How are cognitive systems organised?

Sumeet Agarwal

[Slides based on Bermúdez Ch. 10]

Fundamental questions

1. In what format does a particular cognitive system carry information? [Representation]
2. How does that cognitive system transform information? [Processing]
3. How is the mind organised so that it can function as an information processor?

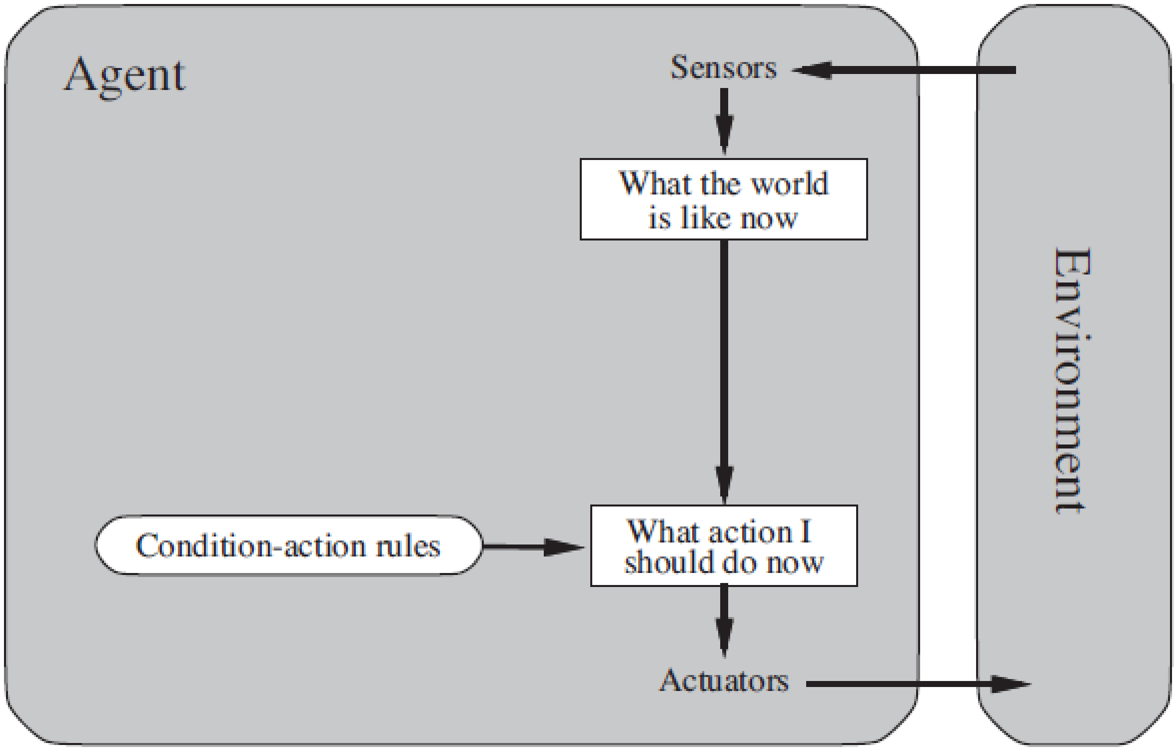
So far we've focused on the first two questions, in the context of different computational models (classical/symbolic, connectionist/neural, Bayesian) and different modalities (vision, language); now we'll think about the third

Architectures for intelligent agents

- Agent: System that perceives environment through *sensory systems* and acts upon it through *effector systems*
- *E.g.*, robots or software agents
- Types:
  - Simple reflex agent
  - Goal-based agent
  - Learning agent

Simple reflex agent

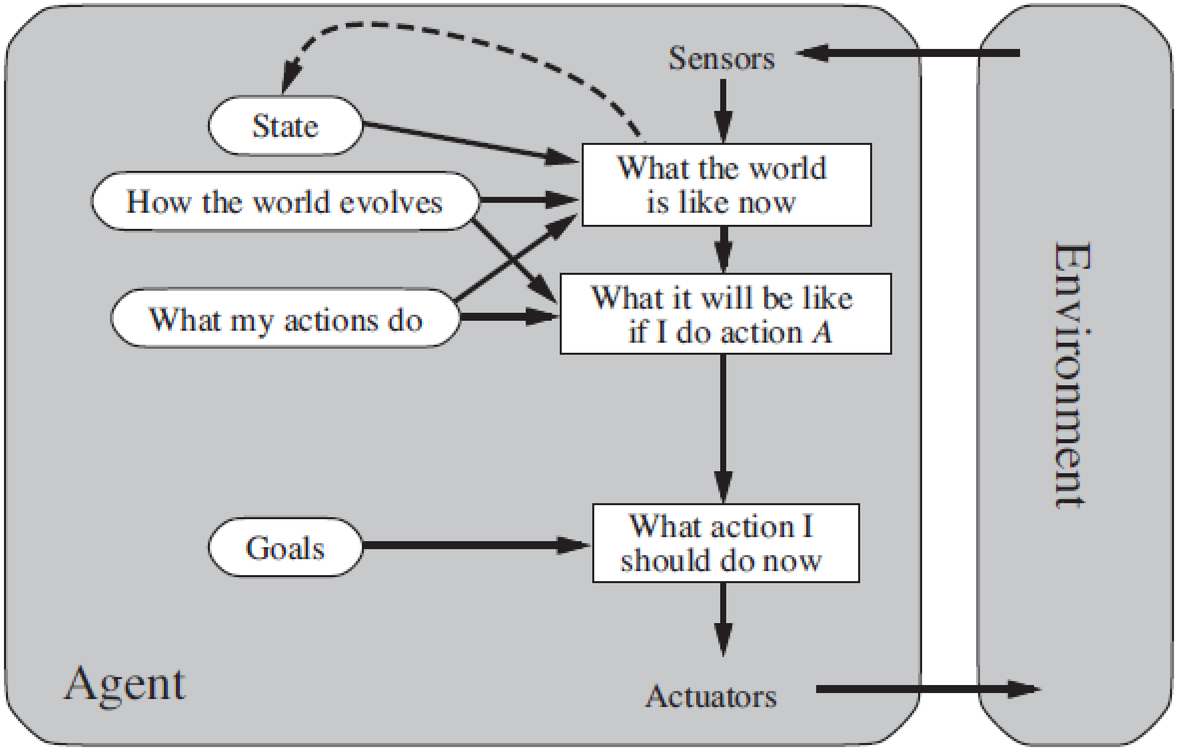
- Direct *condition-action rules*; not cognitive



[Russell and Norvig]

Goal-based agent

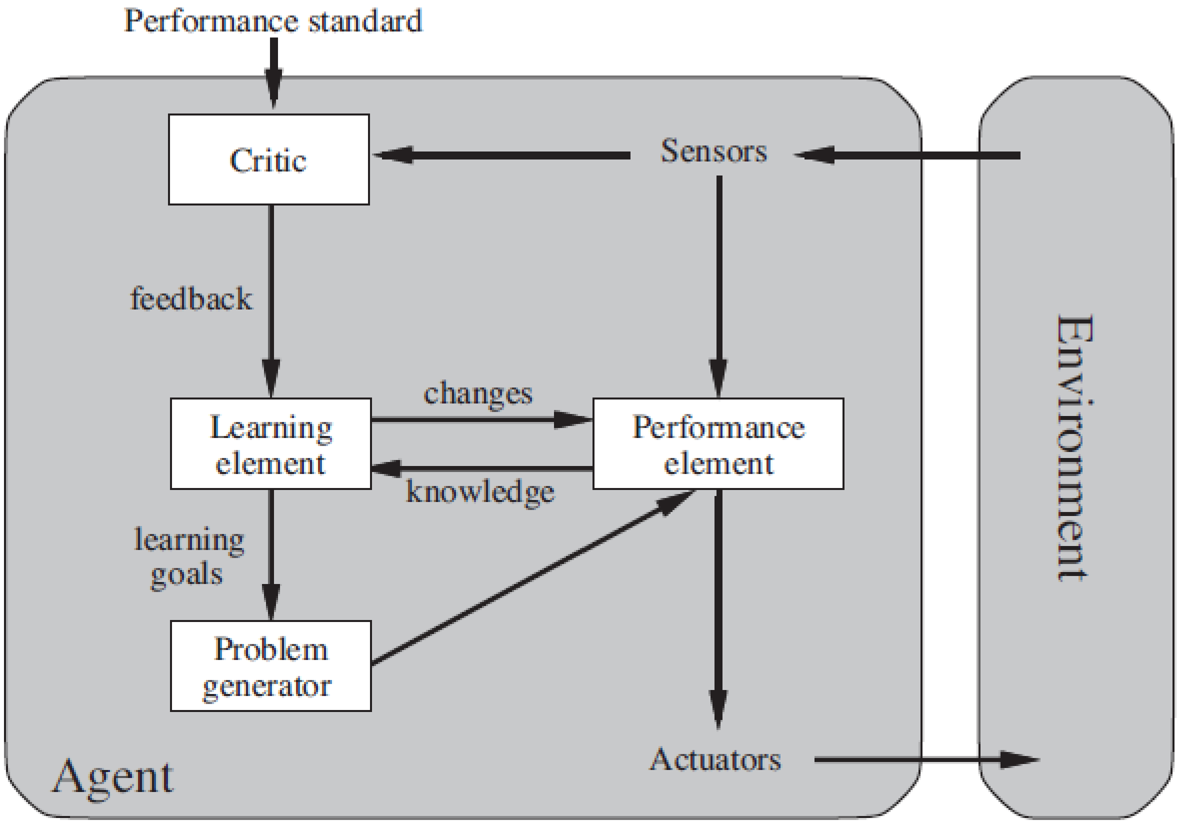
– Combine inputs and goals to decide on actions; primitively cognitive



[Russell and Norvig]

Learning agent

– Feedback as input; genuinely cognitive



[Russell and Norvig]

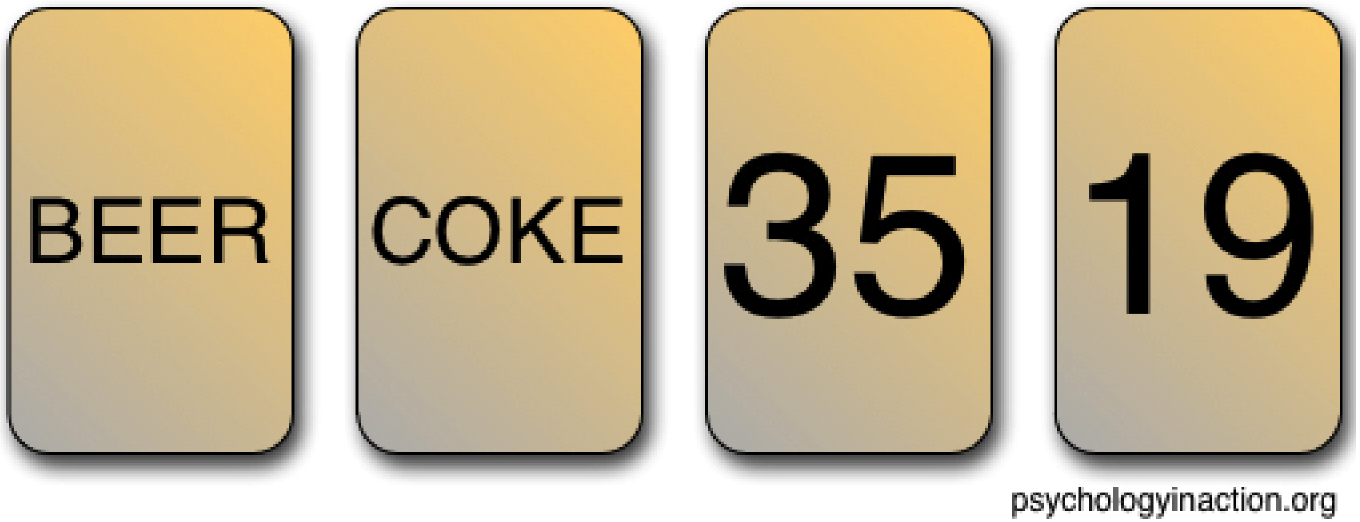
Specialised sub-systems

- Identifying/distinguishing sub-systems?
- Differences between sensory/effector sub-systems and systems in between?
- Differences in representation and processing across sub-systems?
- Autonomy and insulation of different sub-systems?

The Modularity of Mind [Jerry Fodor]

- Defence of organisation as autonomous cognitive sub-systems
- Not only *horizontal* cognitive faculties (domain-general: memory/attention), but also *vertical* (domain-specific: vision/language)
- Modular processes: Domain-specificity, informational encapsulation, mandatory application, speed; sometimes, fixed neural architecture, specific breakdown patterns
- Also central processing: Quinean, isotropic
- **Fodor's First Law of the Nonexistence of Cognitive Science**: “the more global a cognitive process is, the less anybody understands it”
- Massive modularity hypothesis [Leda Cosmides, John Tooby]

Each card has a letter on one side and a number on the other.
Proposed Rule: **If a card has a vowel on one side, then it has an even number on the other side**.
Which card(s) do you need to turn over in order to determine if the rule is true or false?



These four cards represent patrons in a bar, and each card has their drink on one side and their age in years on the other.

Proposed Rule: **If a patron is drinking a beer, then they must be 21 years or older.**

Which card(s) do you need to turn over in order to determine if the rule is being followed?

Solutions

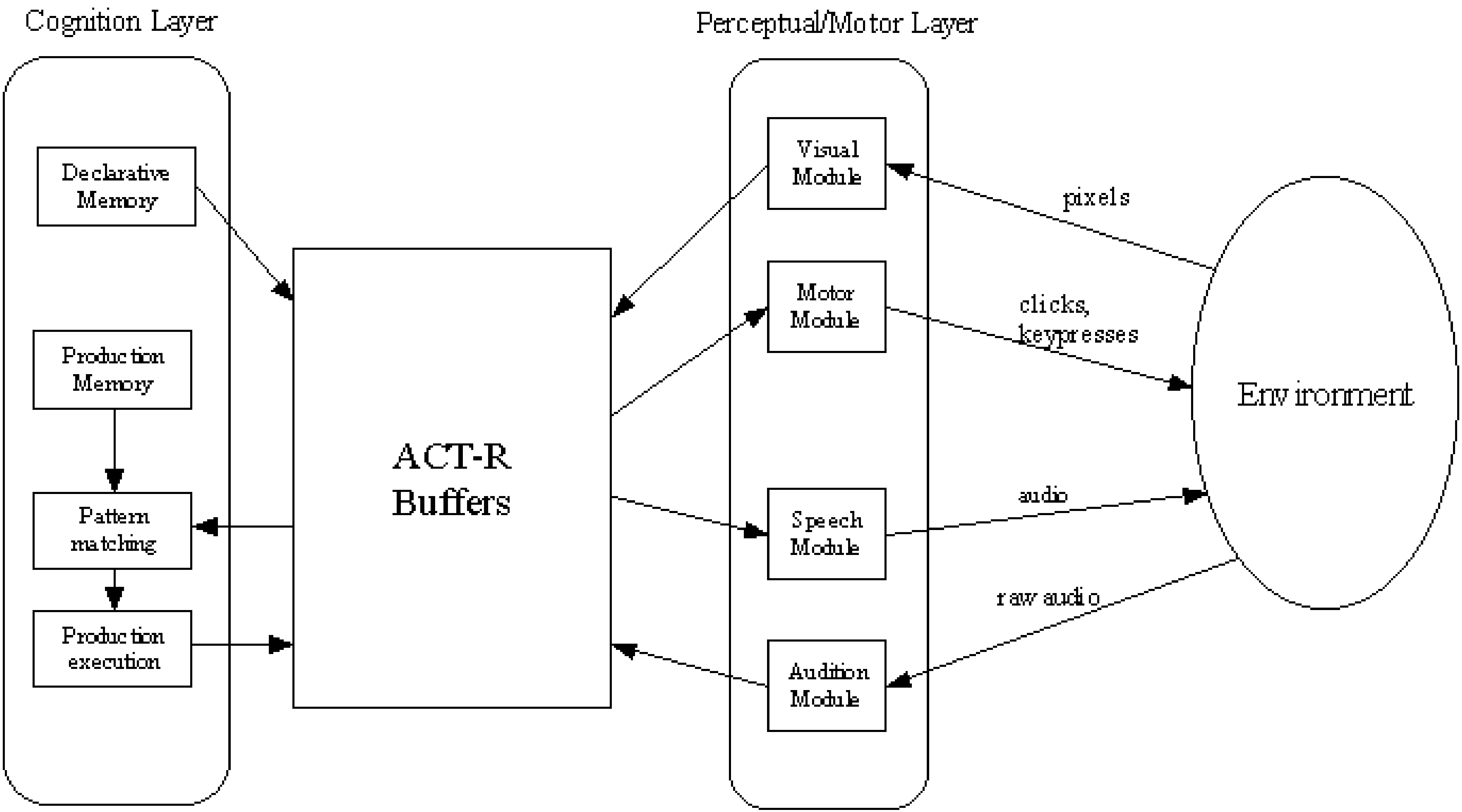
- Puzzle 1: **A 7**
  - Fully correct: 3/16 (19%)
  - **A 4** 4/16; **A D 4 7** 3/16
  - Individual card-wise: 19/37 (51%)
- Puzzle 2: **Beer 19**
  - Fully correct: 8/17 (47%)
  - **Beer 35** 3/17; **35** 3/17
  - Individual card-wise: 22/28 (79%)

Massive modularity hypothesis

- Based on evolutionary psychology and the idea of Darwinian modules, e.g., cheater detection
- Each module evolved to solve a specific adaptive task which enhanced evolutionary fitness
- Mind is only such modules, because no domain-general architecture could have solved adaptive problems
  - Argument from error: Fitness criteria are domain-specific
  - Argument from statistics and learning: Patterns in the world are domain-specific
- Darwinian modules not same as Fodorean; can also be achieved by domain-specific *knowledge*?

Hybrid architectures: *ACT-R/PM*

- Adaptive Control of Thought – Rational/Perceptual-Motor [John R. Anderson]
- Symbolic and neural computational models not mutually exclusive (symbolic better for highly structured and sharply defined problems, neural better for perceptual and pattern recognition); ACT-R/PM seeks to hybridise
- Cognitive architecture: Similar to a programming language. Toolkit for constructing cognitive models.
- Another cognitive architecture: *Soar* (originally, State Operator And Result). Only symbolic, rule-based computation, no neural networks



Cognition Layer
Perceptual/Motor Layer
Declarative Memory
Production Memory
Pattern matching
Production execution
ACT-R Buffers
Visual Module
Motor Module
Speech Module
Audition Module
pixels
clicks, keypresses
audio
raw audio
Environment

ACT-R/PM

- Consists of cognition layer and perceptual-motor layer, interacting through buffers
- Two types of knowledge: declarative (*knowledge-that*) and procedural (*knowledge-how*)
- Declarative knowledge stored as *chunks*, procedural knowledge as *production rules*; both symbolic
- Seems Fodorean, but no central processor; also not massively modular (cognition layer is domain-general)
- What makes it hybrid is subsymbolic *selection* of production rules in the pattern-matching module, based on *utility*
- Also subsymbolic activation levels of declarative memory

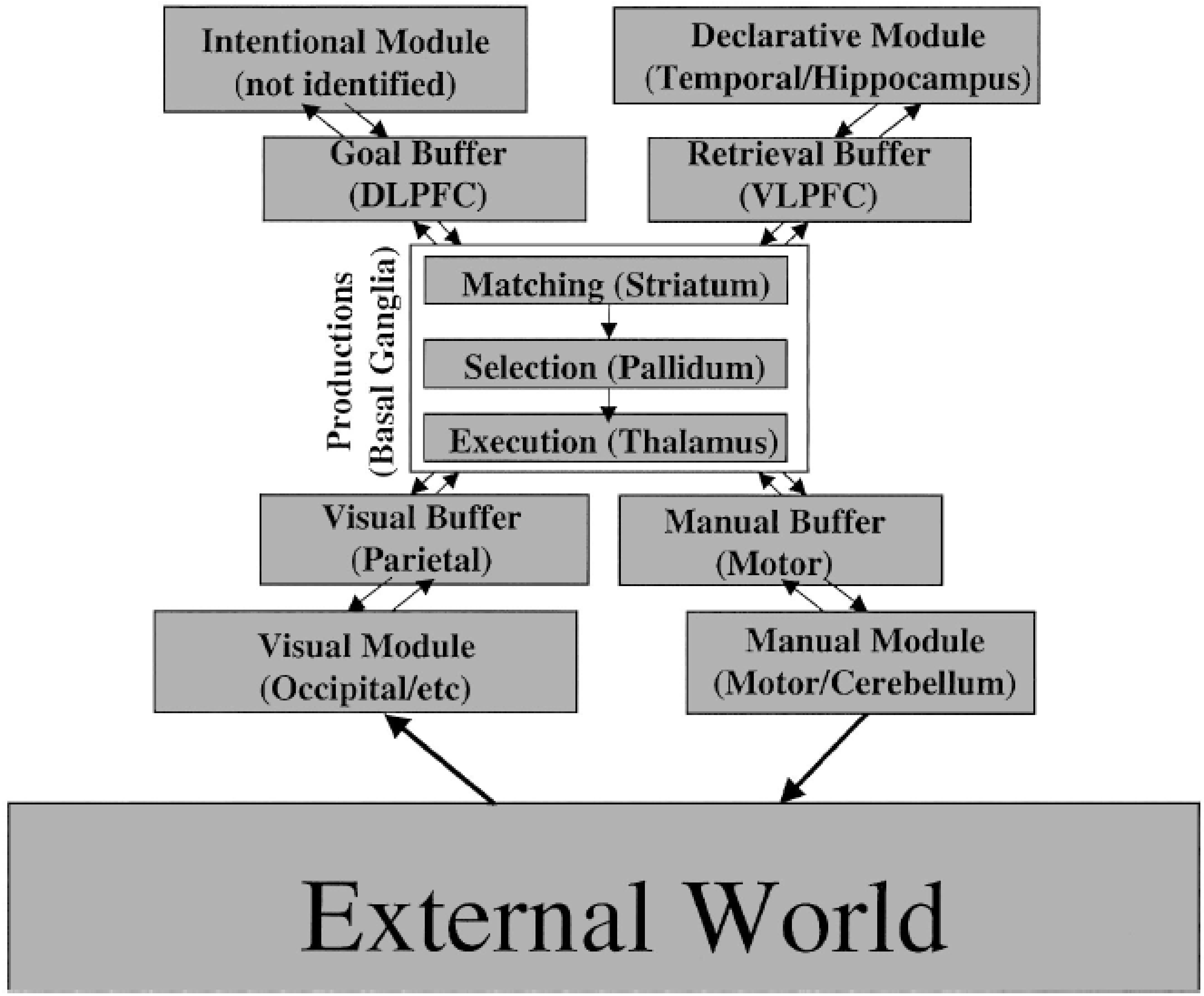
Memory type	Symbolic performance mechanisms	Subsymbolic performance mechanisms	Symbolic learning mechanisms	Subsymbolic learning mechanisms
Declarative chunks	Knowledge (usually facts) that can be directly verbalised	Relative activation of declarative chunks affects retrieval	Adding new declarative chunks to the set	Changing activation of declarative chunks and changing strength of links between chunks
Production rules	Knowledge for taking particular actions in particular situations	Relative utility of production rules affects choice	Adding new production rules to the set	Changing utility of production rules

ACT-R/PM: Summary

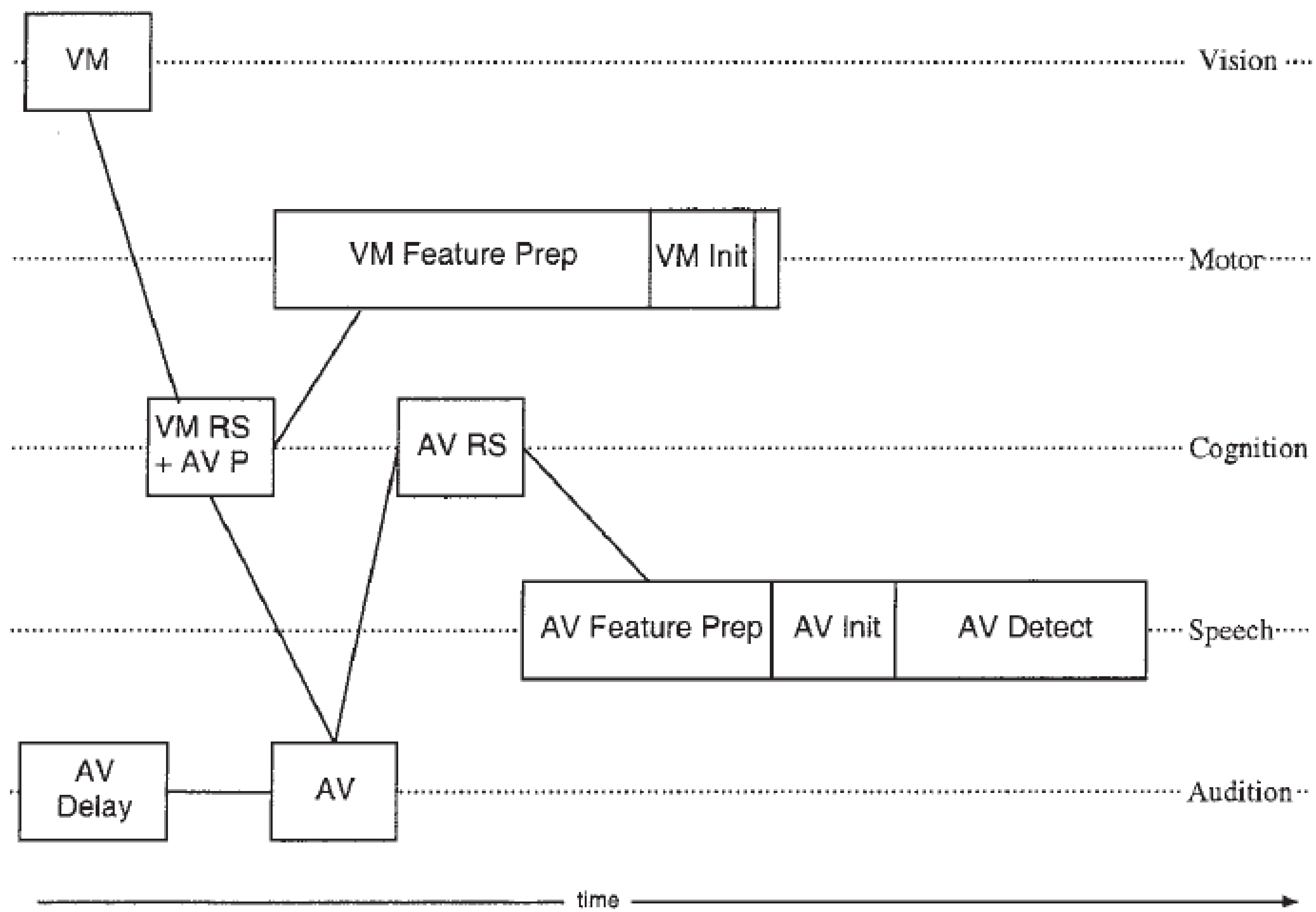
- Both symbolic and subsymbolic information processing; the latter involves pattern matching and updation of activation/utility strengths via learning, so could map onto models like neural networks
- Close connection between organisation of mind and nature of information processing
- Different parts/levels of a cognitive architecture can exploit different information processing models; so models not mutually exclusive but perhaps complementary
- http://act-r.psy.cmu.edu/

“A single system (mind) produces all aspects of behavior. It is one mind that minds them all. Even if the mind has parts, modules, components, or whatever, they all mesh together to produce behavior. Any bit of behavior has causal tendrils that extend back through large parts of the total cognitive system before grounding in the environmental situation of some earlier times. If a theory covers only one part or component, it flirts with trouble from the start. It goes without saying that there are dissociations, independencies, impenetrabilities, and modularities. These all help to break the web of each bit of behavior being shaped by an unlimited set of antecedents. So they are important to understand and help to make that theory simple enough to use. But they don’t remove the necessity of a theory that provides the total picture and explains the role of the parts and why they exist.”

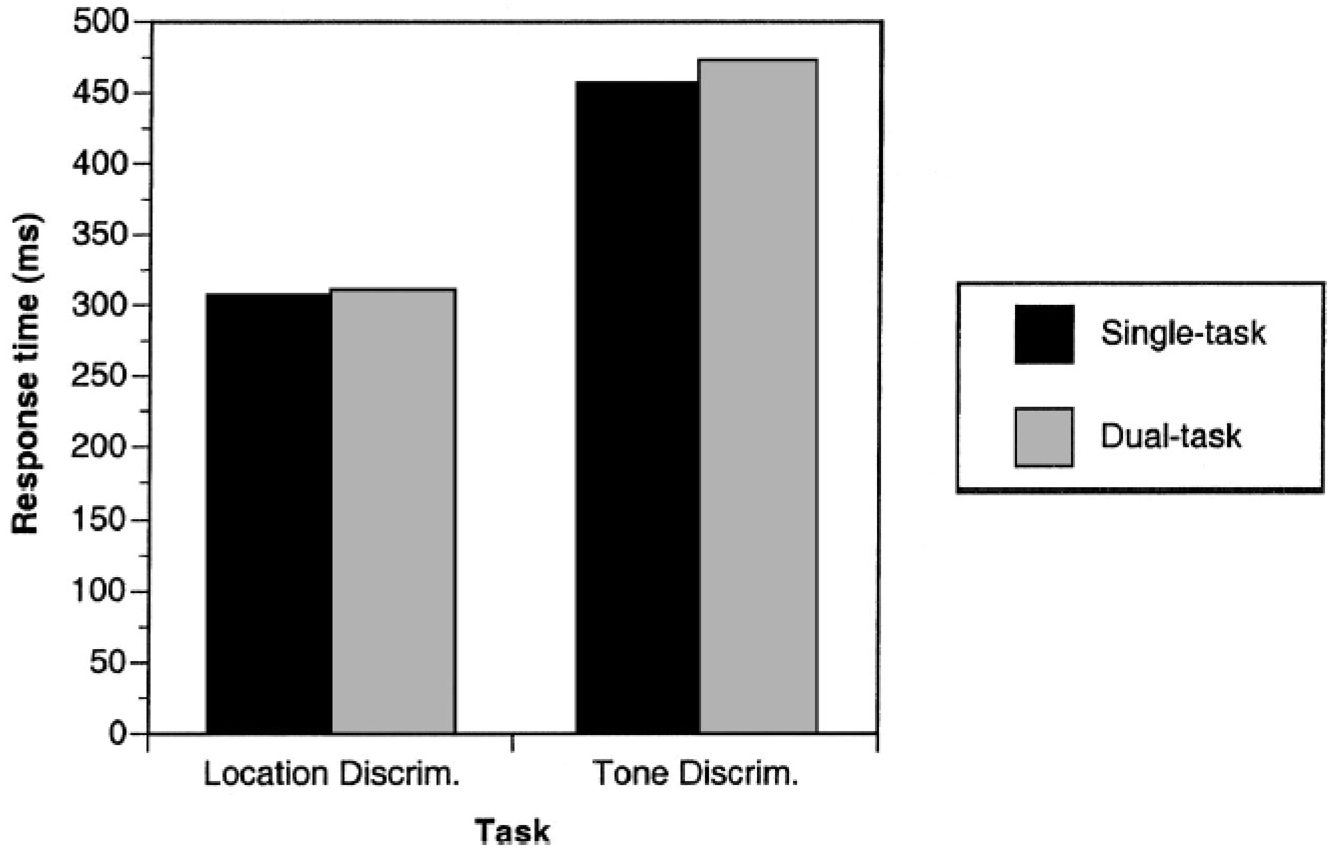
—Allen Newell

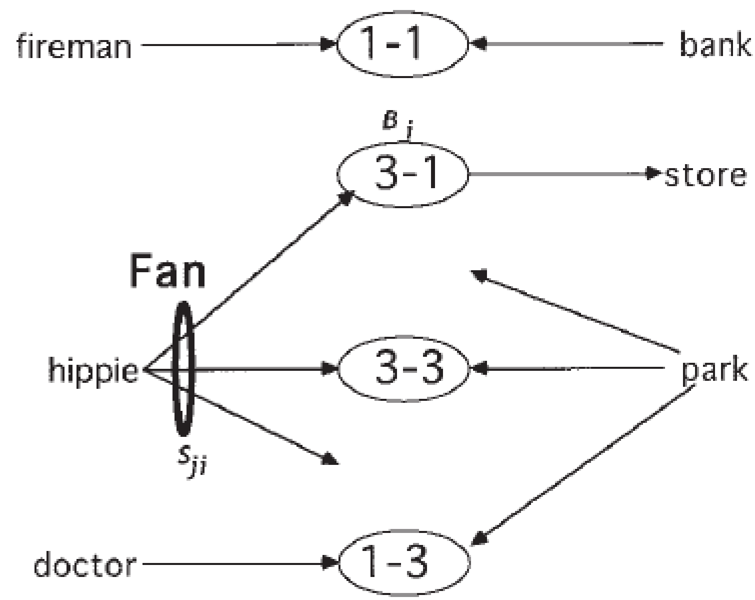


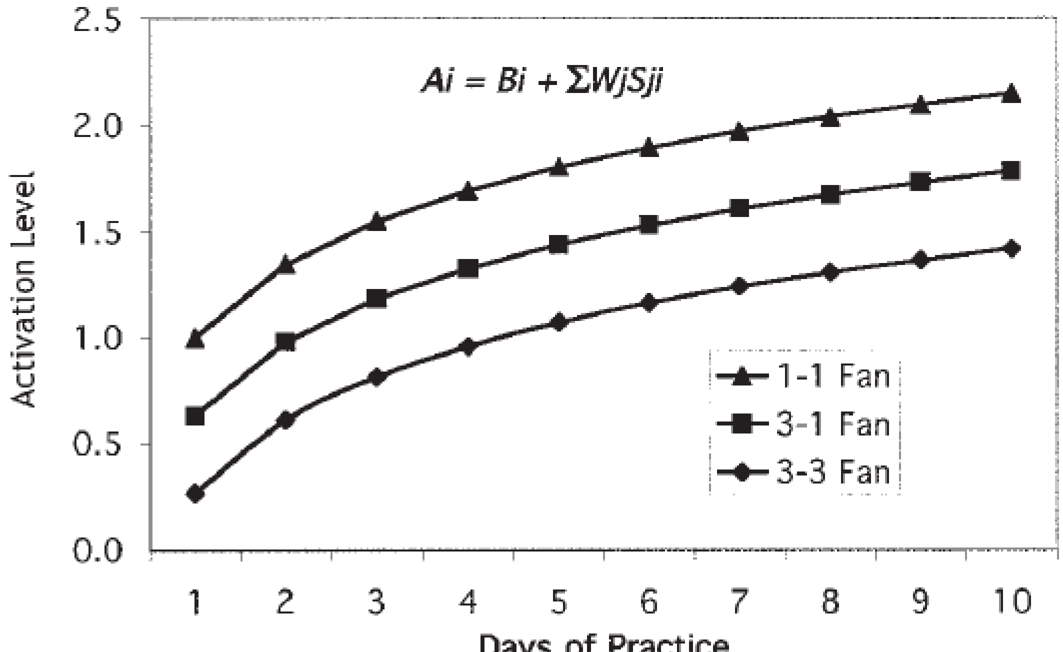
Intentional Module
(not identified)
Declarative Module
(Temporal/Hippocampus)
Goal Buffer
(DLPFC)
Retrieval Buffer
(VLPFC)
Productions
(Basal Ganglia)
Matching (Striatum)
Selection (Pallidum)
Execution (Thalamus)
Visual Buffer
(Parietal)
Manual Buffer
(Motor)
Visual Module
(Occipital/etc)
Manual Module
(Motor/Cerebellum)
External World

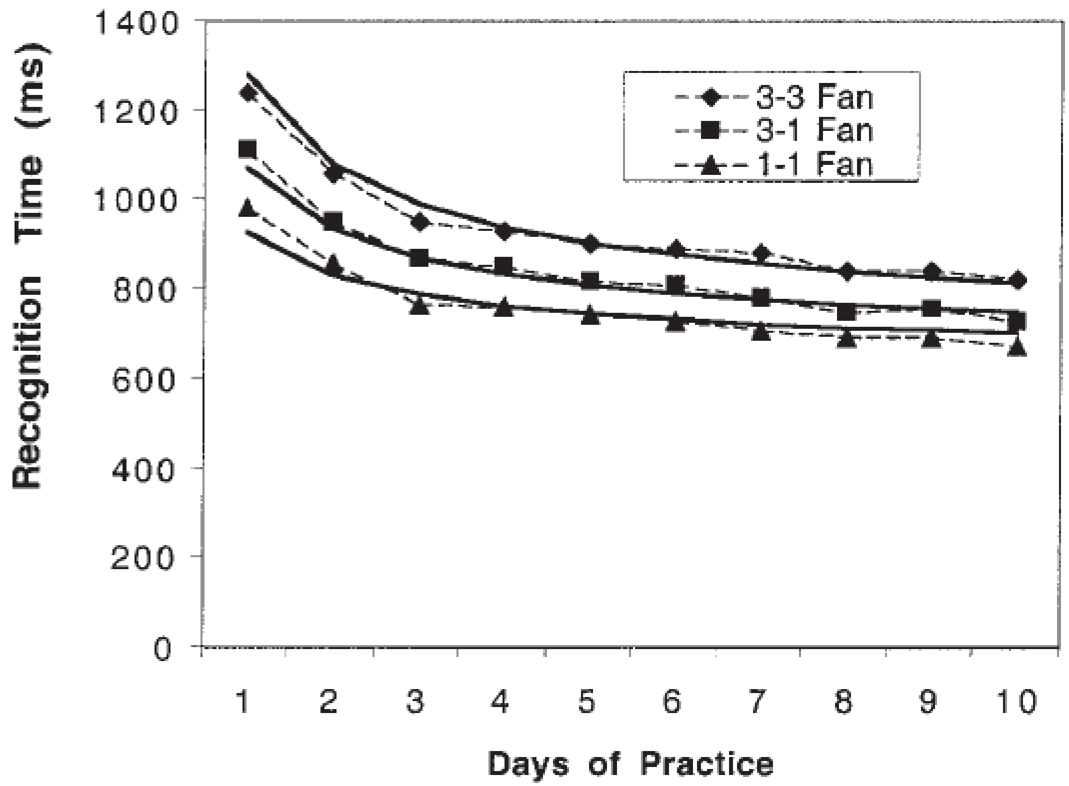


VM
Vision
VM Feature Prep
VM Init
Motor
VM RS
+ AV P
AV RS
Cognition
AV Feature Prep
AV Init
AV Detect
Speech
AV
Delay
AV
Audition
time

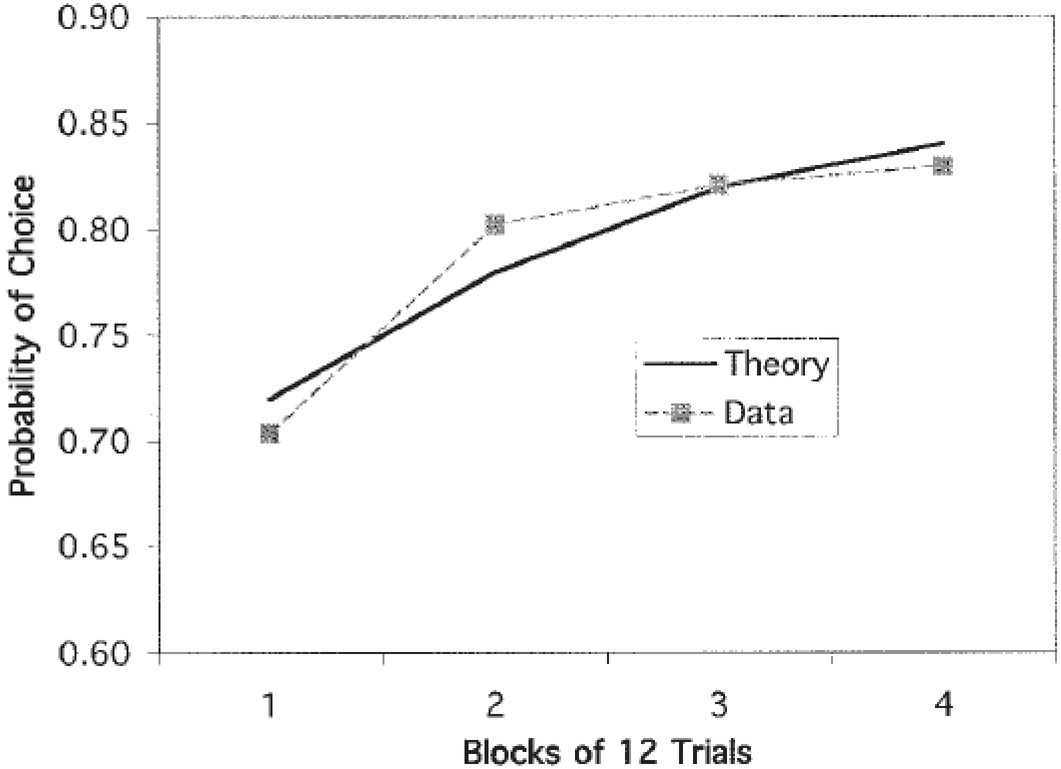

500
450
400
350
300
250
200
150
100
50
0
Response time (ms)
Location Discrim.
Tone Discrim.
Task
Single-task
Dual-task


fireman
1-1
bank
Bi
3-1
store
Fan
hippie
3-3
park
Sji
doctor
1-3


Ai = Bi + ΣWjSji
Activation Level
Days of Practice
1-1 Fan
3-1 Fan
3-3 Fan
2.5
2.0
1.5
1.0
0.5
0.0
1
2
3
4
5
6
7
8
9
10



3-3 Fan
3-1 Fan
1-1 Fan
1400
1200
1000
800
600
400
200
0
Recognition Time (ms)
1
2
3
4
5
6
7
8
9
10
Days of Practice



0.90
0.85
0.80
0.75
0.70
0.65
0.60
Probability of Choice
Theory
Data
1
2
3
4
Blocks of 12 Trials