Design & Implementation of Cognition-enabled Robot Agents

Module 10: Cognitive Architectures
Lecture 2: Example Cognitive Architectures

Institute for Artificial Intelligence
Universität Bremen

Winter Term 2020/21

Lecture Contents

1. Surveys
2. Soar
3. ACT-R
4. Clarion
5. ISAC
6. Lecture summary
7. Recommended reading & videos

Example Cognitive Architectures

Surveys:

Biologically Inspired Cognitive Architectures Society, Comparative Repository of Cognitive Architectures, http://bicasociety.org/cogarch/architectures.htm (25 cognitive architectures)

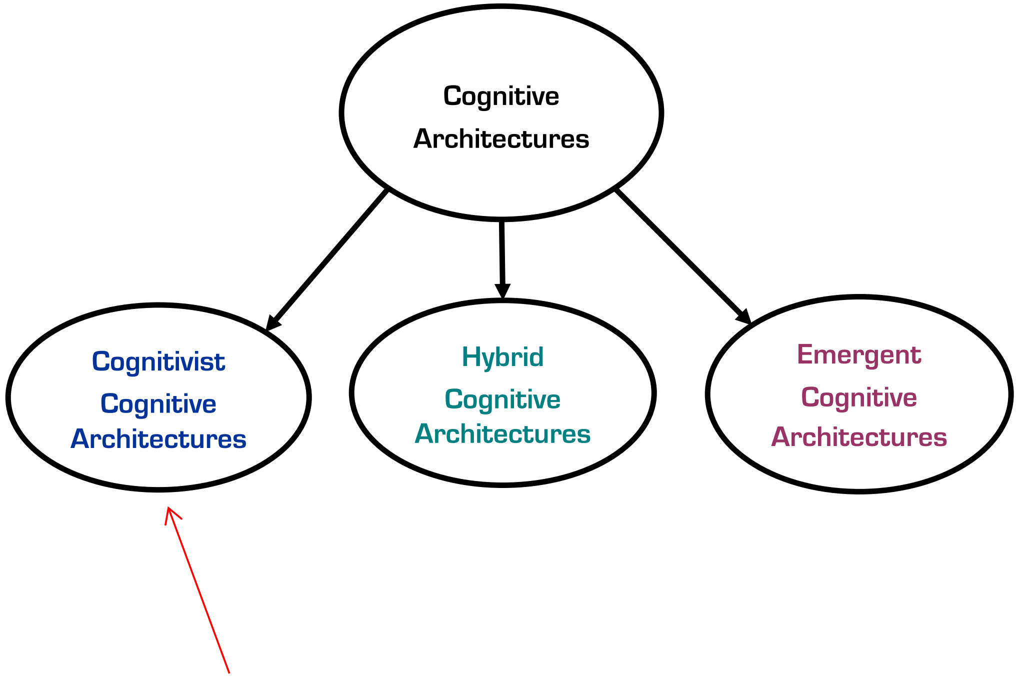
A Survey of Cognitive and Agent Architectures, University of Michigan, http://ai.eecs.umich.edu/cogarch0/ (12 cognitive architectures)

W. Duch, R. J. Oentaryo, and M. Pasquier. "Cognitive Architectures: Where do we go from here?", Proc. Conf. Artificial General Intelligence, 122-136, 2008. (17 cognitive architectures)

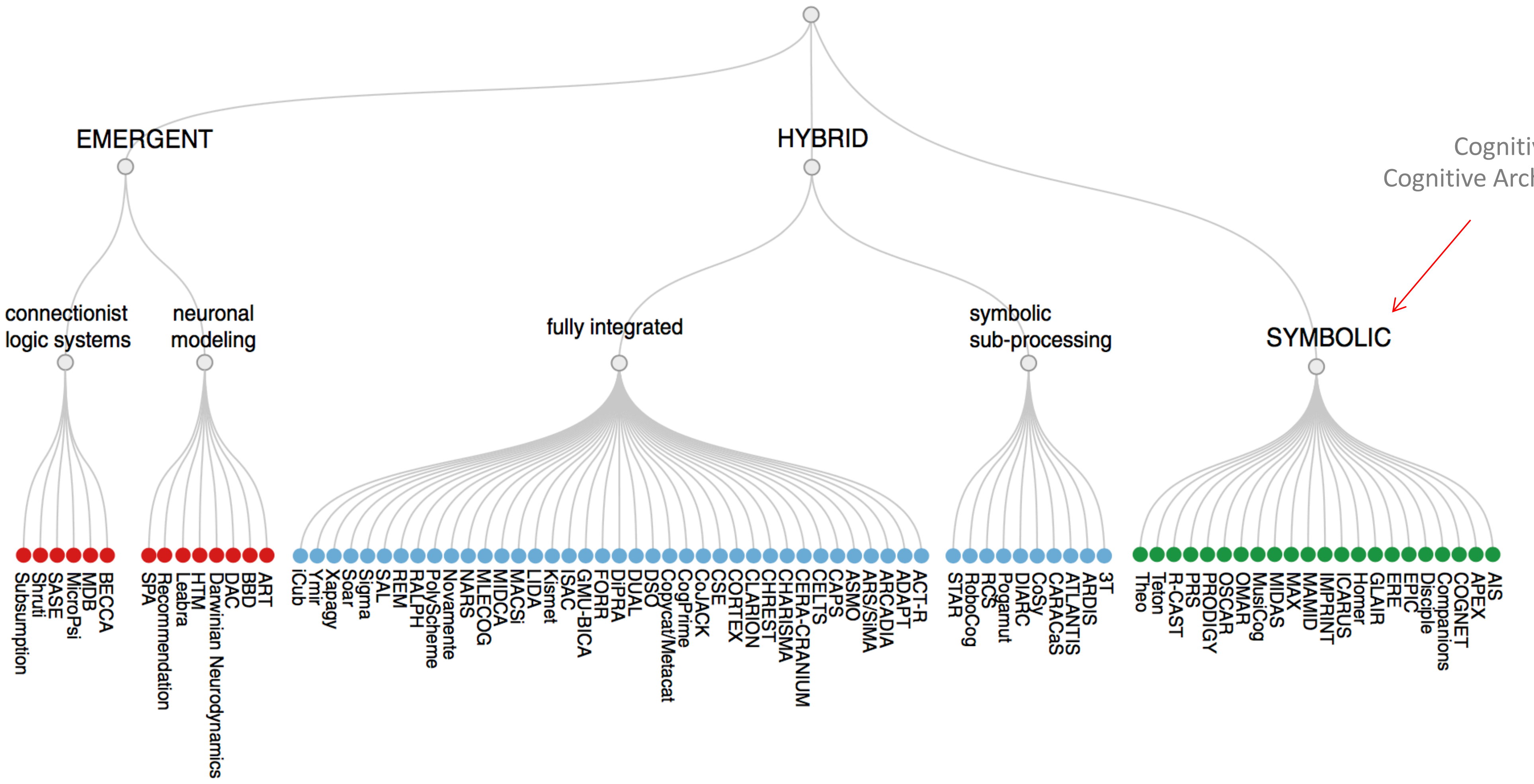
D. Vernon, G. Metta, and G. Sandini, "A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents", IEEE Transactions on Evolutionary Computation, Vol. 11, No. 2, pp. 151-180, 2007. (14 cognitive architectures)

D. Vernon, C. von Hofsten, and L. Fadiga. "A Roadmap for Cognitive Development in Humanoid Robots", Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010. Chapter 5 and Appendix I (20 cognitive architectures)

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020. (84 cognitive architectures)

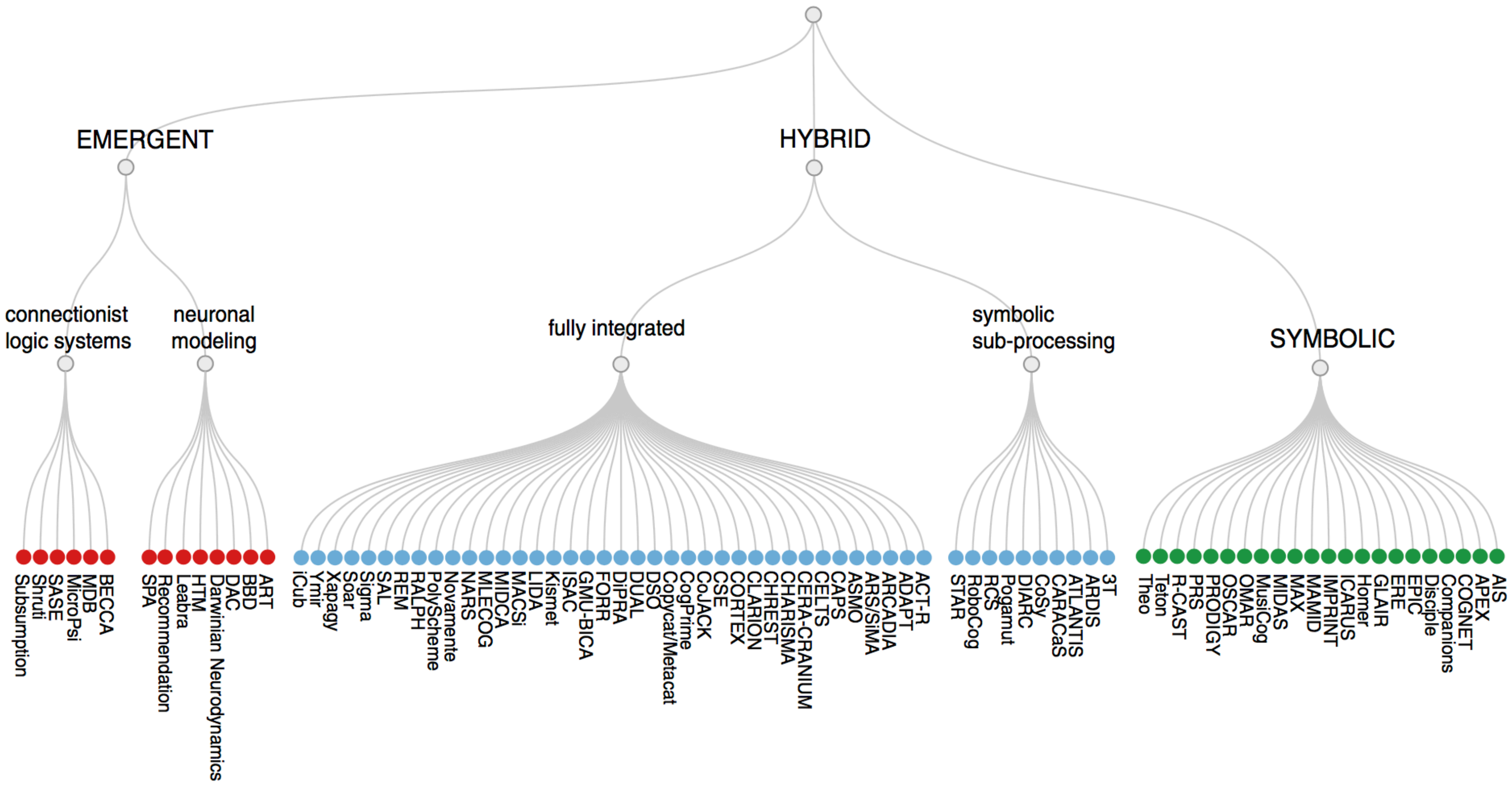


Kotseruba and Tsotsos refer to these as Symbolic Architectures

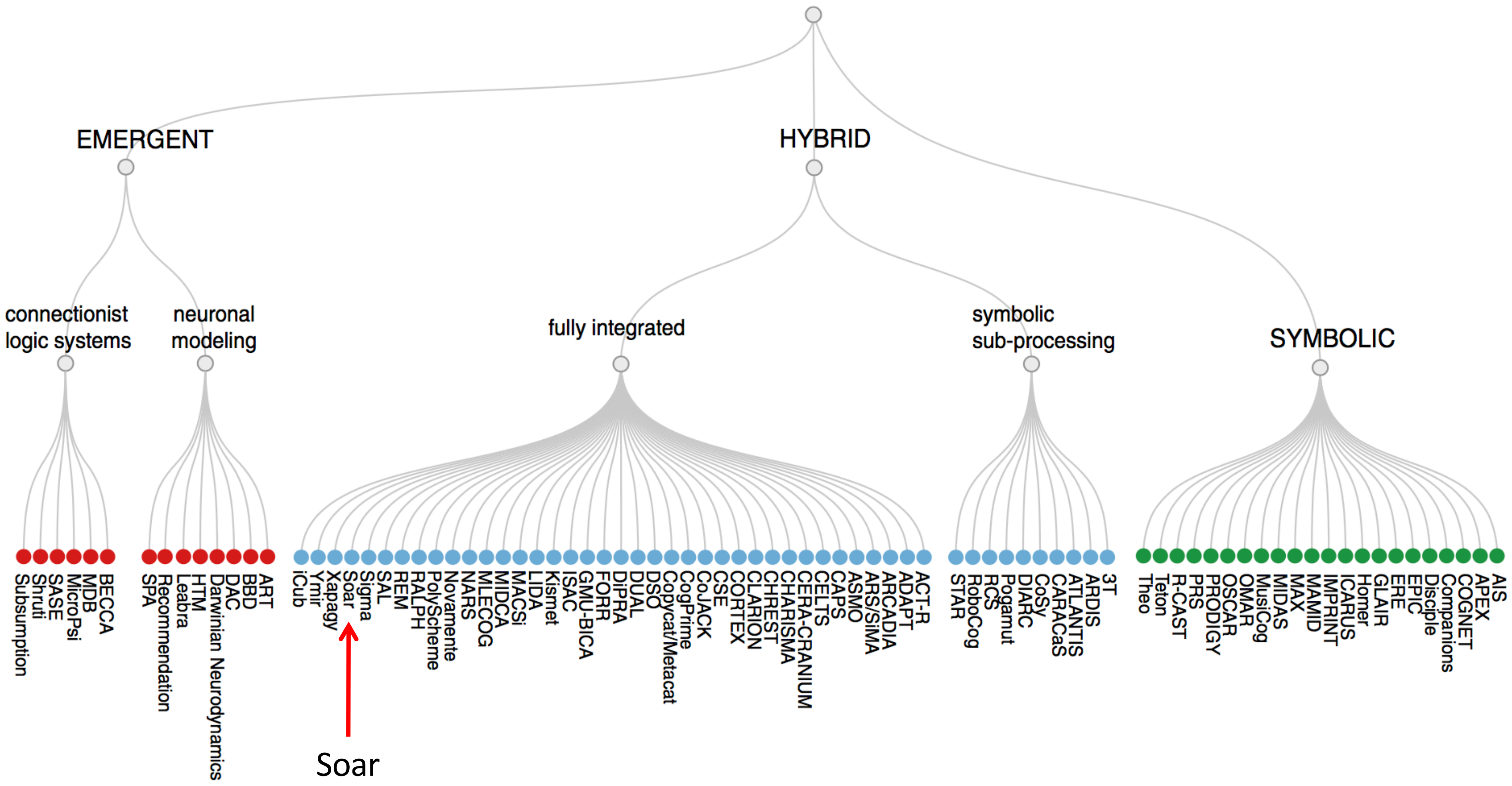


Cognitivist
Cognitive Architectures

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

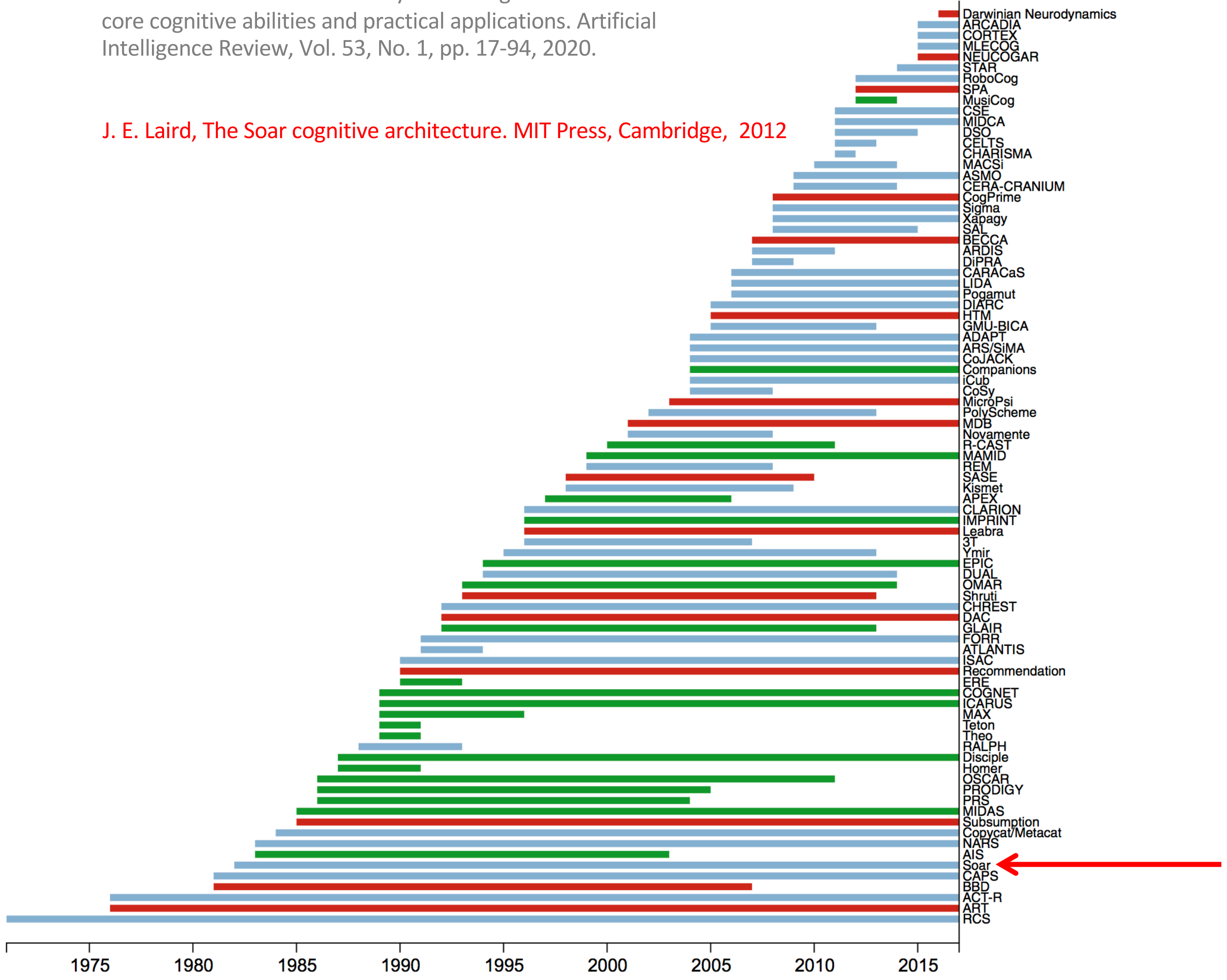


We will briefly sample four of the most well-known cognitive architectures

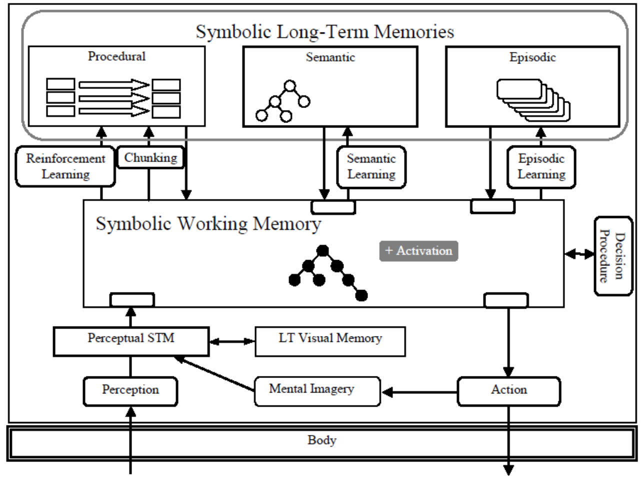

EMERGENT
HYBRID
connectionist logic systems
neuronal modeling
fully integrated
symbolic sub-processing
SYMBOLIC
Subsumption
Shruti
SASE
MicroPsi
MDB
BECCA
SPA
Recommendation
Leabra
HTM
Darwinian Neurodynamics
DAC
BBD
ART
iCub
Ymir
Xapagy
Soar
Sigma
SAL
REM
RALPH
PolyScheme
Novamente
NARS
MLECOG
MIDCA
MACSi
LIDA
Kismet
ISAC
GMU-BICA
FORR
DiPRA
DUAL
DSO
Copycat/Metacat
CogPrime
CoJACK
CSE
CORTEX
CLARION
CHREST
CHARISMA
CERA-CRANIUM
CELTS
CAPS
ASMO
ARS/SiMA
ARCADIA
ADAPT
ACT-R
STAR
RoboCog
RCS
Pogamut
DIARC
CoSy
CARACaS
ATLANTIS
ARDIS
3T
Theo
Teton
R-CAST
PRS
PRODIGY
OSCAR
OMAR
MusiCog
MIDAS
MAX
MAMID
IMPRINT
ICARUS
Homer
GLAIR
ERE
EPIC
Disciple
Companions
COGNET
APEX
AIS
Soar

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

J. E. Laird, The Soar cognitive architecture. MIT Press, Cambridge, 2012

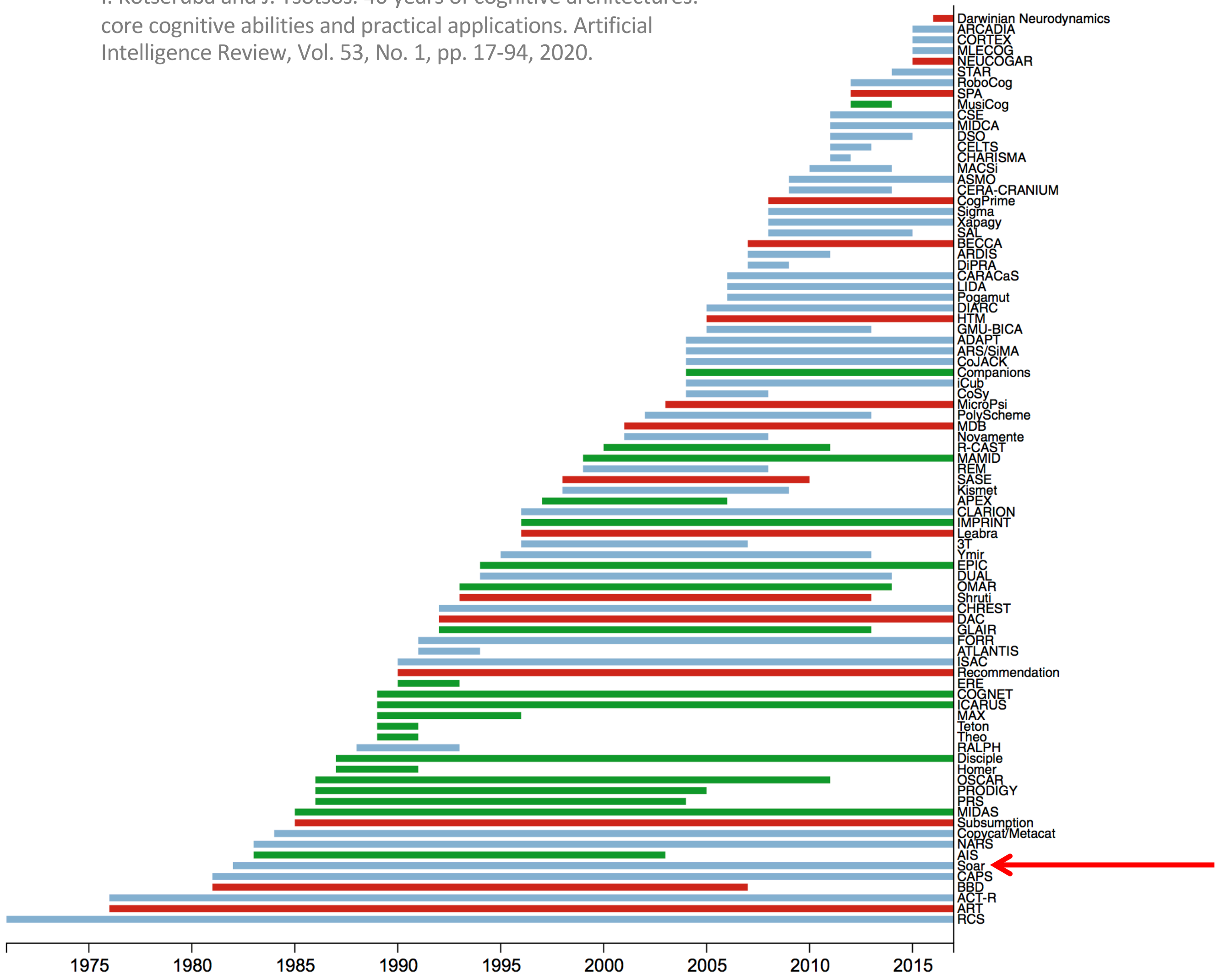


Soar



- 1983 – 2021; now version 9.6
- Production (rule-based) system
- Cyclic operation
  1. Production cycle: fire all rules that match information in the symbolic working memory; update memory, fire all rules …
  2. Decision cycle: select an action
- Universal sub-goaling: create a new goal and expose more knowledge when an impasse is encountered
- Learns a new rule when an impasse is resolved

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.



160 A Catalogue of Cognitive Architectures

A.1 Cognitivist Cognitive Architectures

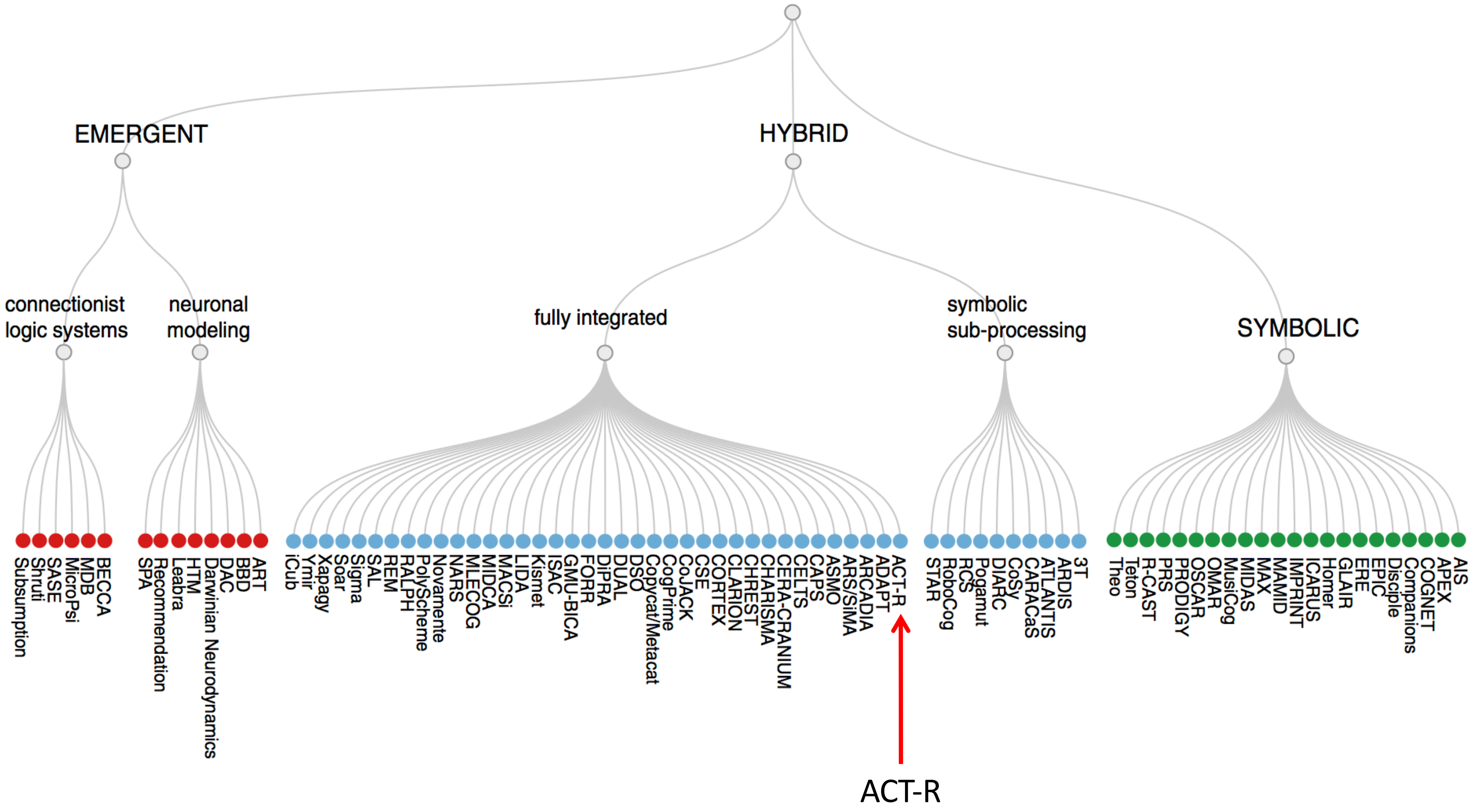
A.1.1 The Soar Cognitive Architecture

The Soar system [211, 326, 220, 222] is Newell's candidate for a Unified Theory of Cognition [271] and, as such, it is an architypal cognitivist cognitive architecture (as well as being an iconic one).It is a production (or rule-based) system[1] that operates in a cyclic manner, with a production cycle and a decision cycle. It operates as follows. First, all productions that match the contents of declarative (working) memory fire. A production that fires may alter the state of declarative memory and cause other productions to fire. This continues until no more productions fire. At this point, the decision cycle begins in which a single action from several possible actions is selected. The selection is based on stored action preferences. Thus, for each decision cycle there may have been many production cycles. Productions in Soar are low-level; that is to say, knowledge is encapsulated at a very small grain size.

One important aspect of the decision process concerns a process known as *universal sub-goaling*. Since there is no guarantee that the action preferences will be unambiguous or that they will lead to a unique action or indeed any action, the decision cycle may lead to an 'impasse'. If this happens, Soar sets up an new state in a new problem space — sub-goaling — with the goal of resolving the impasse. Resolving one impasse may cause others and the sub-goaling process continues. It is assumed that degenerate cases can be dealt with (*e.g.* if all else fails, choose randomly between two actions). Whenever an impasse is resolved, Soar creates a new production rule which summarizes the processing that occurred in the sub-state in solving the sub-goal. Thus, resolving an impasse alters the system super-state, *i.e.* the state in which the impasse originally occurred. This change is called a result and becomes the outcome of the production rule. The condition for the production rule to fire is derived from a dependency analysis: finding what declarative memory items matched in the course of determining the result. This change in state is a form of learning and it is the only form that occurs in Soar, *i.e.* Soar only learns new production rules. Since impasses occur often in Soar, learning is pervasive in Soar's operation.

[1] A production is effectively an IF-THEN condition-action pair. A production system is a set of production rules and a computational engine for interpreting or executing productions.

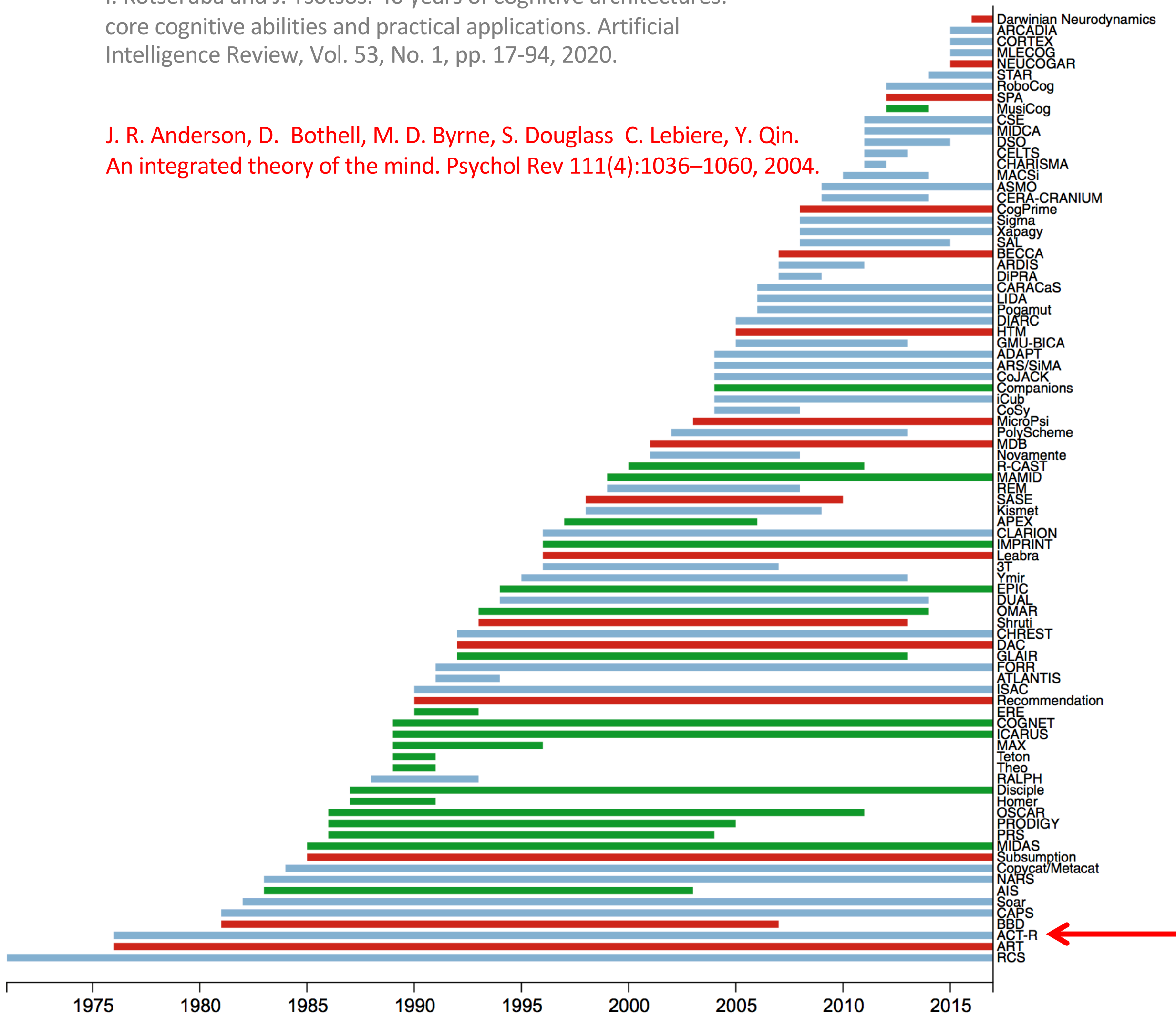
D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.



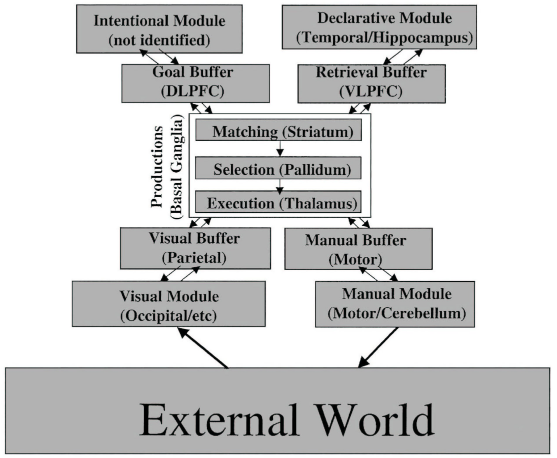
I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass C. Lebiere, Y. Qin. An integrated theory of the mind. Psychol Rev 111(4):1036–1060, 2004.



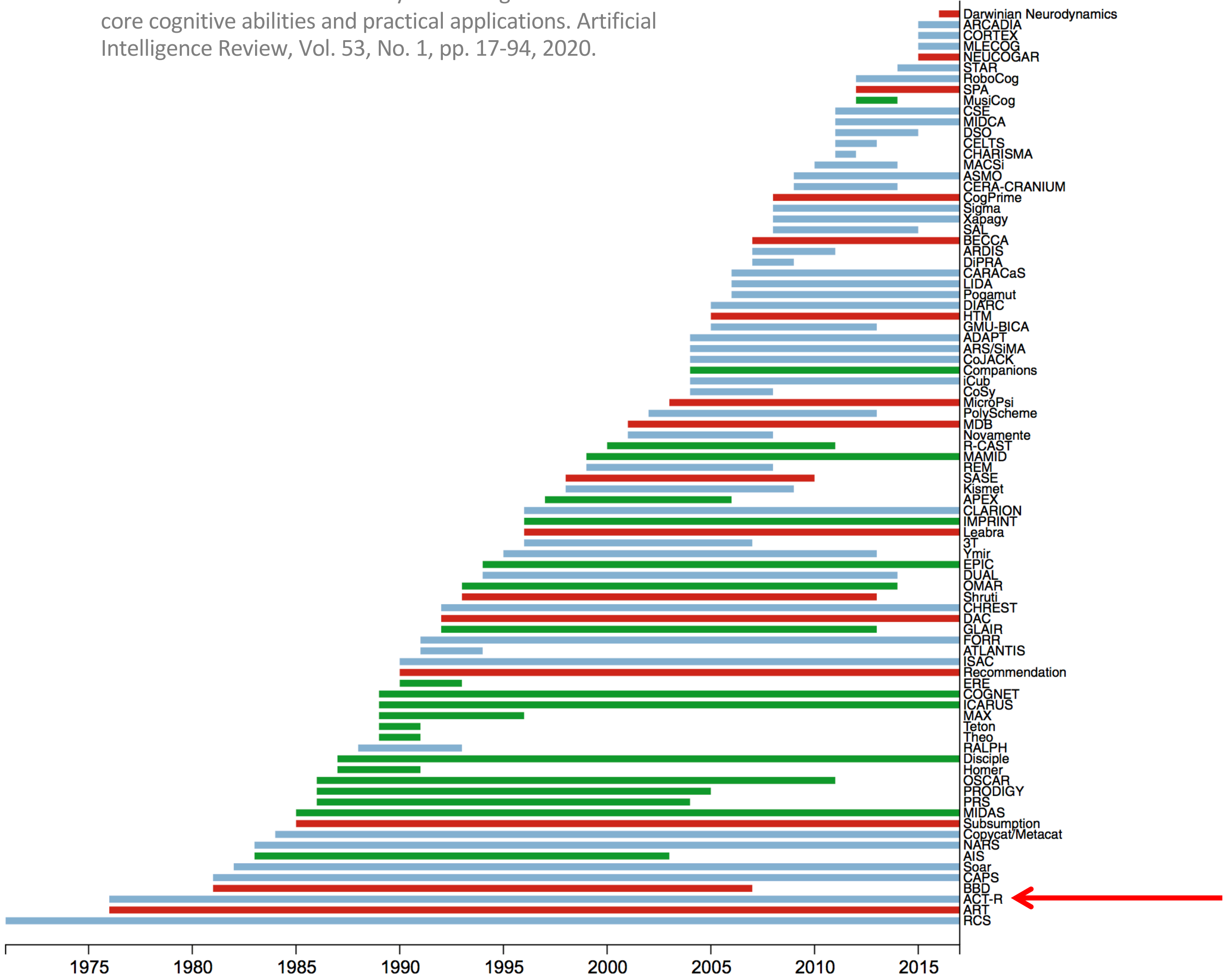
ACT-R



- 1996, 2004; now version 7
- Production system with five modules: Intentional, Declarative, Visual, Manual, Production
- Most modules are mapped onto a specific area in the brain
- Cyclic operation: executes one production per cycle
  - Pattern of information in the buffers is recognized
  - A single production is selected and fires
  - The buffers are updated

Artificial Intelligence

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.



162 A Catalogue of Cognitive Architectures

A.1.3 *ACT-R — Adaptive Control of Thought - Rational*

The ACT-R [6, 7] cognitive architecture is a widely-regarded candidate for a unified theory of cognition. It focusses on modular decomposition and offers a theory of how these modules are integrated to produce coherent cognition. The architecture comprises five specialized modules, each devoted to processing a different kind of information (see Figure A.1). There is a vision module for determining the identity and position of objects in the visual field, a manual module for controlling hands, a declarative module for retrieving information from long-term information, and a goal module for keeping track of the internal state when solving a problem. Finally, it also has a production system that coordinates the operation of the other four modules. It does this indirectly via four buffers into which each module places a limited amount of information.

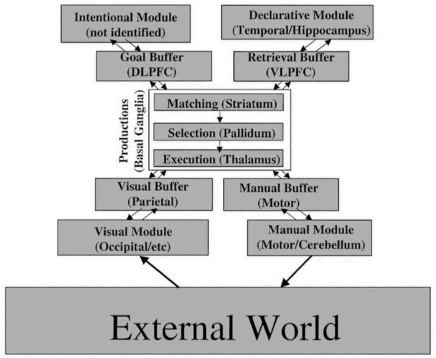
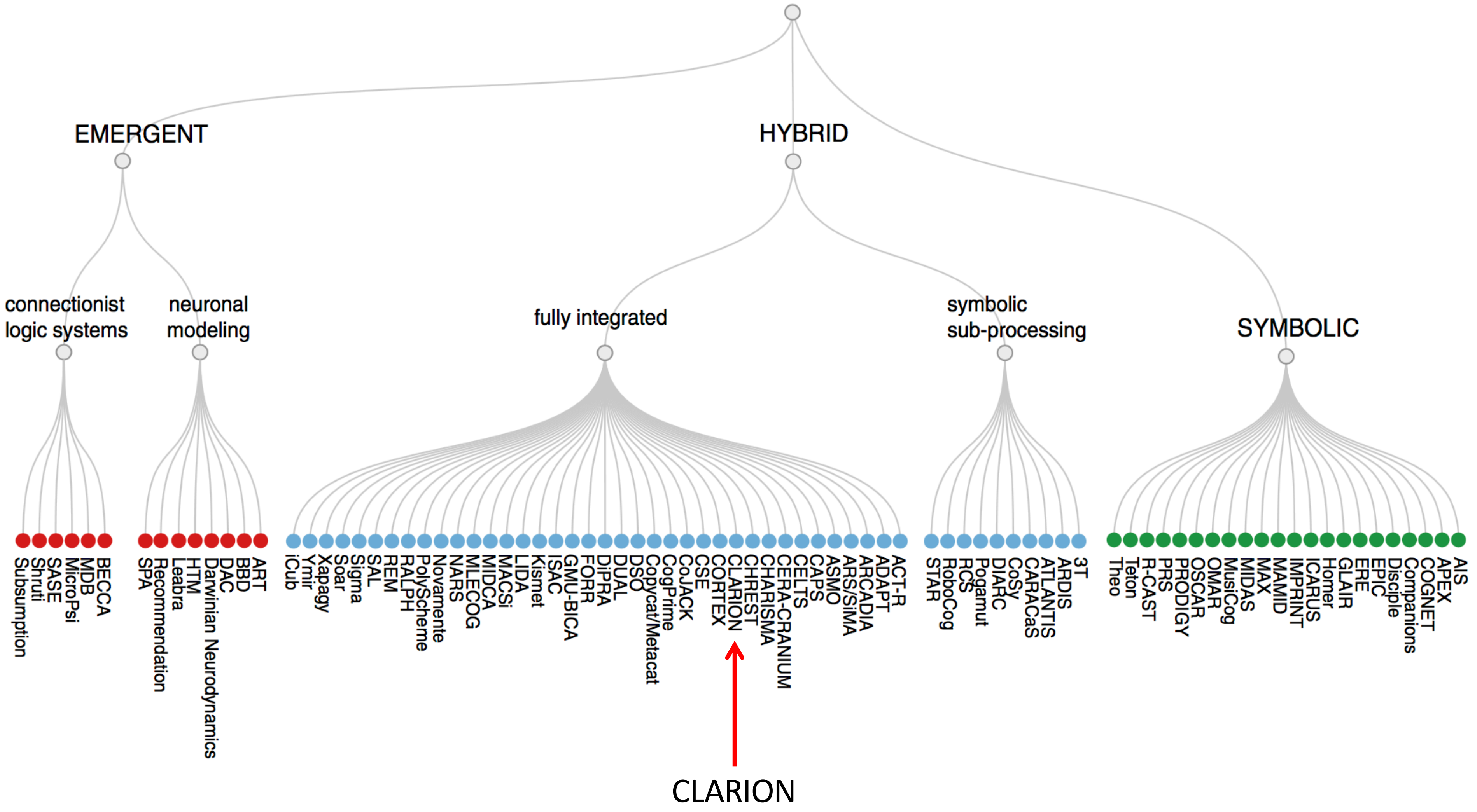


Fig. A.1 The ACT-R Cognitive Architecture (from [7])

ACT-R operates in a cyclic manner in which the patterns of information held in the buffers (and determined by external world and internal modules) are recognized, a single production fires, and the buffers are updated. It is assumed that this cycle takes approximately 50 ms.

There are two serial bottle-necks in ACT-R. One is that the content of any buffer is limited to a single declarative unit of knowledge, called a 'chunk'. This implies that only one memory can be retrieved at a time and indeed that a single object can be encoded in the visual field at any one time. The second bottle-neck is that only one production is selected to fire in any one cycle. This contrasts with both Soar and

D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.



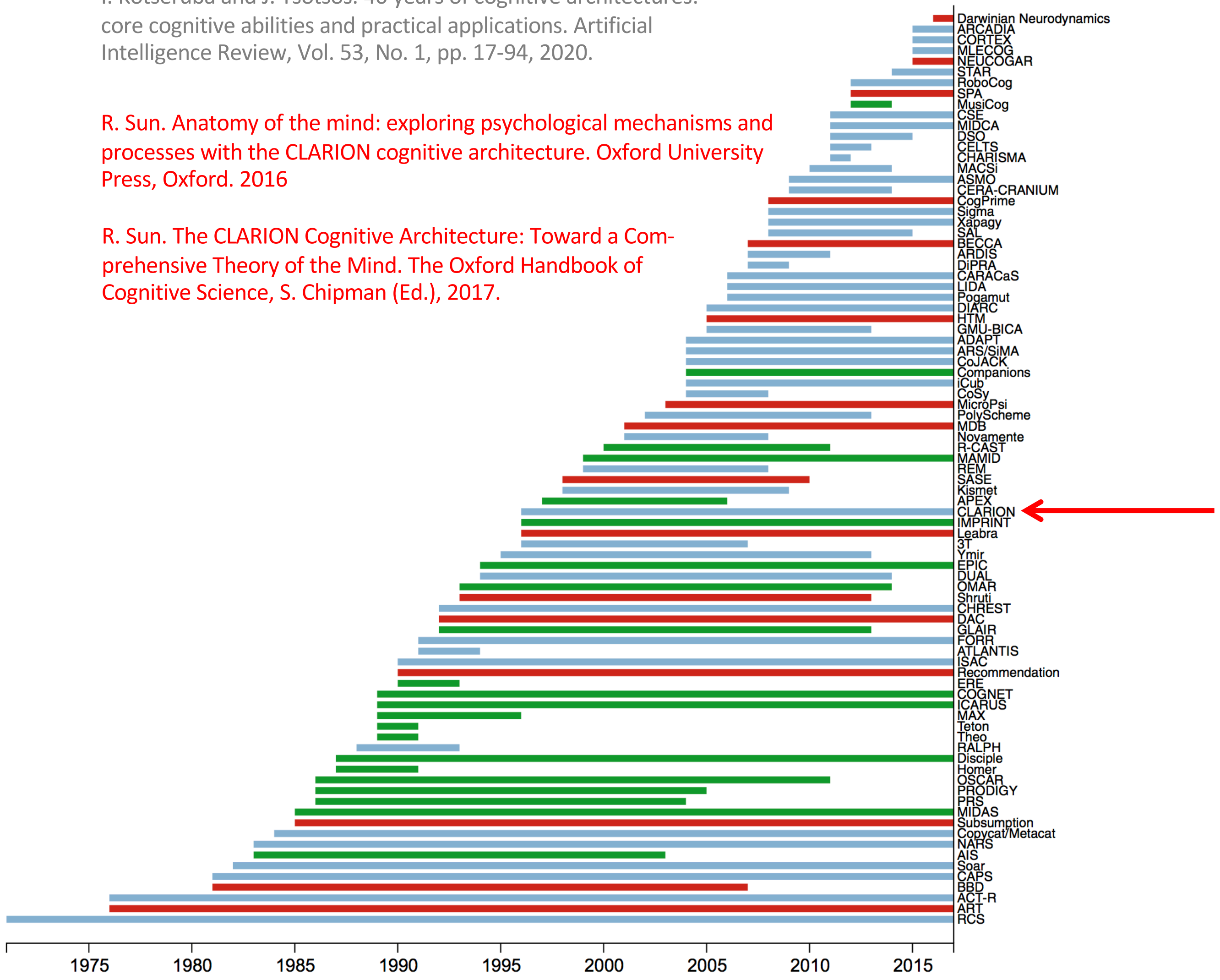
I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

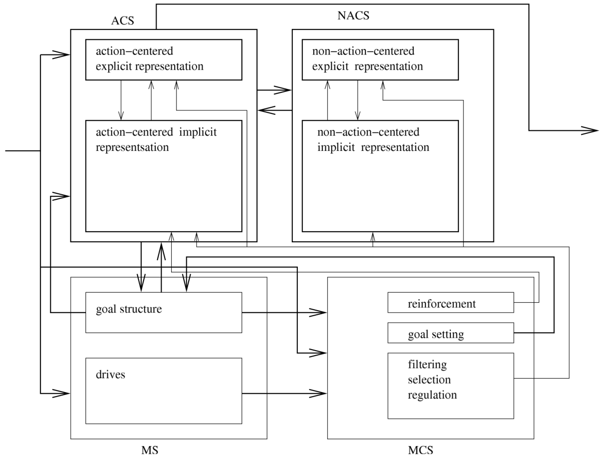
CLARION

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

R. Sun. Anatomy of the mind: exploring psychological mechanisms and processes with the CLARION cognitive architecture. Oxford University Press, Oxford. 2016

R. Sun. The CLARION Cognitive Architecture: Toward a Comprehensive Theory of the Mind. The Oxford Handbook of Cognitive Science, S. Chipman (Ed.), 2017.





Four sub-systems

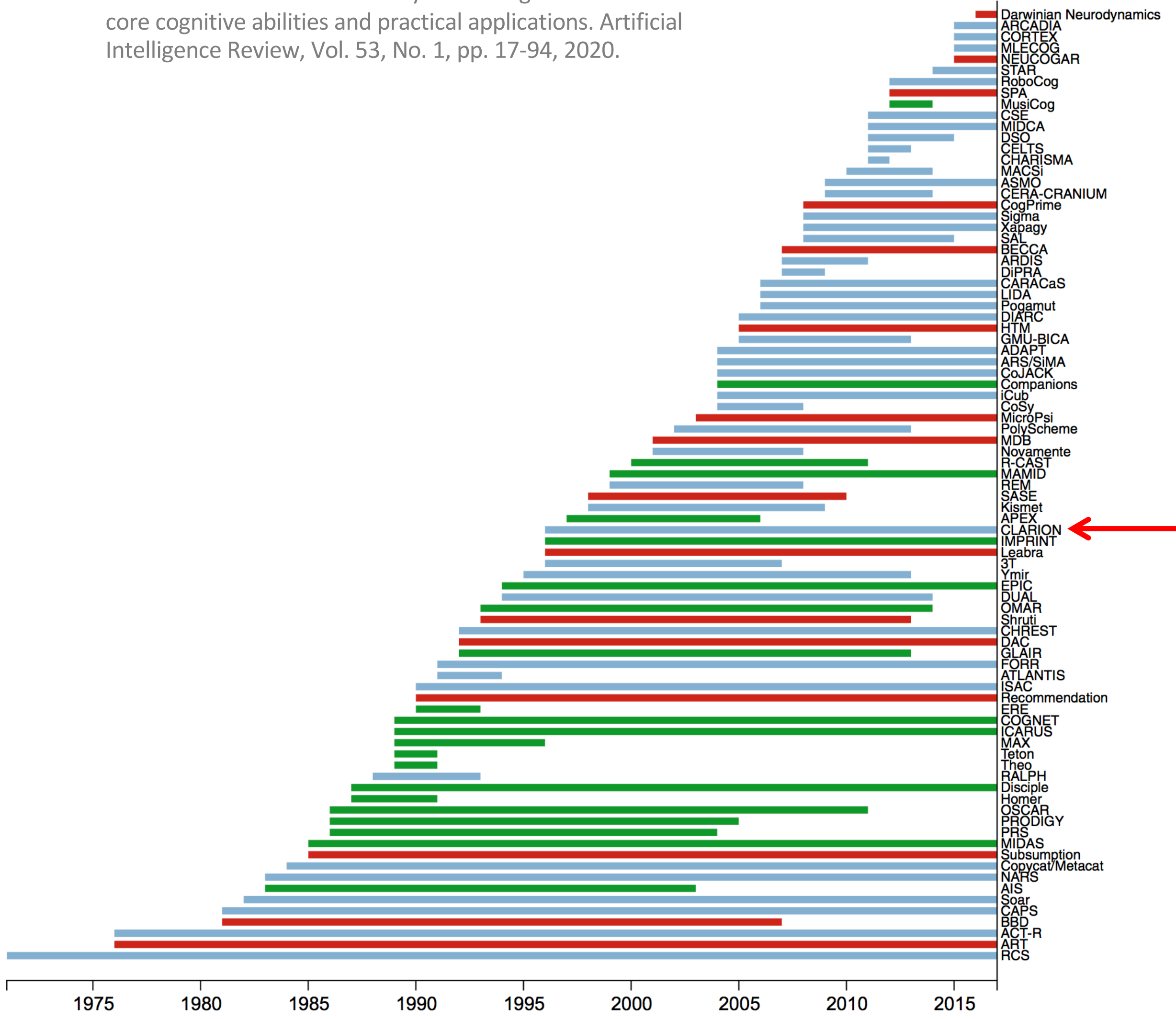
1. ACS – Action-centred subsystem
2. NACS – Non-action-centred subsystem
3. MS – Motivational subsystem
4. MCS – meta-cognitive subsystem

All four subsystems have two levels of knowledge representation

- Implicit connectionist bottom level
- Explicit symbolic top level

Implicit and explicit levels interact and cooperate both in action selection and in learning

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.



194 A Catalogue of Cognitive Architectures

A.3.6 *The CLARION Cognitive Architecture*

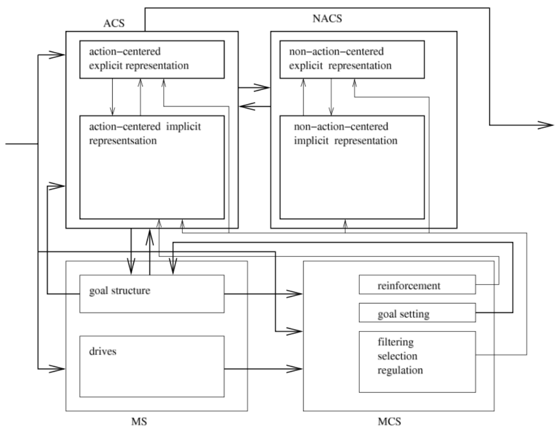


Fig. A.10 The CLARION hybrid cognitive architecture (from [364]). ACS stand for the action-centered subsystem, NACS for the non-action-centred subsystem, MS for the motivational subsystem, and MCS for the meta-cognitive subsystem. All four subsystems have two types of representation: implicit (connectionist) and explicit (symbolic).

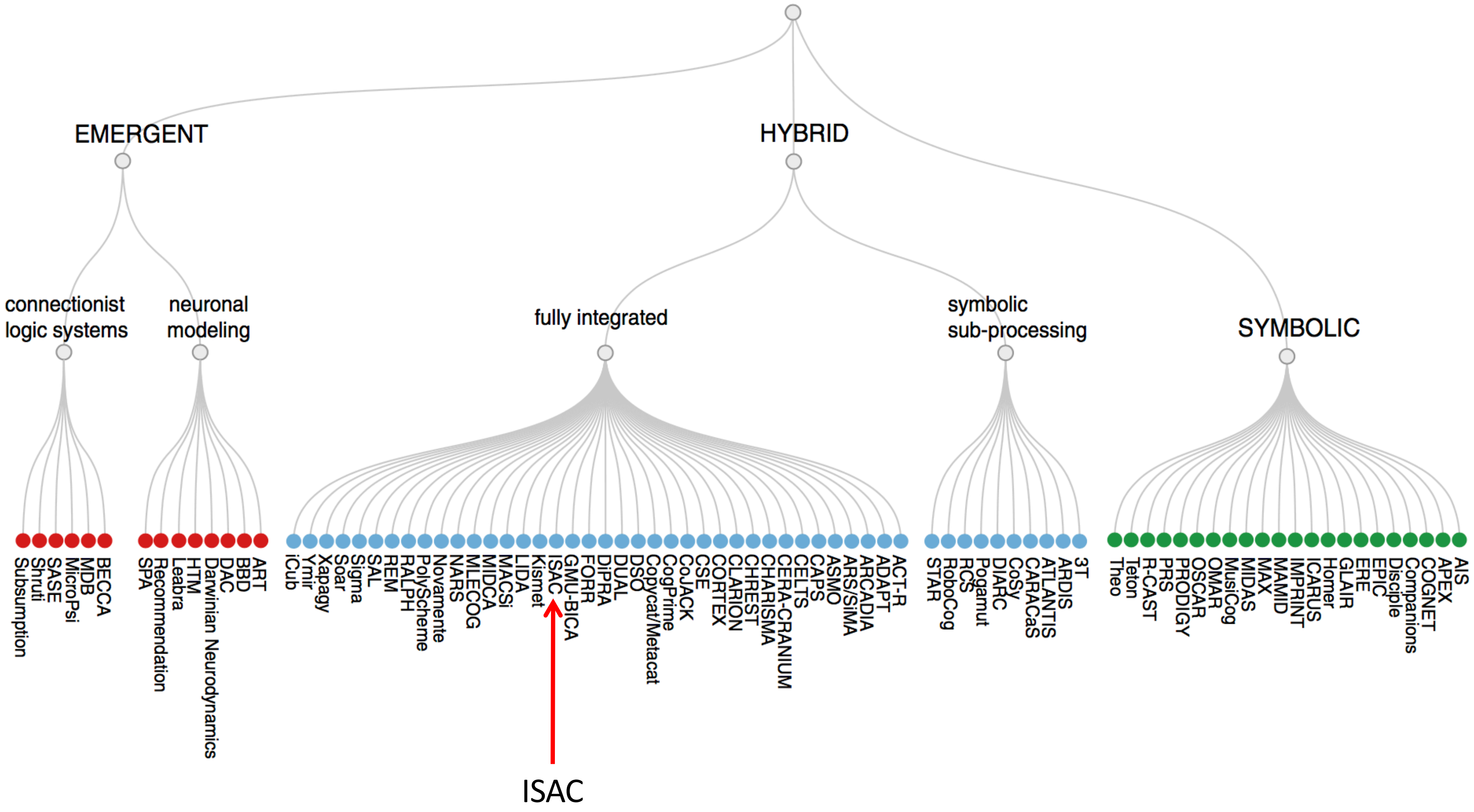
CLARION [362, 363, 364] is an architypal hybrid cognitive architecture, deploying both connectionist and symbolic representations. It comprises four subsystems:

1. An action-centred subsystem (ACS);
2. A non-action-centred subsystem (NACS);
3. A motivational subsystem (MS);
4. A meta-cognitive subsystem (MCS).

All four subsystems have two levels of knowledge representation: an implicit connectionist bottom level and an explicit symbolic top level. The implicit and explicit levels interact and cooperate both in action selection and in learning.

The action-centred subsystem controls both external physical movements and internal "mental" operations. Given some observational state, i.e. a set of sensory features, the bottom level evaluates the desirability of all possible actions. The desirability is learned by reinforcement learning using the Q-Learning algorithm [392]. At the same time, the top level identifies possible actions from a rule network, again based on the observed sensory features. The bottom-level and top-level action are compared and the most appropriate top-level action is selected and executed. The

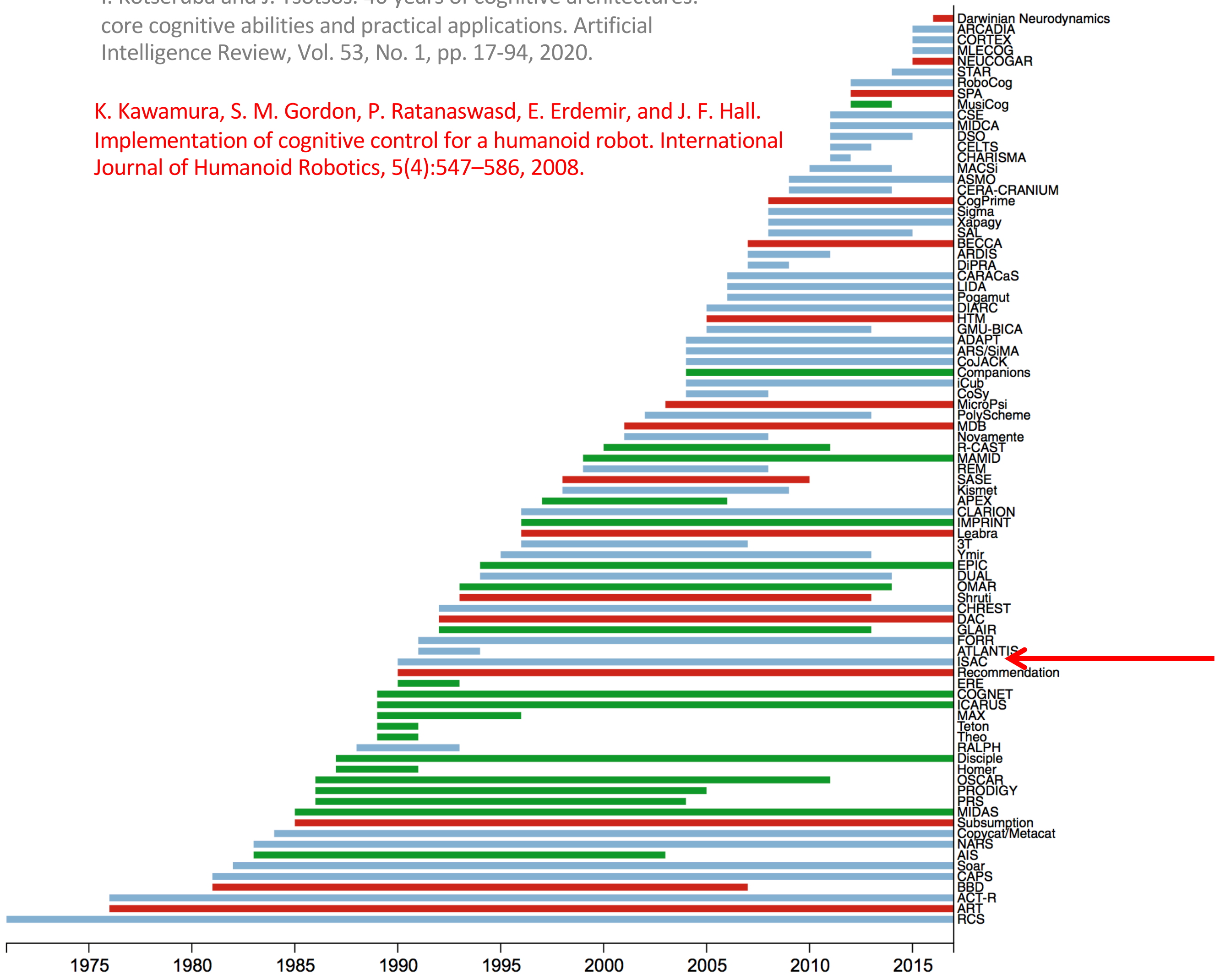
D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.



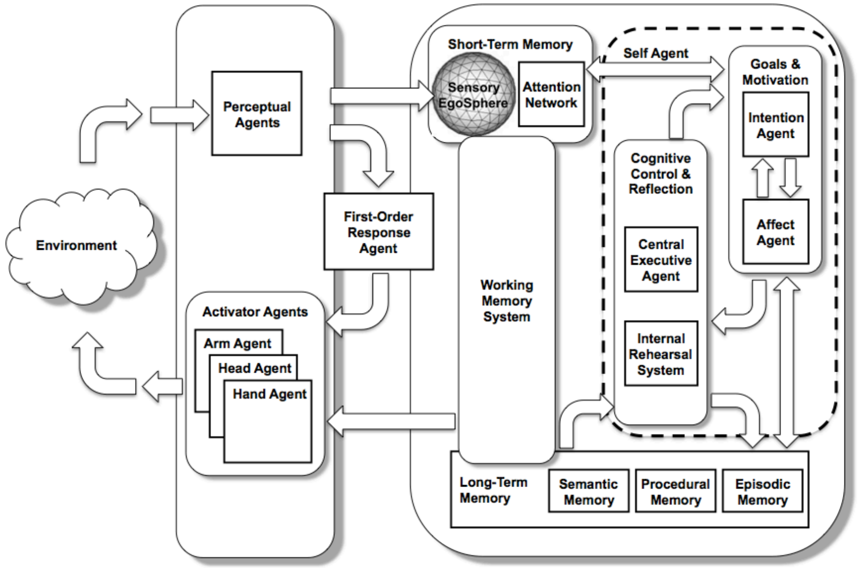
We will now study one of these cognitive architectures in a little more detail

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

K. Kawamura, S. M. Gordon, P. Ratanaswasd, E. Erdemir, and J. F. Hall. Implementation of cognitive control for a humanoid robot. International Journal of Humanoid Robotics, 5(4):547–586, 2008.



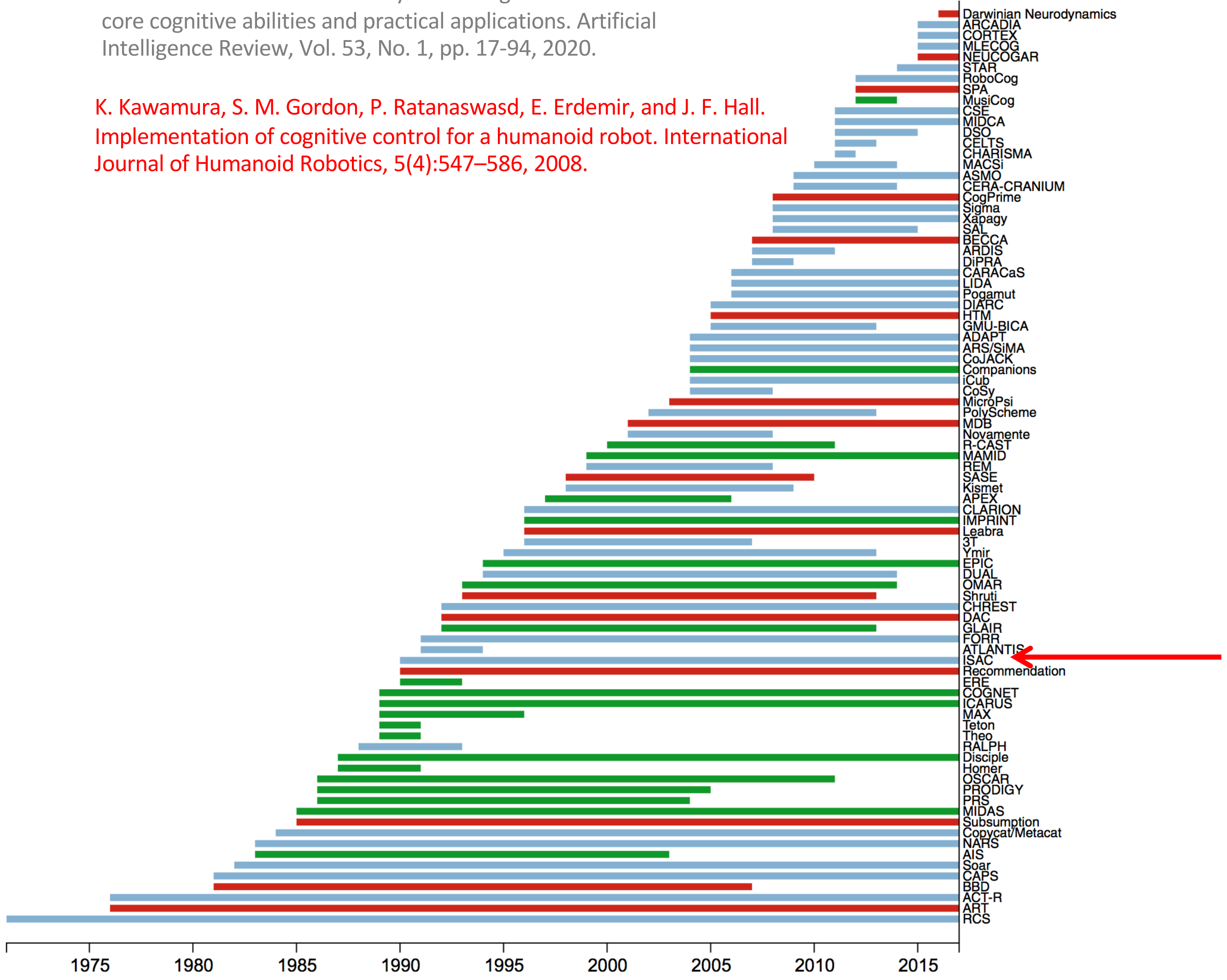
ISAC

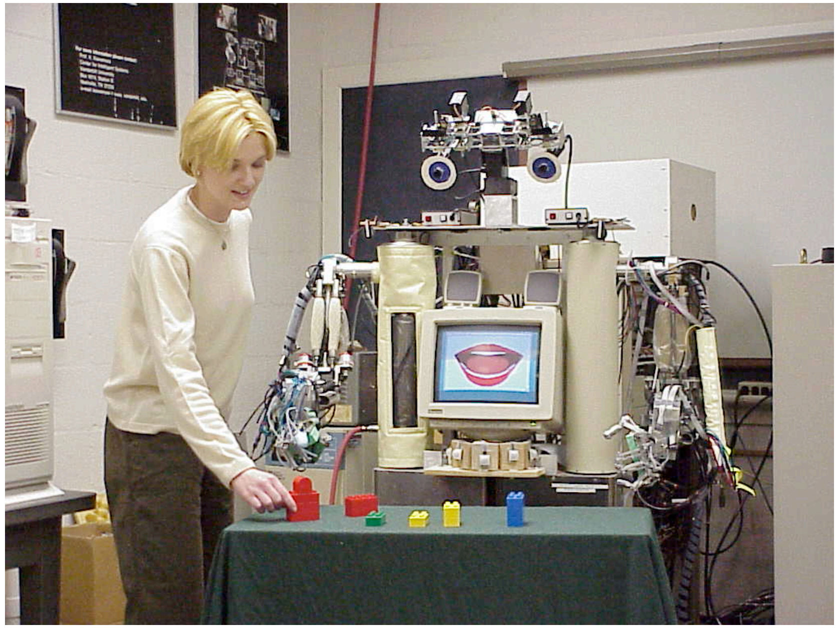


ISAC

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

K. Kawamura, S. M. Gordon, P. Ratanaswasd, E. Erdemir, and J. F. Hall. Implementation of cognitive control for a humanoid robot. International Journal of Humanoid Robotics, 5(4):547–586, 2008.

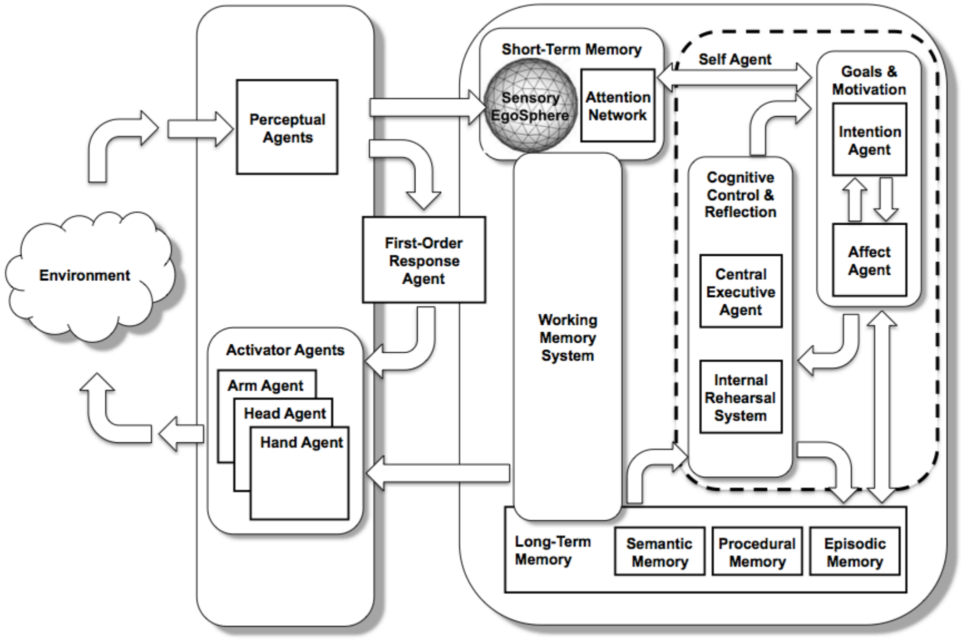




ISAC

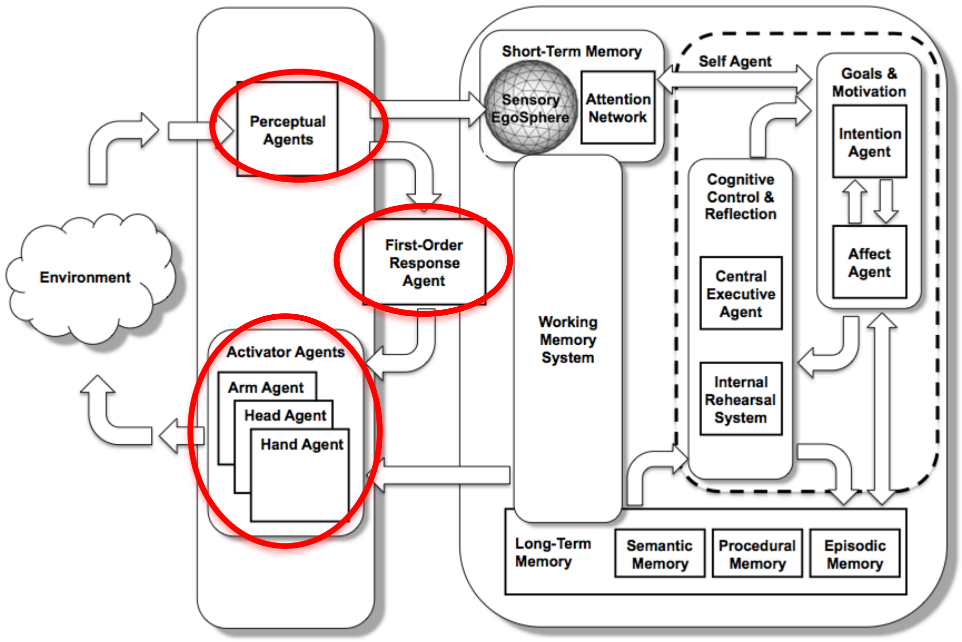
ISAC — Intelligent Soft Arm Control

- Hybrid cognitive architecture for an upper torso humanoid robot (also called ISAC)
- Comprises an integrated collection of software agents and associated memories
- Agents operate asynchonously and communicate with each other by message passing



ISAC

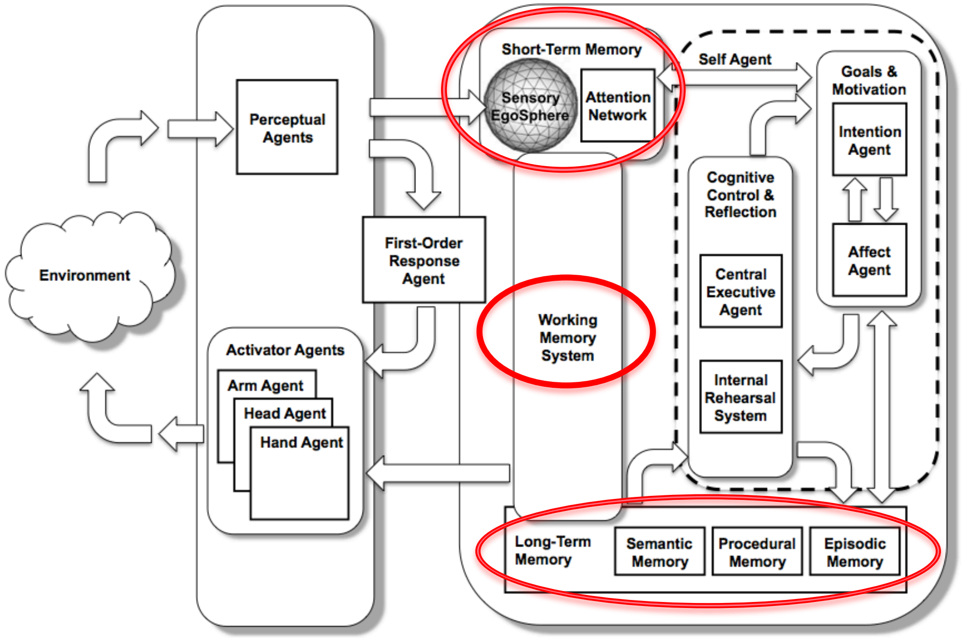
- Activator agents for motion control
- Perceptual agents
- First-order Response Agent (FRA) to effect reactive perception-action control



ISAC

Three memory systems

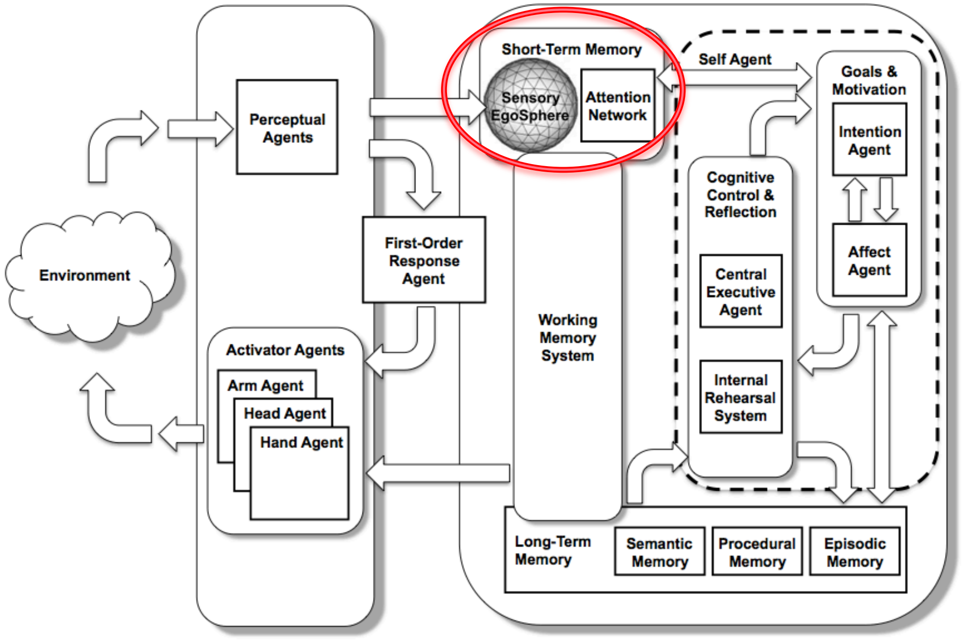
- Short-term memory (STM)
- Long-term memory (LTM)
- Working memory system (WMS)



ISAC

Short-term Memory

- Robot-centred spatio-temporal memory of the current perceptual events
- Sensory EgoSphere (SES)
  - Discrete representation of what is happening around the robot
  - Represented by a geodesic sphere indexed by two angles
- Attentional network
  - Determines the perceptual events that are most relevant



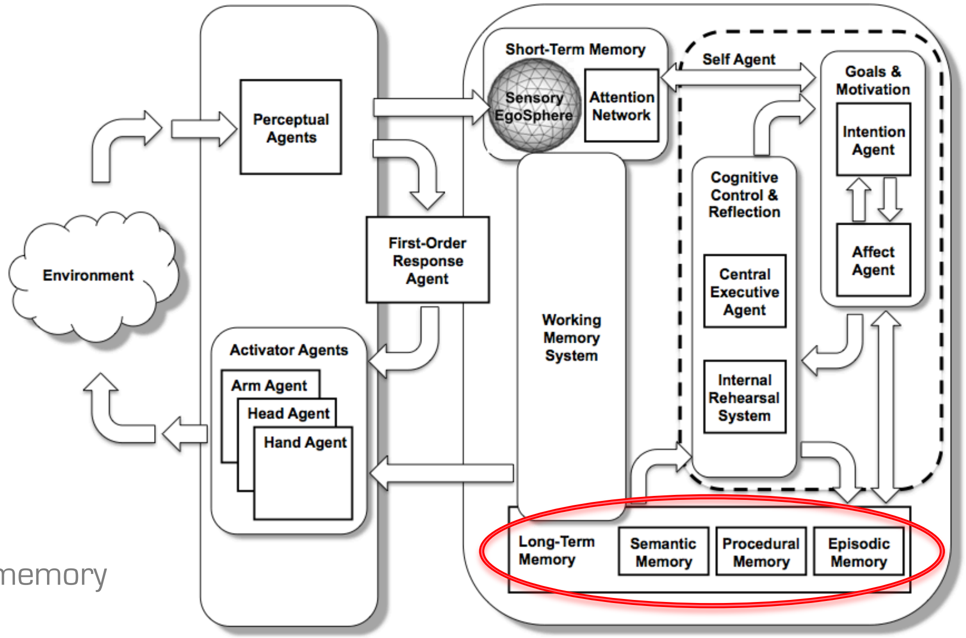
ISAC

Long-term Memory

- Stores information about the robot's learned skills and past experiences
- Semantic memory
- Episodic memory
- Procedural memory

Robot's declarative memory of the facts it knows

Representations of the motions it can perform

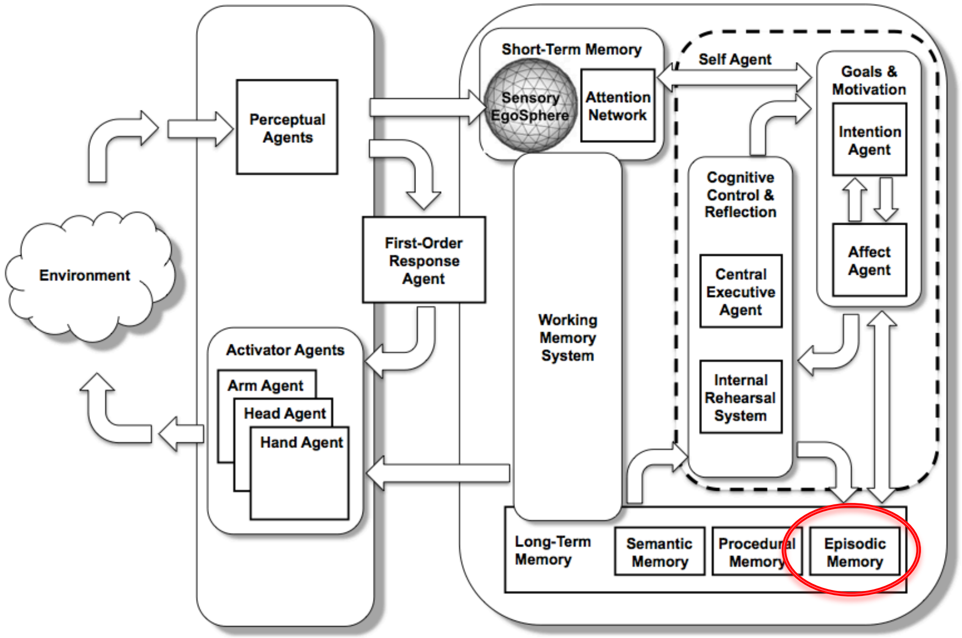


ISAC

Episodic memory

Abstracts past experiences & creates links or associations between them

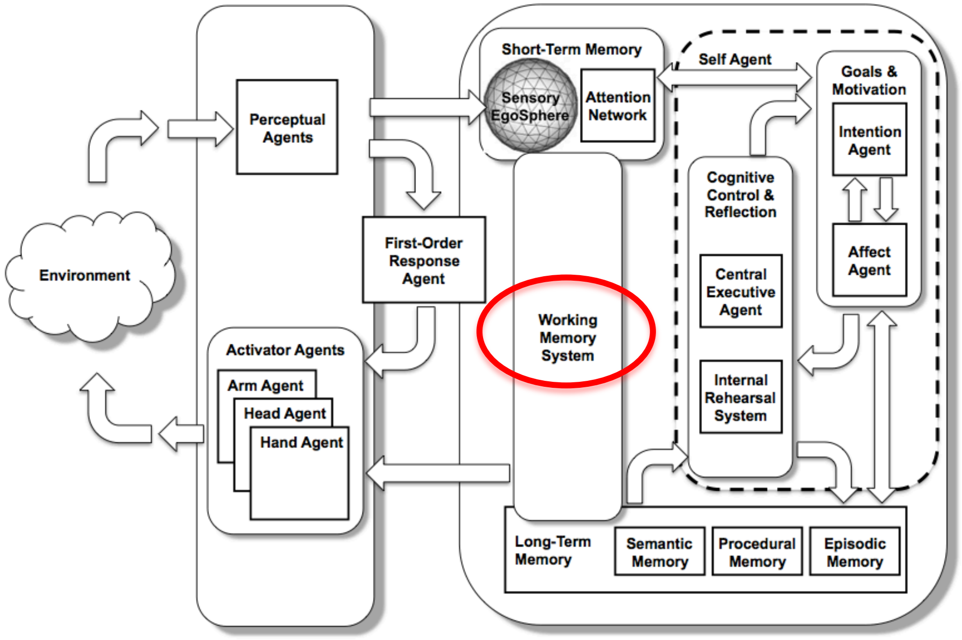
- External situation — i.e. task-relevant percepts from the SES
- Goals
- Emotions — i.e. internal evaluation of the perceived situation
- Actions
- Outcomes that arise from actions
- Valuations of these outcomes — e.g. how close they are to the desired goal state and any reward received at a result



ISAC

Working Memory System

- Temporarily stores information that is related to the task currently being executed
- A type of cache memory for STM and the information it stores, called chunks
- Encapsulates expectations of future reward (learned using a neural network)



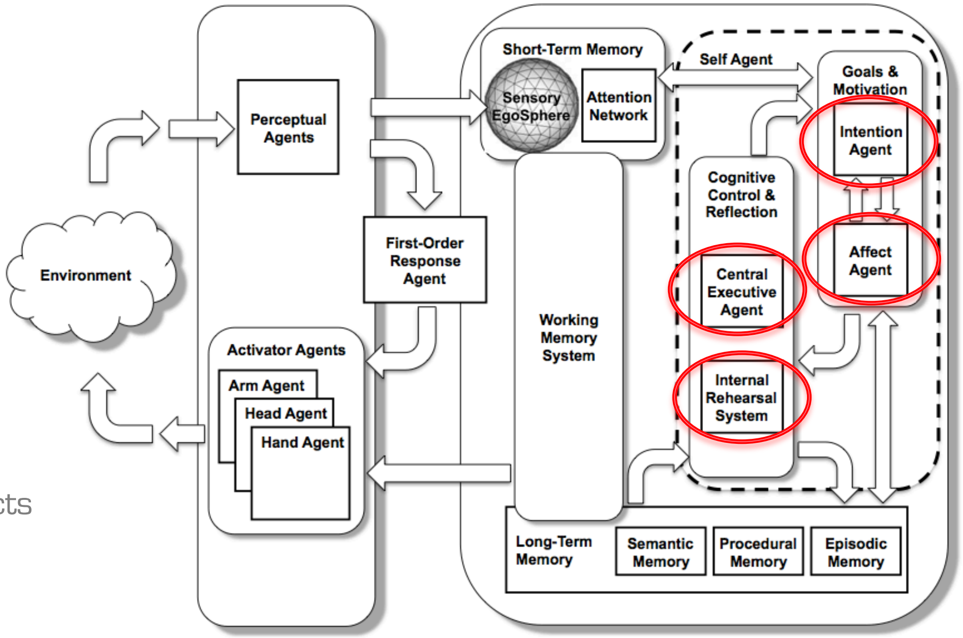
ISAC

Cognitive behaviour is achieved through the interaction of several agents

- Cognitive Control & Reflection sub-system
  - Central Executive Agent (CEA)
  - Internal Rehearsal System

Simulates the effects of possible actions

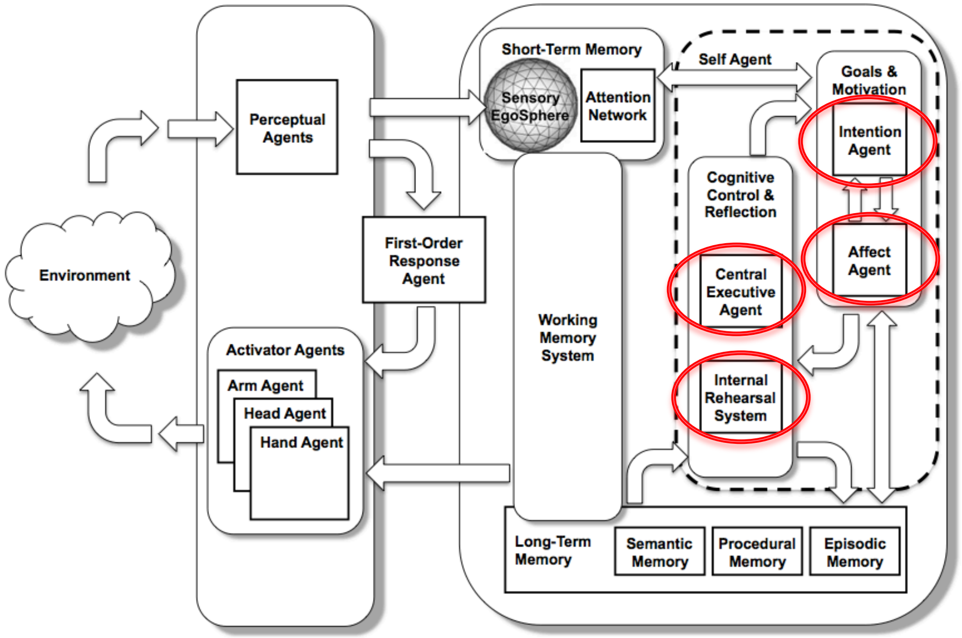
- Goals & Motivation sub-system
  - Intention Agent
  - Affect Agent



ISAC

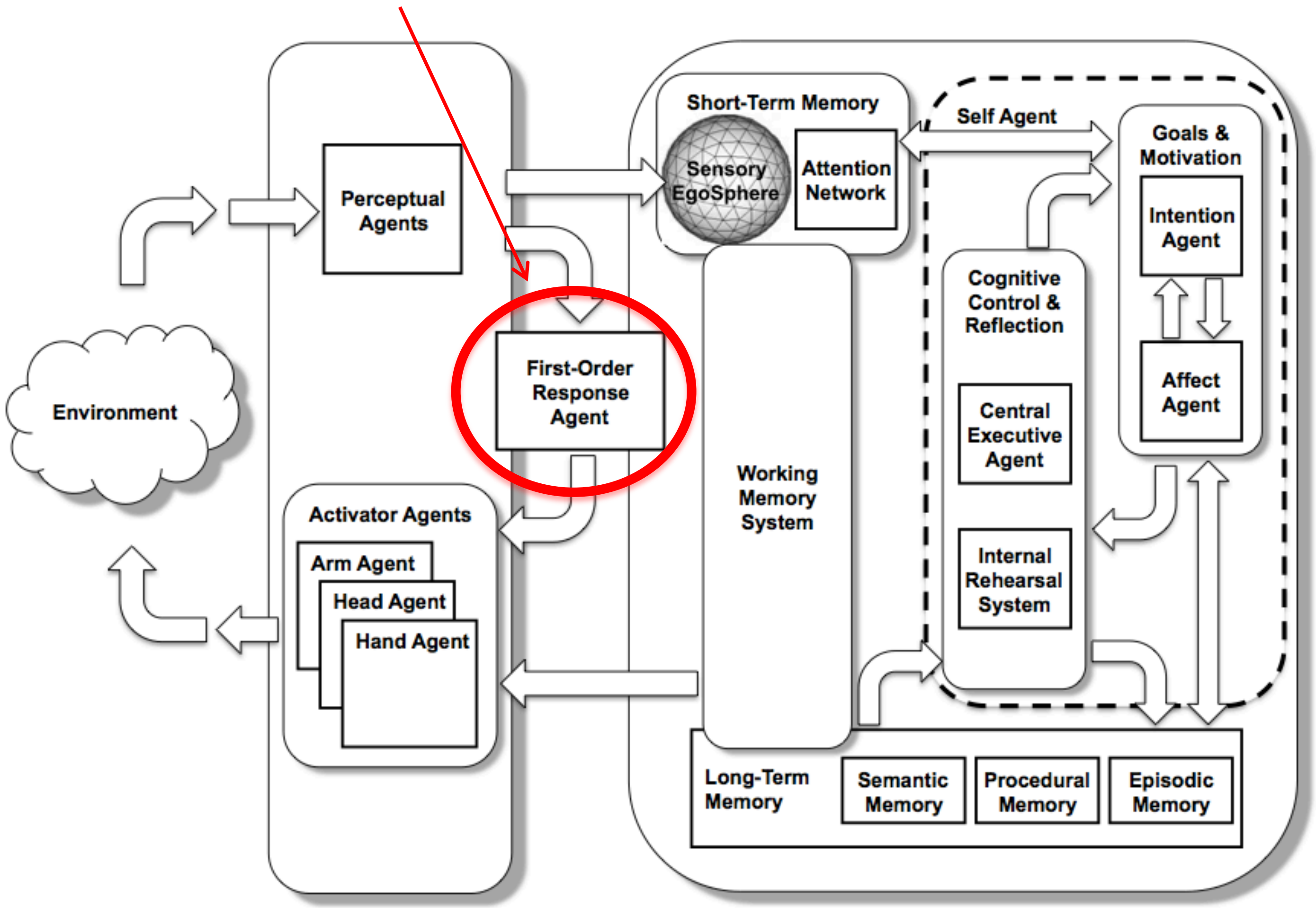
Cognitive behaviour is achieved through the interaction of several agents

- The CEA is responsible for cognitive control
- Invokes the skills required to perform some given task on the basis of the current focus of attention and past experiences
- The goals are provided by the Intention Agent
- Decision-making is modulated by the Affect Agent



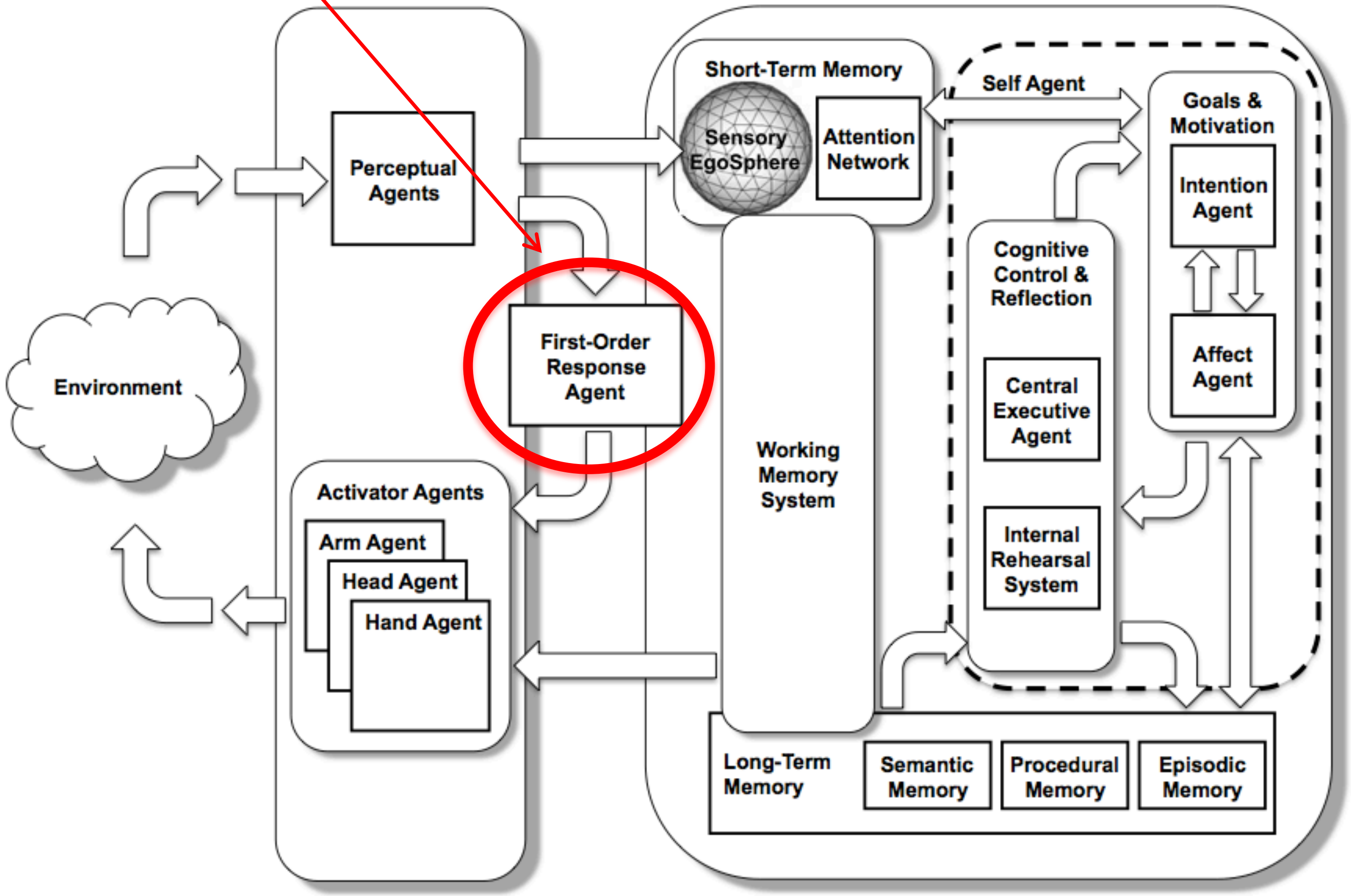
ISAC

Normally, the First-order Response Agent (FRA) produces reactive responses to sensory triggers



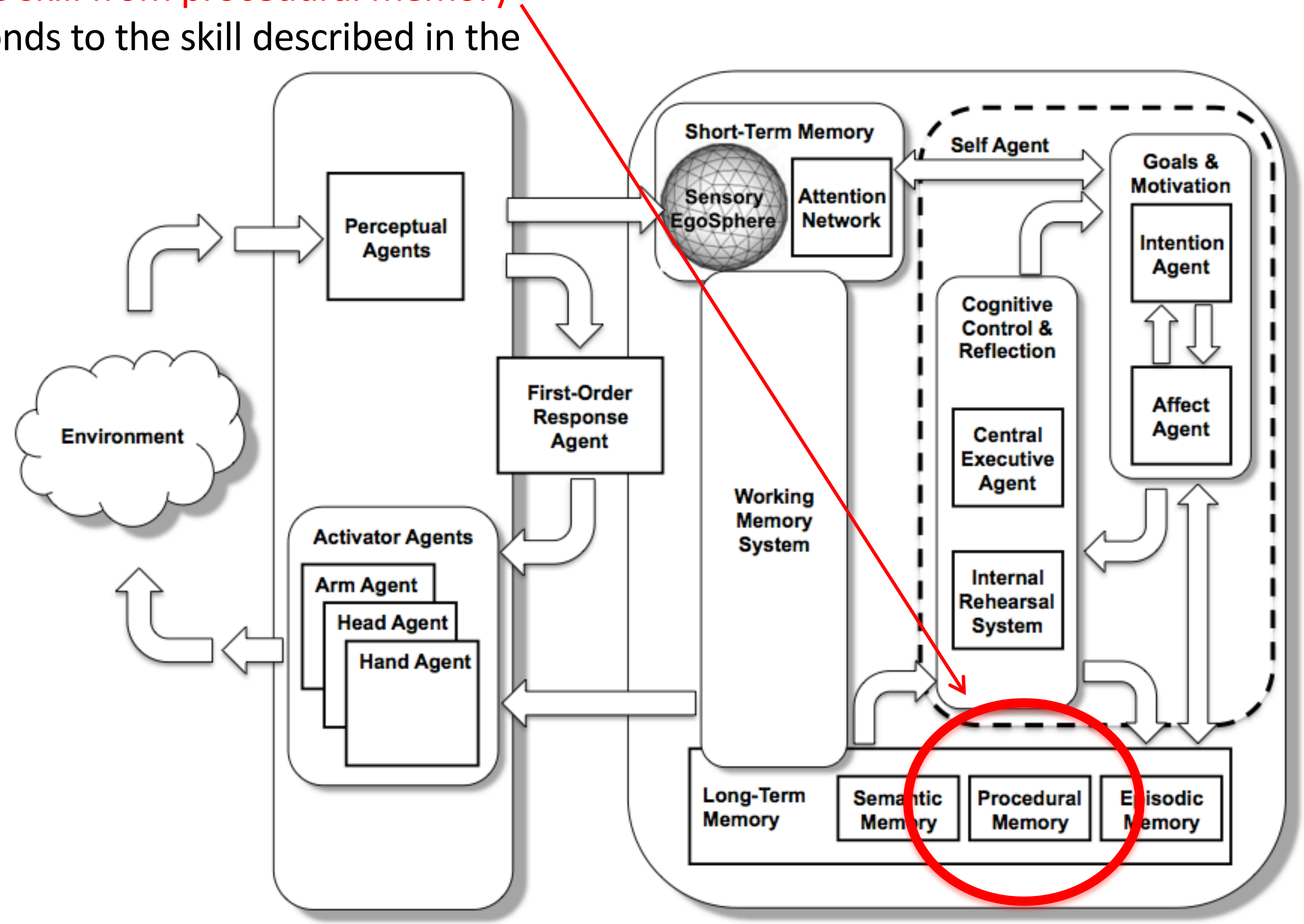
ISAC

First-order Response Agent (FRA)
is also responsible for executing tasks



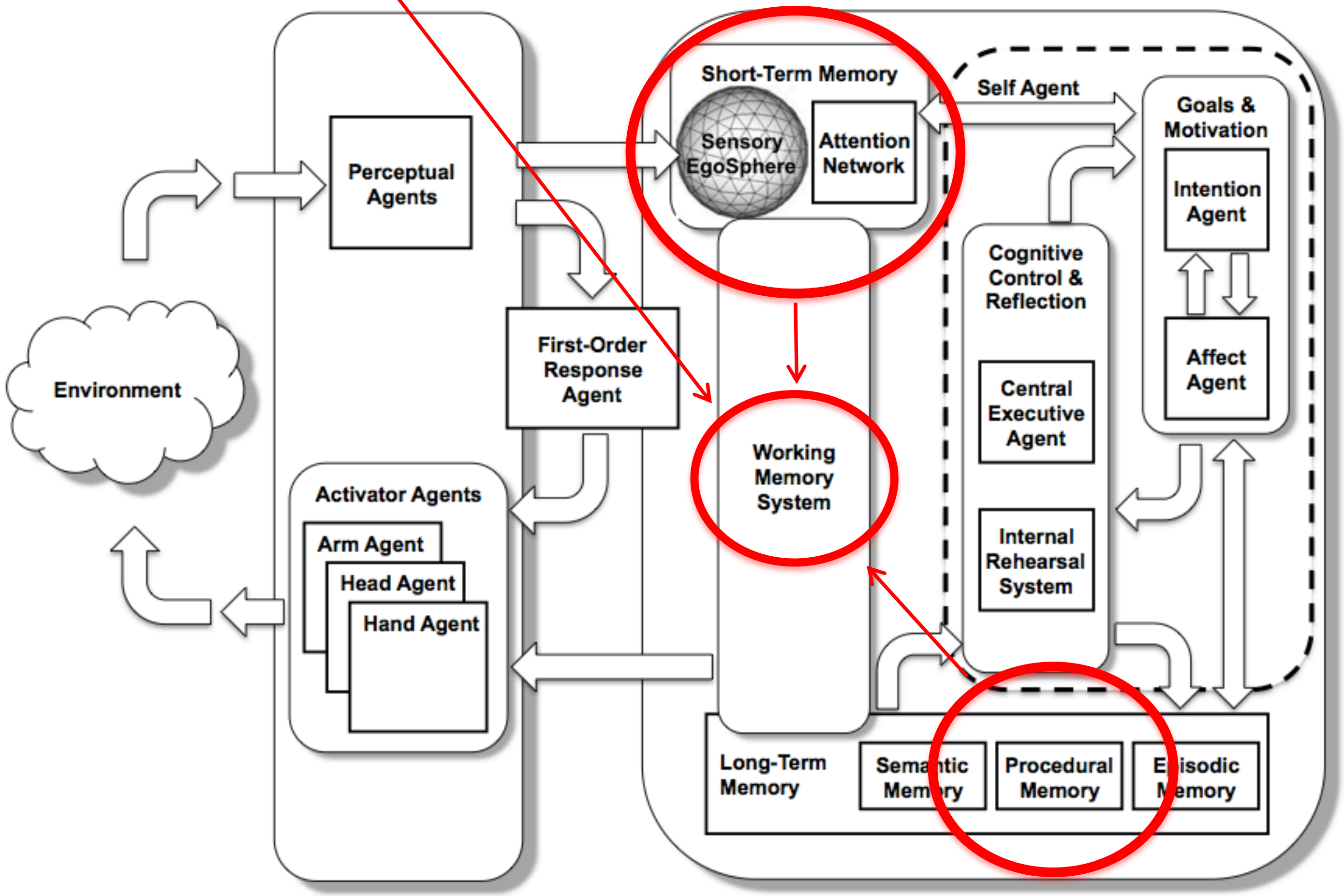
ISAC

When a task is assigned by a human,
the FRA retrieves the skill from procedural memory in LTM that corresponds to the skill described in the task information



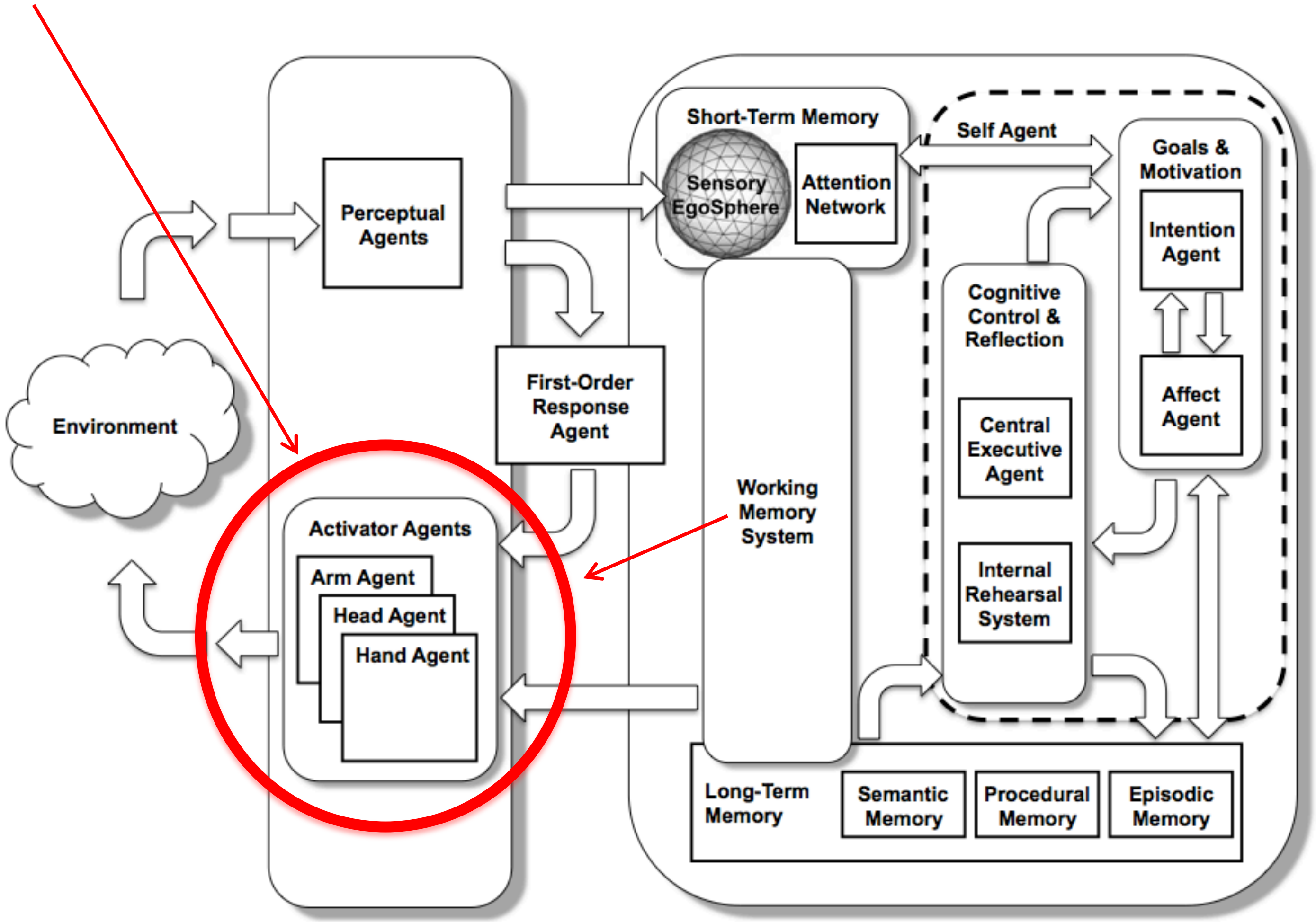
It then places it in the Working Memory System along with the current percept

ISAC



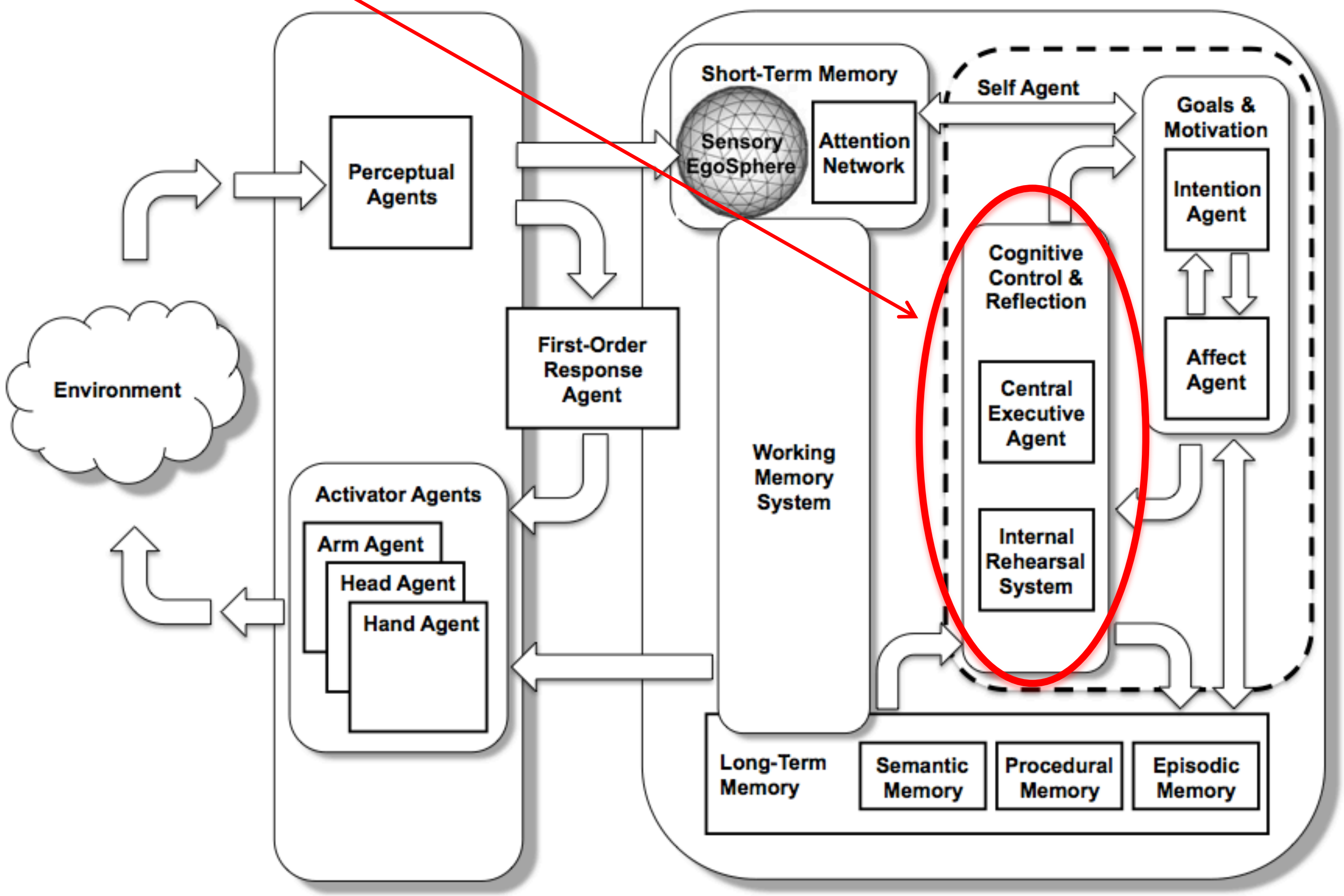
The Activator Agent then executes it, suspending execution whenever a reactive response is required

ISAC



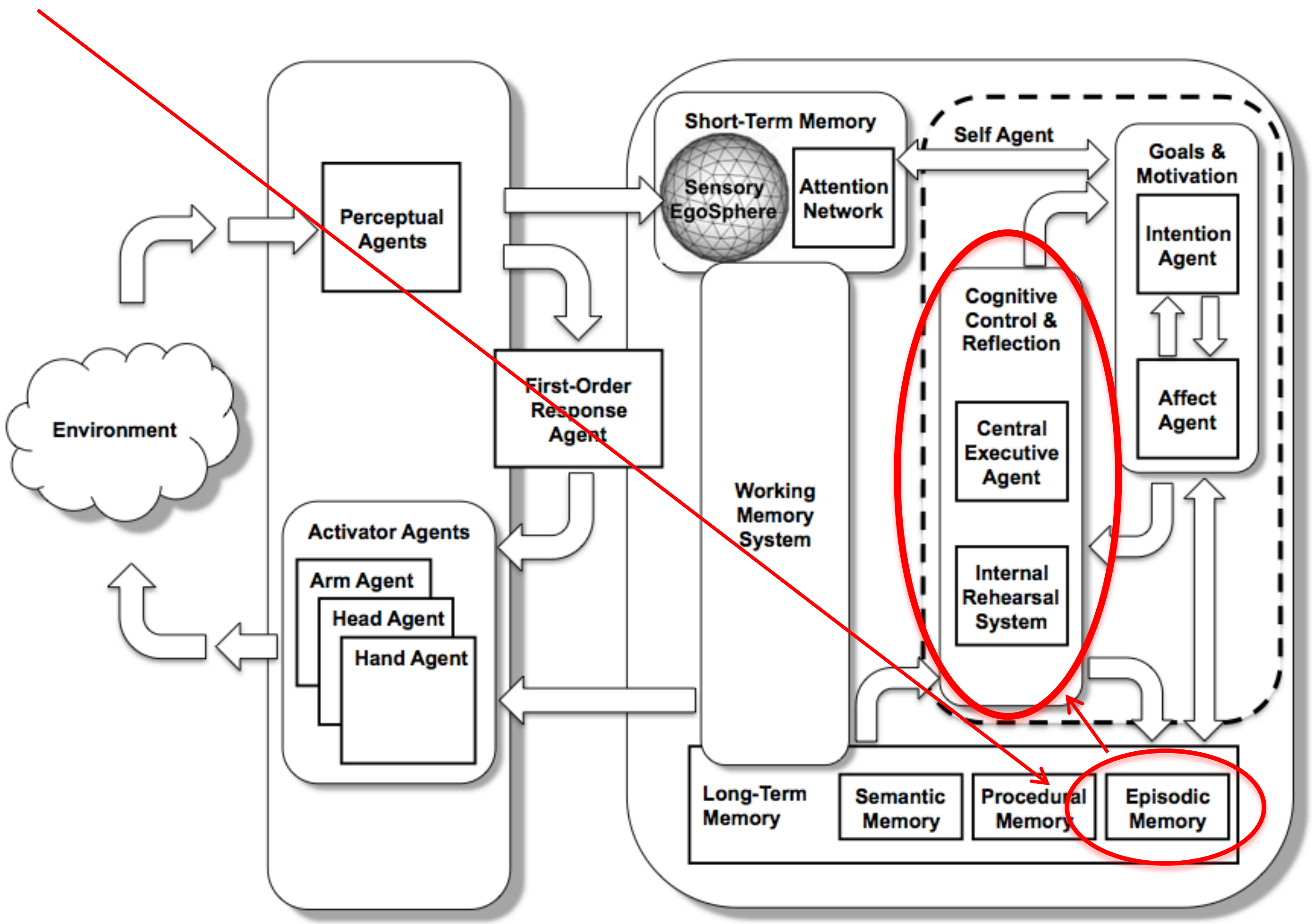
ISAC

If the FRA finds no matching skill for the task, the Central Executive Agent takes over



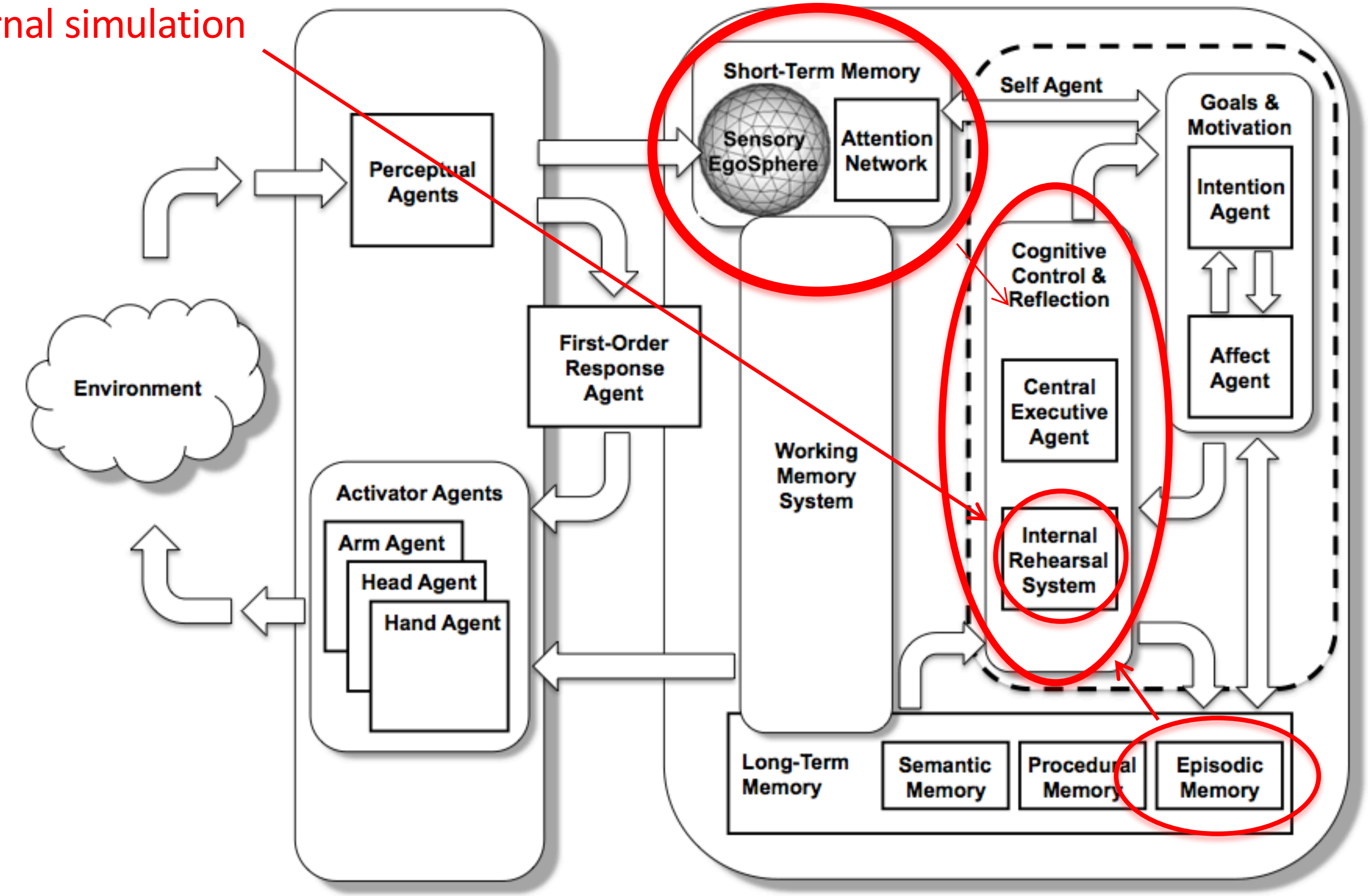
ISAC

Recalls from episodic memory past experiences and behaviours that contain information similar to the current task



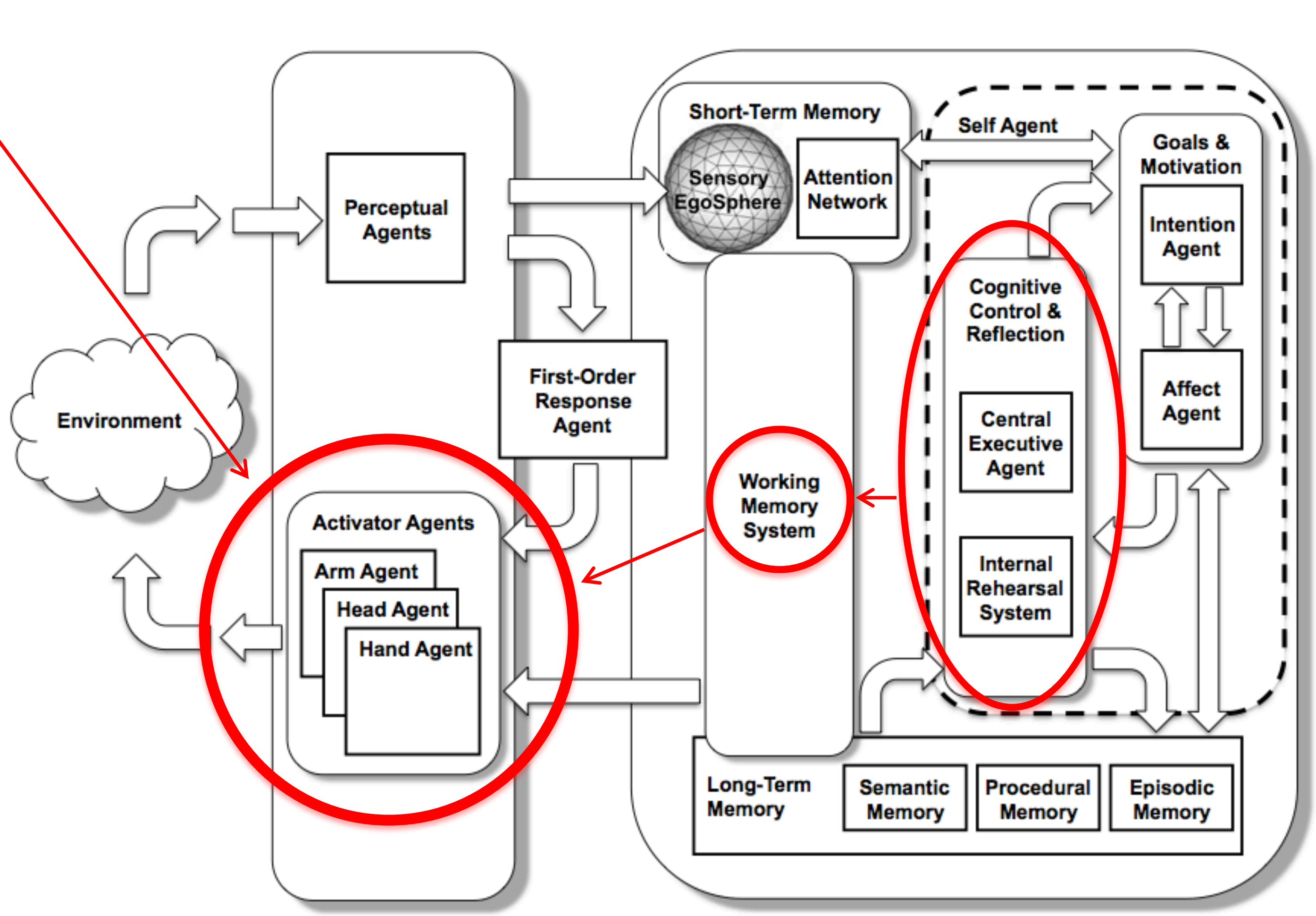
ISAC

Select a behaviour-percept pair,
based on the current percept in the SES, its relevance,
and the likelihood of successful execution as
determined by internal simulation



ISAC

This is then placed in working memory and the Activator Agent executes the action



Lecture Summary

1. There are in excess of 84 cognitive architectures
2. They can be classified as cognitivist, emergent, and hybrid
3. Cognitivist cognitive architectures are often referred to as symbolic cognitive architectures
4. The most famous include Soar, ACT-R, and CLARION
5. The ISAC cognitive architecture was designed for cognitive robotics and comprises most of the components one expects in a cognitive architecture
6. In ISAC, task execution is handled either by the First Response Agent or by the Central Executive Agent

Recommended Reading

I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, 2020

D. Vernon, C. von Hofsten, and L. Fadiga. "A Roadmap for Cognitive Development in Humanoid Robots", Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010; Chapter 5 and Appendix A.

Recommended Videos

J. Laird. Open Research and the Soar Cognitive Architecture.
https://www.youtube.com/watch?v=2pNsfBj7XSA&feature=youtu.be