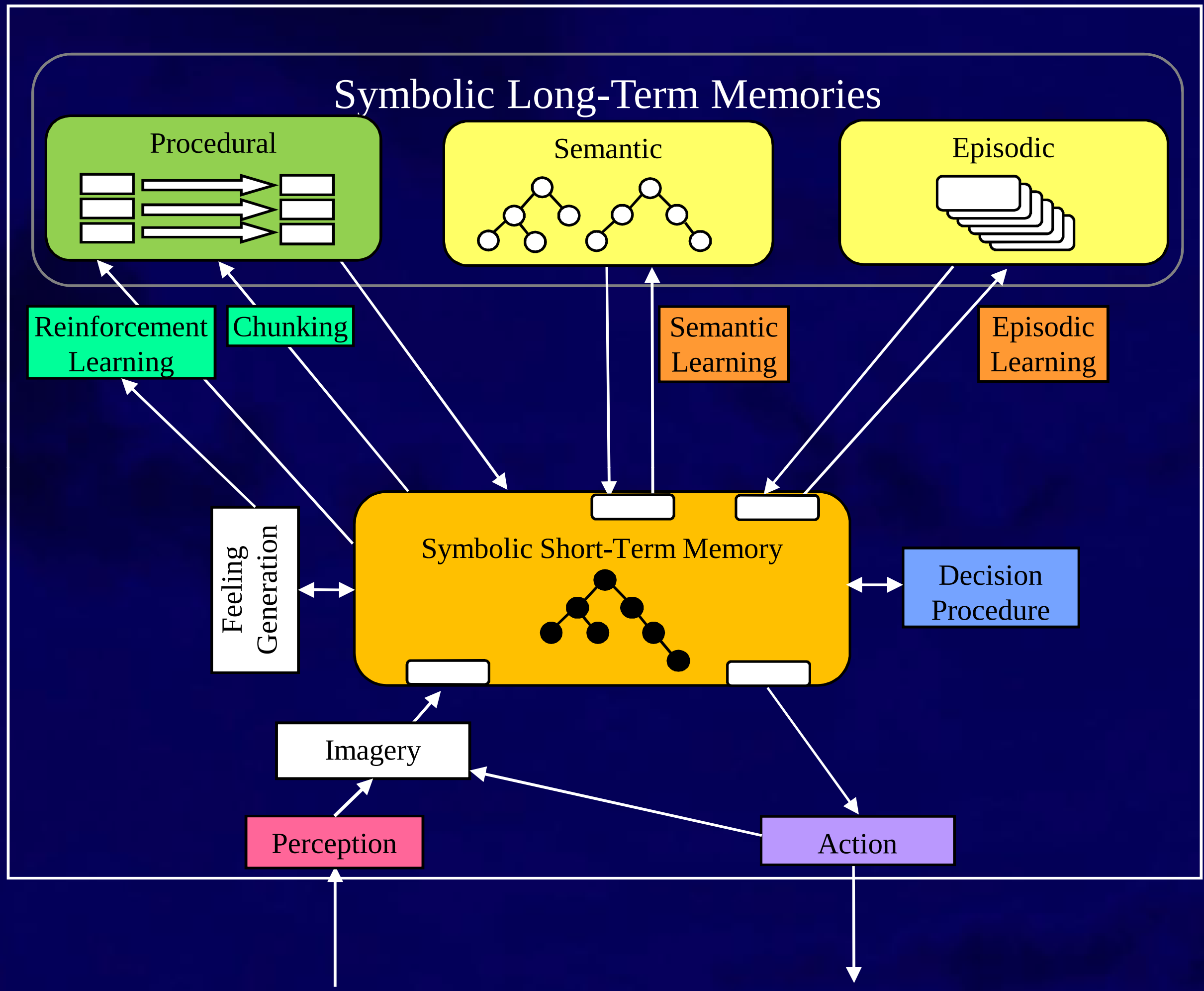
Classification of Active Architectures



Common Architectural Structure

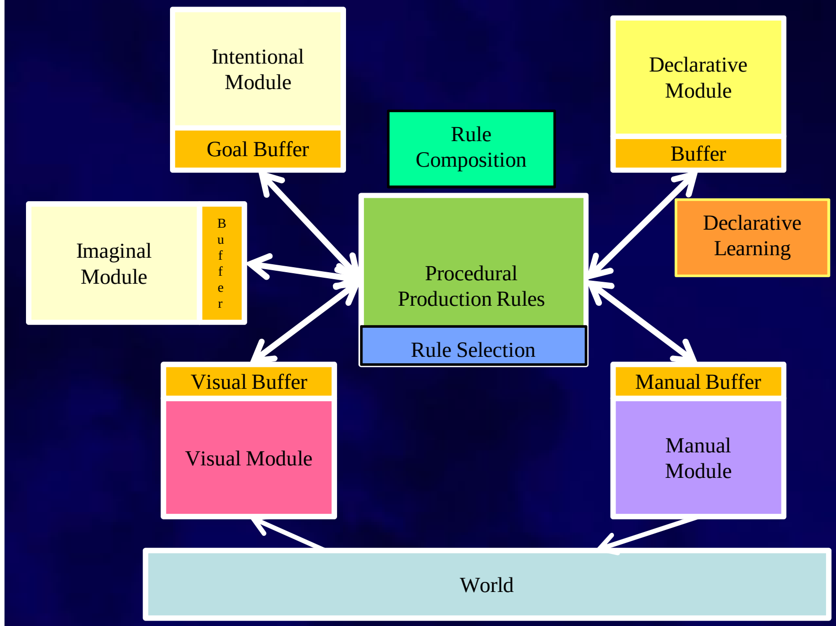


Soar

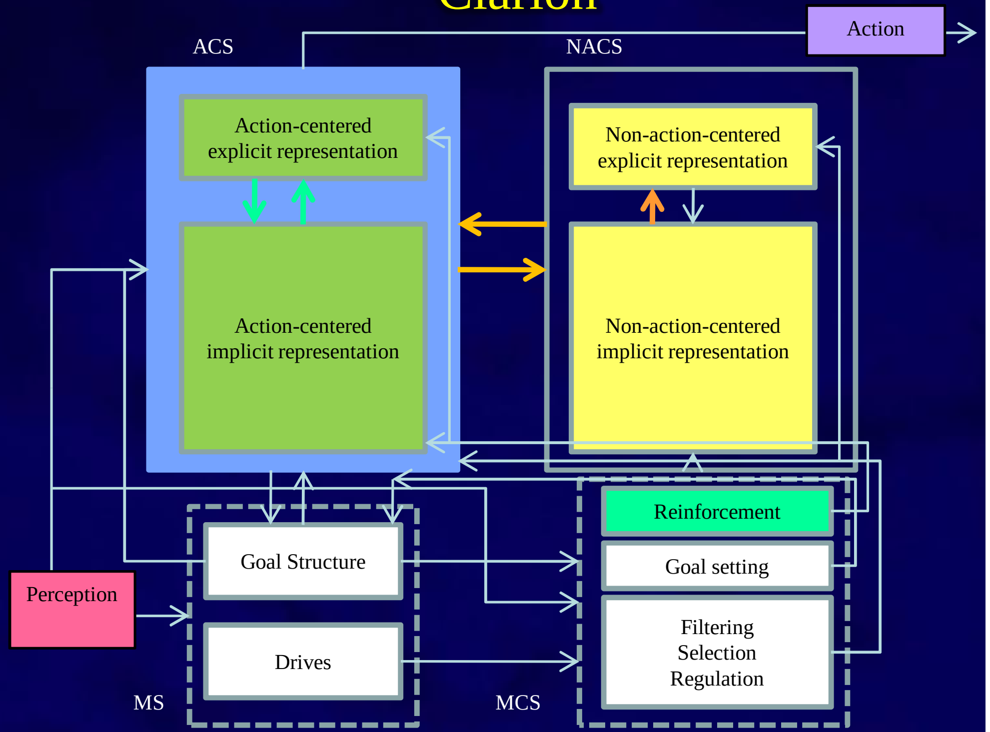


Symbolic Long-Term Memories
Procedural
Semantic
Episodic
Reinforcement Learning
Chunking
Semantic Learning
Episodic Learning
Feeling Generation
Symbolic Short-Term Memory
Decision Procedure
Imagery
Perception
Action

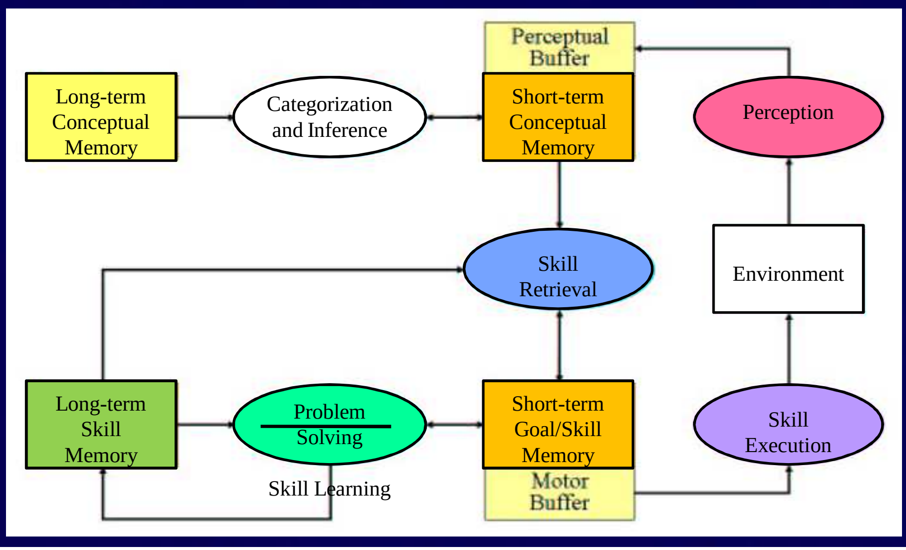
ACT-R 6



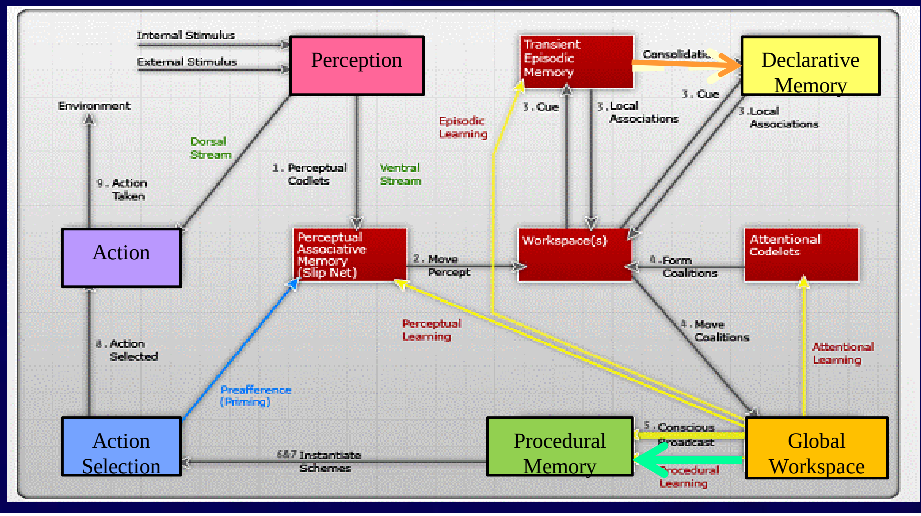
Clarion



ICARUS



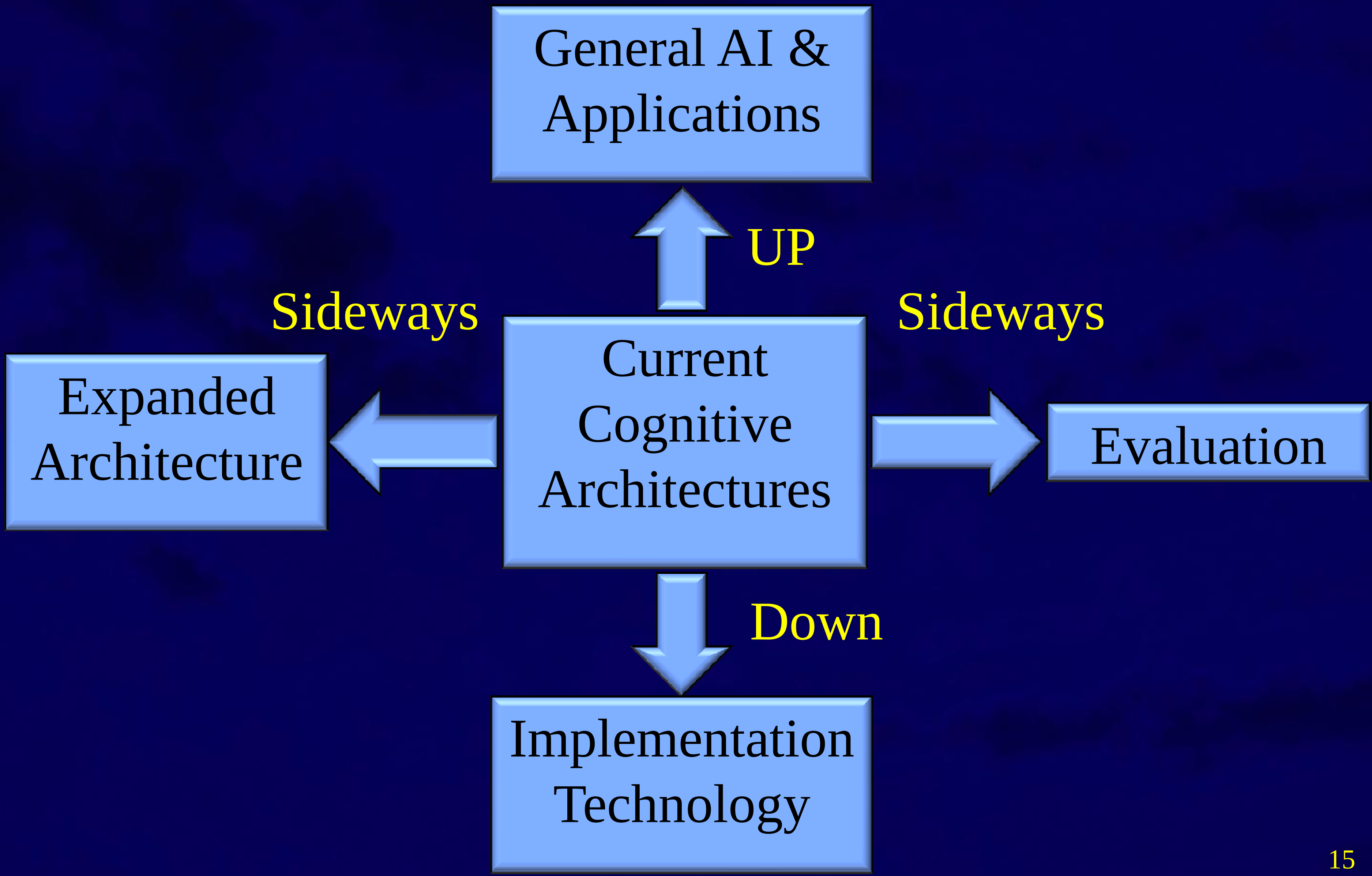
LIDA



Common Processing Across Many Architectures

- Complex behavior arises from sequence of simple decisions over internal and external actions controlled by knowledge
  - No monolithic plans
  - Significant internal parallelism, limited external parallelism
  - For cognitive modeling, ~50msec is basic cycle time of cognition
- <u>Knowledge access is assumed to be bounded to maintain reactivity</u>
- Symbolic long- & short-term knowledge representation
  - Procedural & semantic (Clarion also has non-symbolic)
  - Relational representations (-Clarion)
- Non-symbolic representation for action selection
- Learning is incremental & on-line (-LIDA)

Future of Cognitive Architecture



Up
Toward General Intelligent Agents

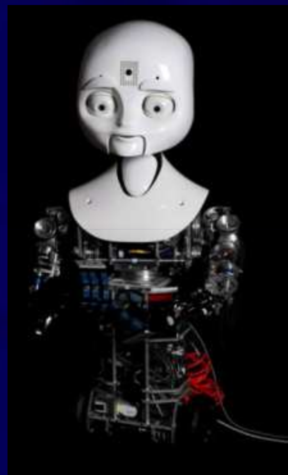
- Many more complex, knowledge-rich capabilities
  – Natural language
  – Planning
  – Spatial, temporal, meta-reasoning, …
  – Reflection to improve performance, develop strategies
- Interactions between those capabilities
  – Natural langue interaction to aid planning
  – Planning during natural language generation
- Social agents that exist for days and weeks perform many different tasks
- Learning is everywhere (*wild learning*)
  – From imitation, instruction, experience, reflection, …
  – *Transition from programming to training, learning by experience*

- Model behavior outside standard psychology experiments

<u>Up</u>

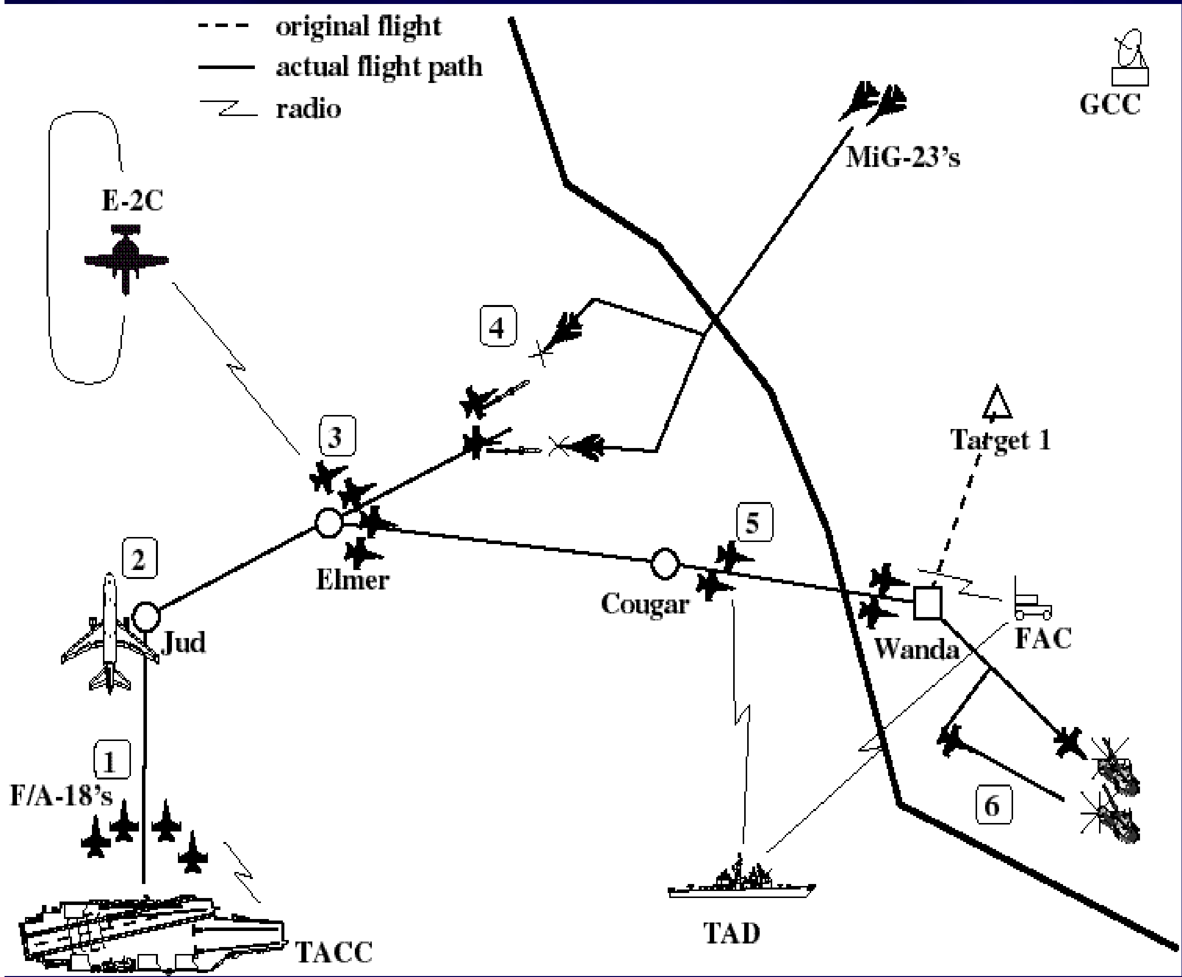
Applications

- ???
- Intelligent assistants
  - PAL: CALO/RADAR
  - Companions
- AI for computer games
- Intelligent robots
- AI for training & education



TacAir-Soar

[1997]



Controls simulated aircraft in real-time training exercises (>3000 entities)

Flies all U.S. air missions

Dynamically changes missions as appropriate

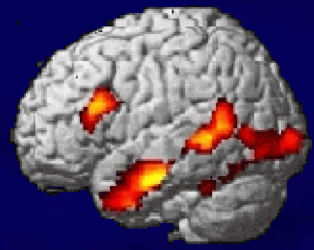
Communicates and coordinates with computer and human controlled planes

>8000 rules

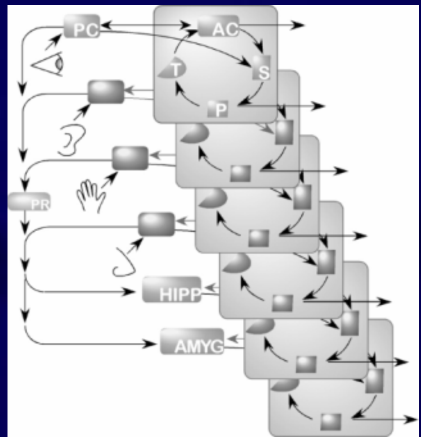
Down

Modeling: Map onto the Brain

- Map onto structure of human brain:
  - ACT-R & MRI



- Use neurologically inspired models of architecture components
  - ACT-R & Leabra
- Build up from models of neural circuits to cognitive processes
  - Arbib, Granger, Grossberg, …



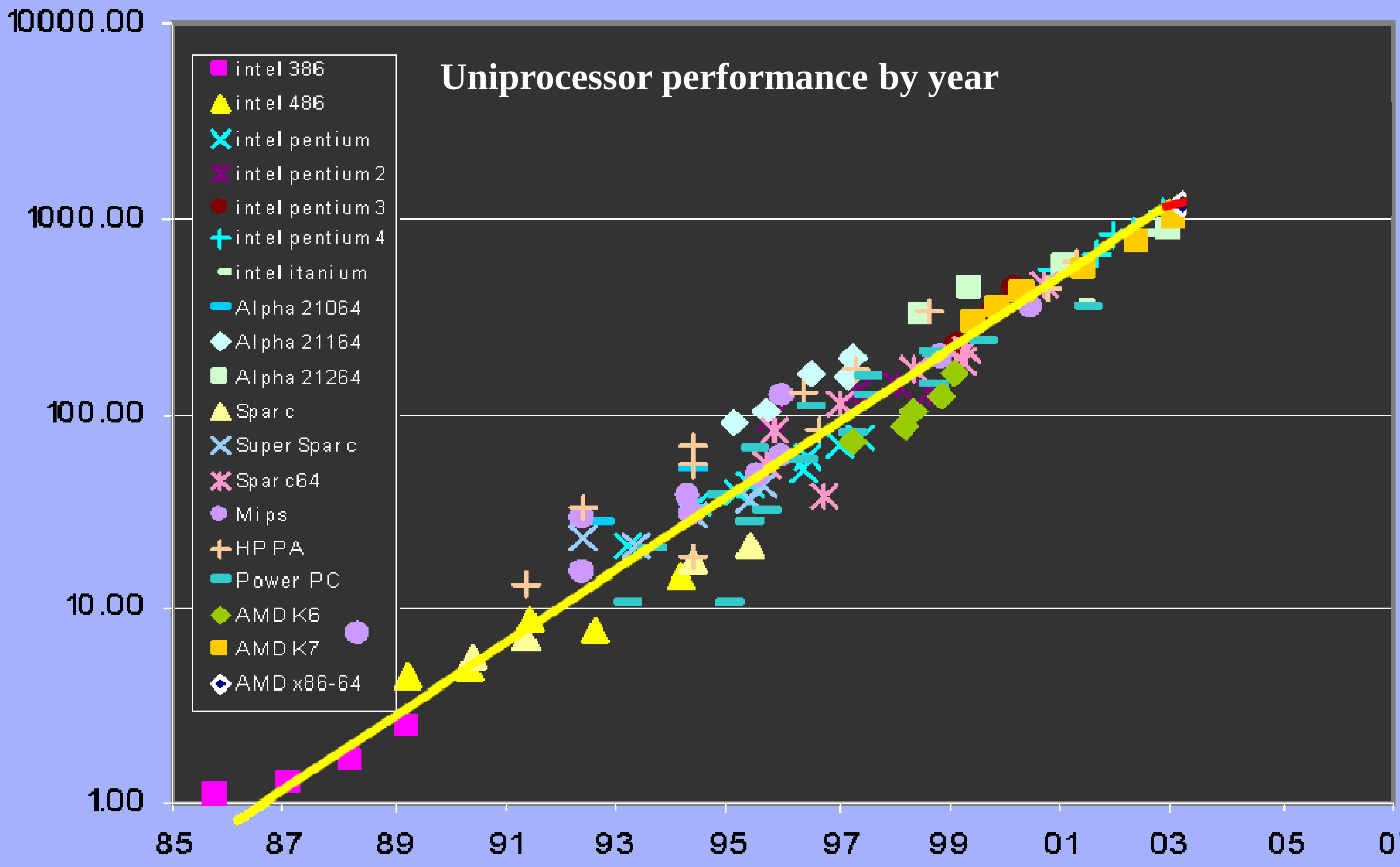
Down

Functionality: Scaling up

- Challenge
  - Real applications will require *huge* knowledge bases
    - 8,000 rules in TacAir-Soar
    - 3,00,000 facts in OpenCyc
  - Real learning leads to lots of knowledge
  - Architectures assume constant time memory retrieval

- Common response:
  - "Don't worry, Moore's Law will save us."

Moore’s Law



Power is the new limiter

[Courtesy Mark Horowitz, Stanford]

Down
Parallelism for Scaling

- Coarse-grain:
  - Multi-core & multi-processor clusters [Companions]
  - But Amdahl's law – still stuck with most costly process
- Fine-grain: New hardware architectures
  - FPGAs for memories
  - GPUs for imagery
  - ???
- Available technology can (should?) impact cognitive architecture design

Sideways

- Expanding set of architectural components & capabilities
- Evaluation

Sideways
Architectural (?) Capabilities

- Vision & motor control
- Categorization, classification, …
- Analogy
- Emotion
- Drives and Motivation [origin of goals]
- Non-symbolic representations, reasoning, learning
  - Mental Imagery
  - Probability

Sideways:
Evaluation

- No common tasks
- No common metrics
- No agreed upon evaluation methodology

- Need comparisons and tests for generality
  - Common tasks, metrics, evaluation methodology

Conclusion

- We are in a “Golden Age” of cognitive architecture
- Lots of exciting research ahead
  - Modeling: connections to the brain
  - Functionality: toward general human-level AI
- Many challenges:
  - Performance:
    - Ubiquitous learning
    - Scaling to large knowledge while maintaining reactivity
  - Applications
    - Do we really need human-level AI?
  - Evaluation
    - How do we get people to work on common problems and compare
  - Consolidation
    - Bring together best ideas
    - Connect to rest of AI