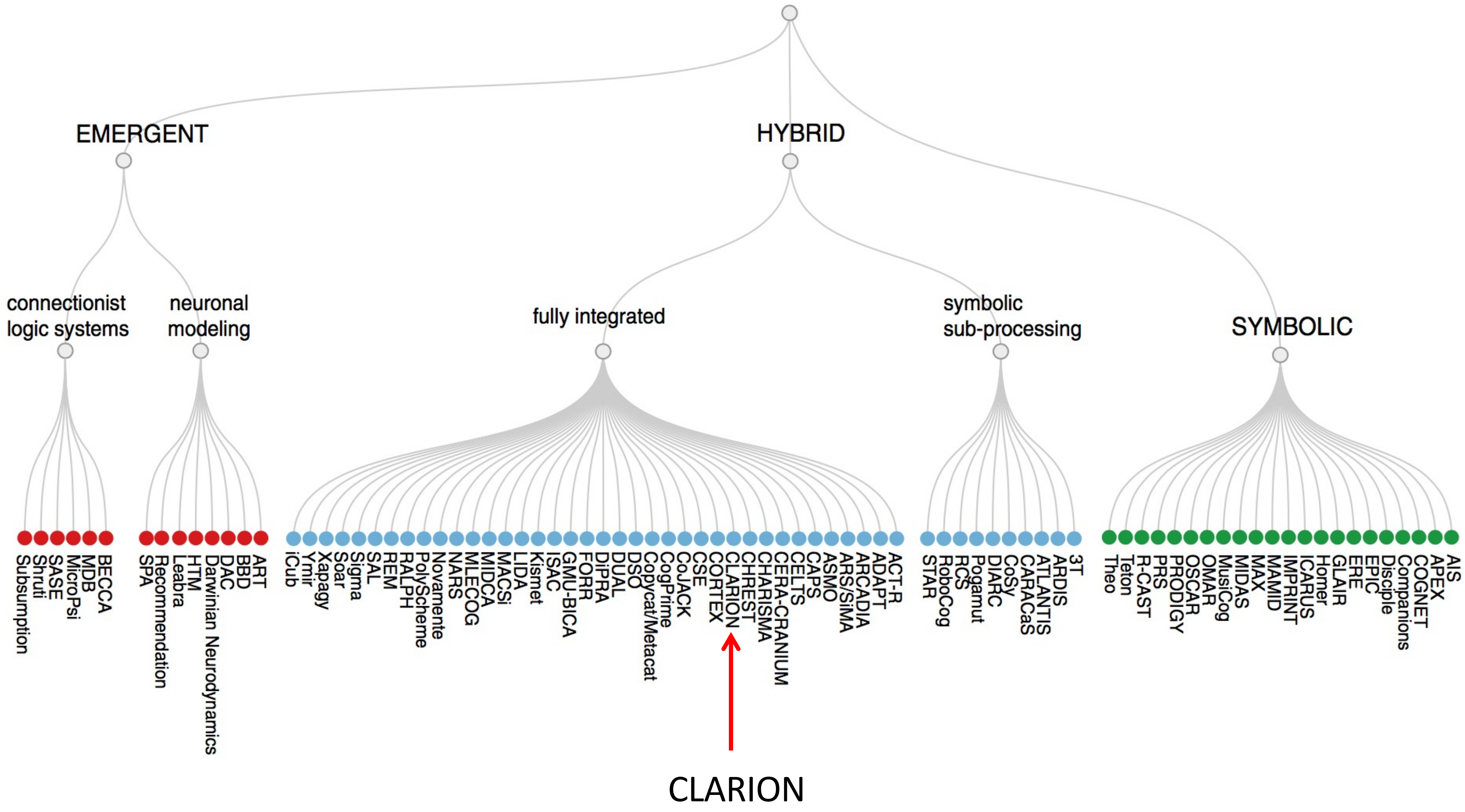
Artificial Cognitive Systems

Module 3: Cognitive Architectures

Lecture 3: Example cognitive architectures: Clarion, ISAC

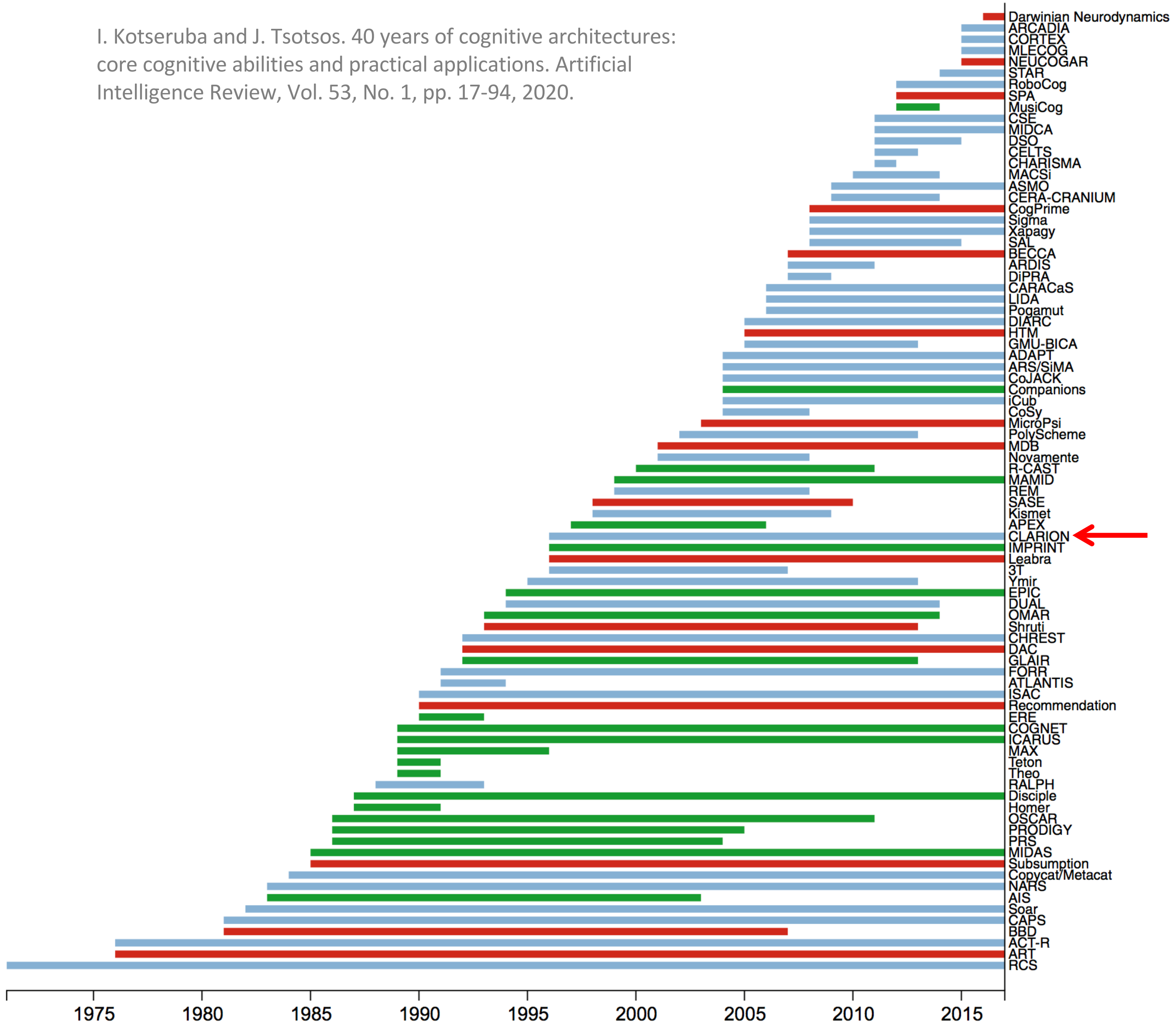
David Vernon
Carnegie Mellon University Africa

www.vernon.eu



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
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.



194 A Catalogue of Cognitive Architectures

A.3.6 *The CLARION Cognitive Architecture*

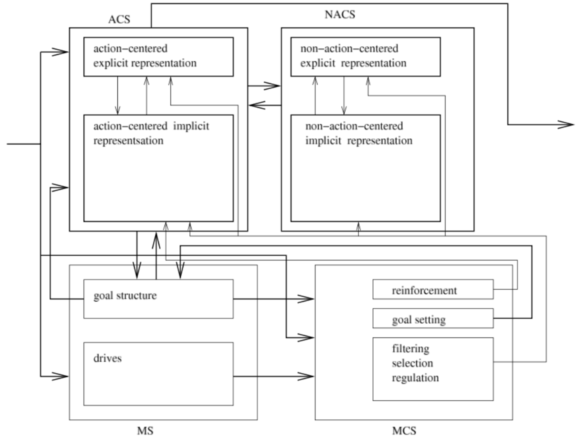


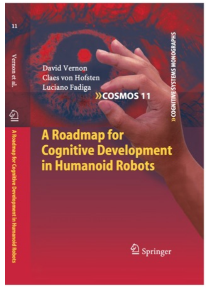
Fig. A.10 The CLARION hybrid cognitive architecture (from [364]). ACS stand for the action-centered subsystem, NACS for the non-action-centred subsystem, MS for the motivational subsystem, and MCS for the meta-cognitive subsystem. All four subsystems have two types of representation: implicit (connectionist)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

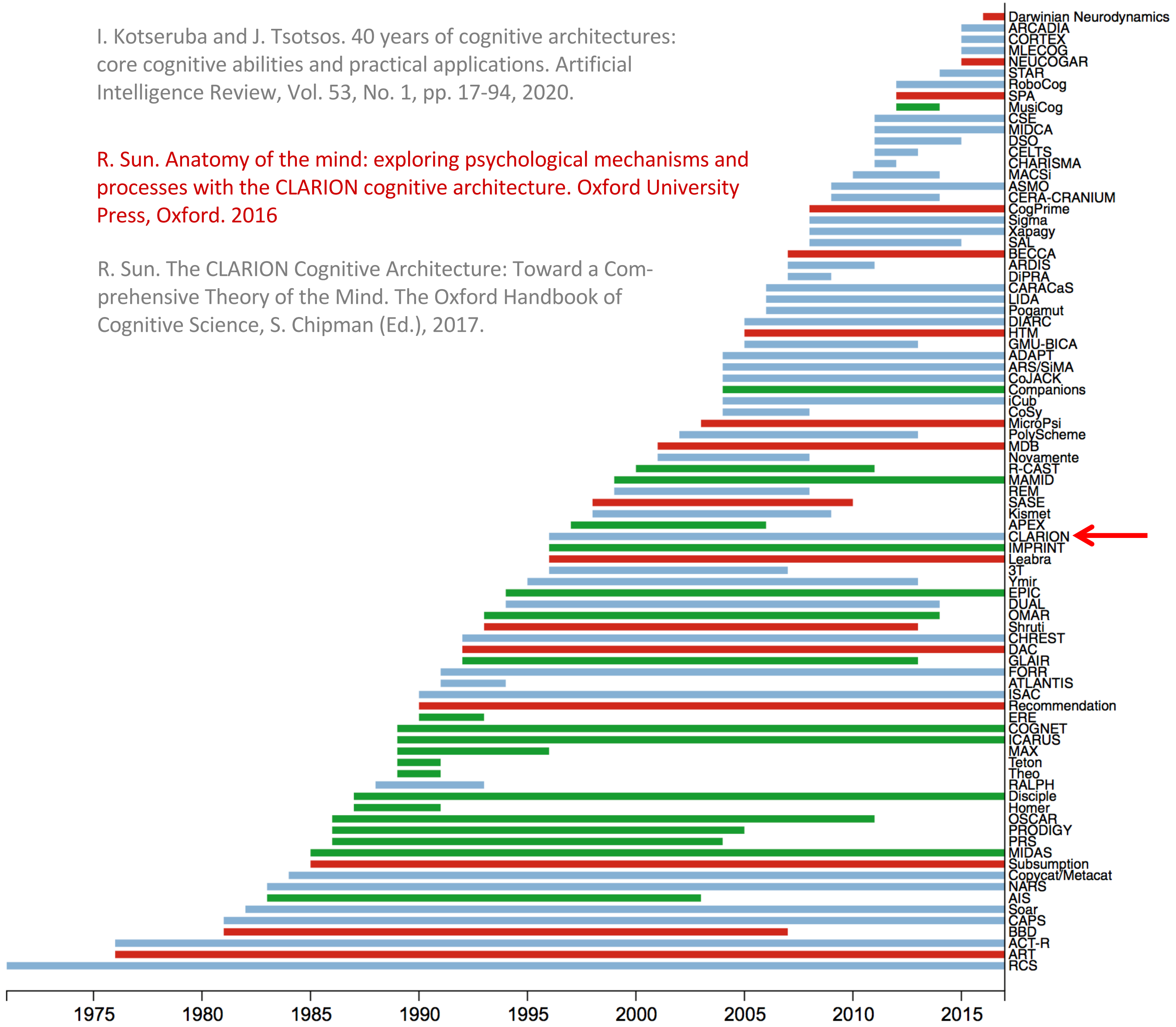


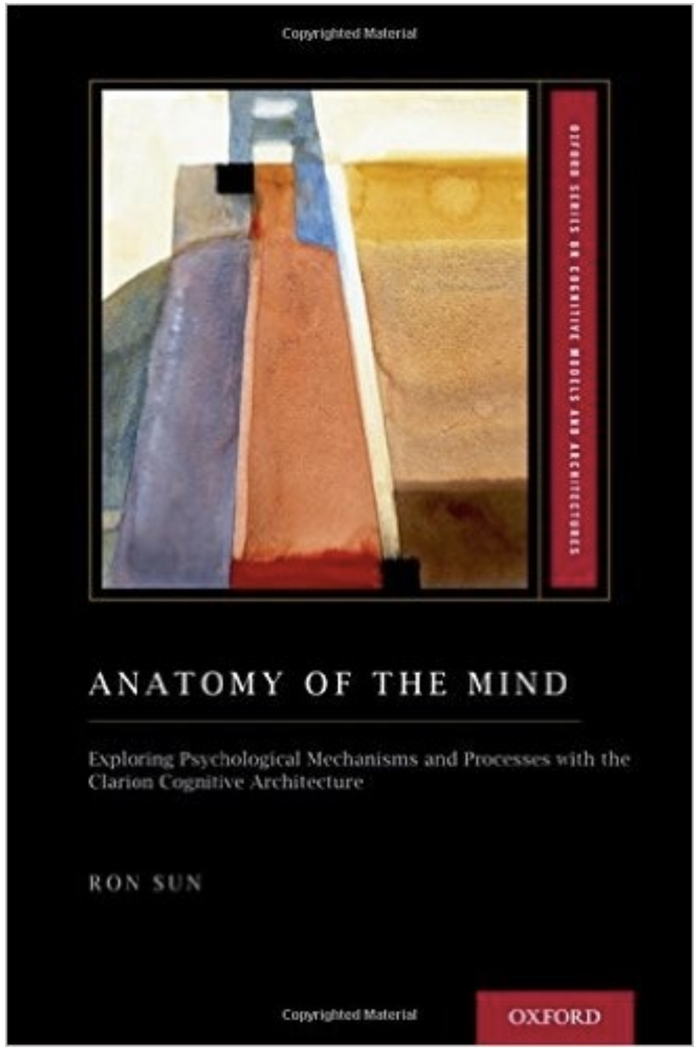
D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.

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.

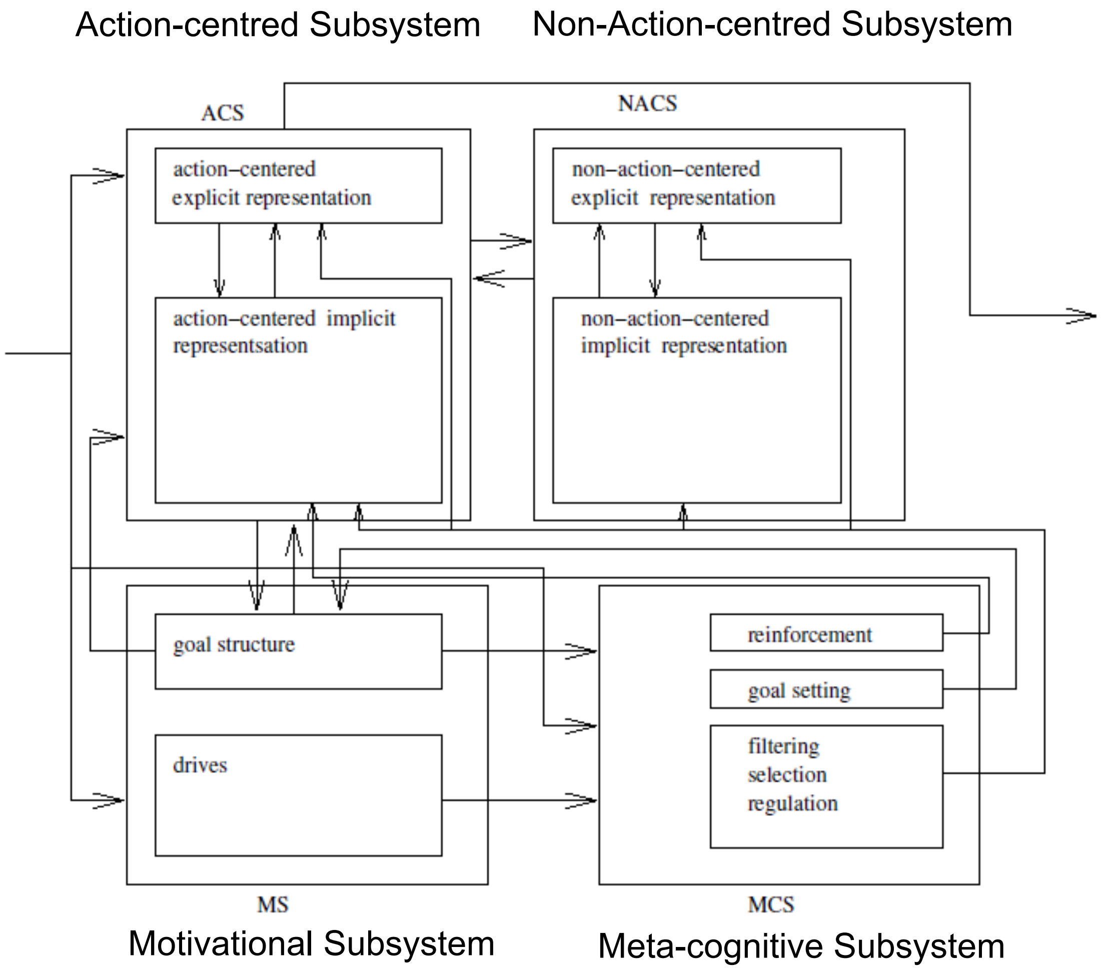
R. Sun. Anatomy of the mind: exploring psychological mechanisms and processes with the CLARION cognitive architecture. Oxford University Press, Oxford. 2016

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





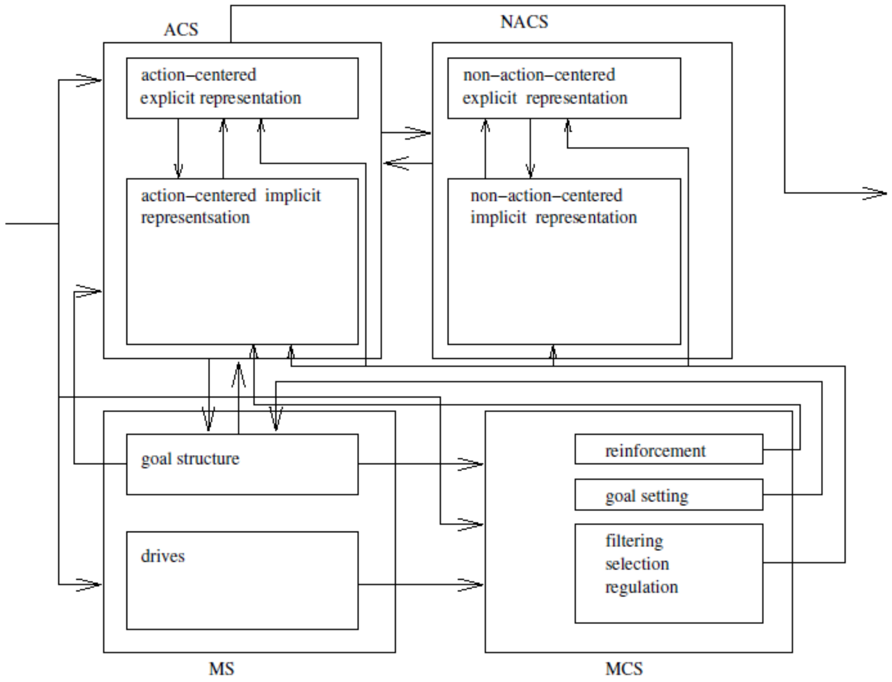
CLARION



Action-centred Subsystem
Non-Action-centred Subsystem
ACS
NACS
action–centered
explicit representation
non–action–centered
explicit representation
action–centered implicit
representsation
non–action–centered
implicit representation
goal structure
drives
reinforcement
goal setting
filtering
selection
regulation
MS
MCS
Motivational Subsystem
Meta-cognitive Subsystem

CLARION

- Hybrid cognitive architecture
  - Symbolic representations
  - Connectionist representations
- Four sub-systems
  - ACS – Action-centred subsystem
  - NACS – Non-action-centred subsystem
  - MS – Motivational subsystem
  - MCS – meta-cognitive subsystem

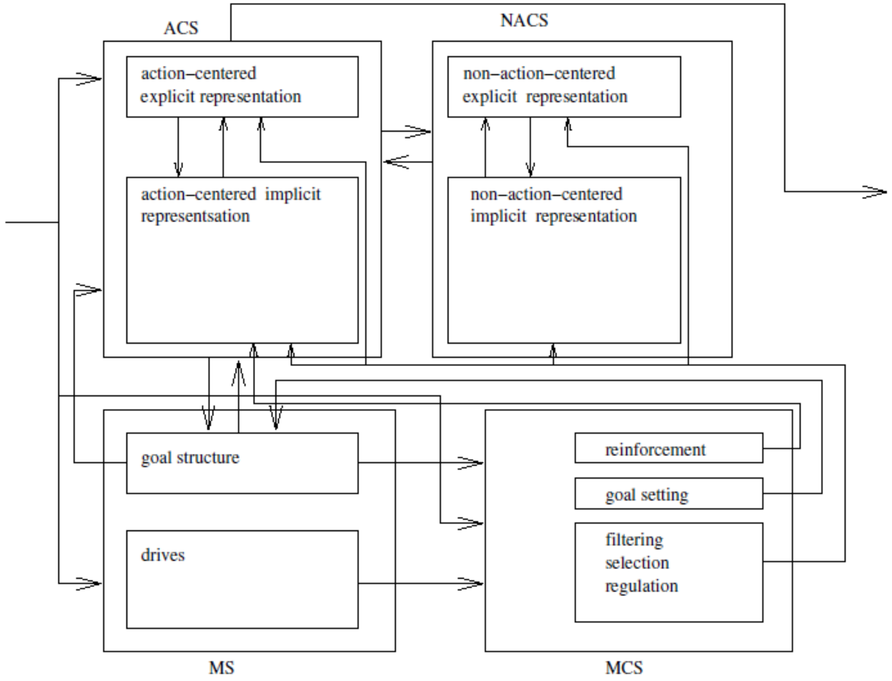


CLARION

- 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

- Able to learn with or without a priori domain-specific knowledge

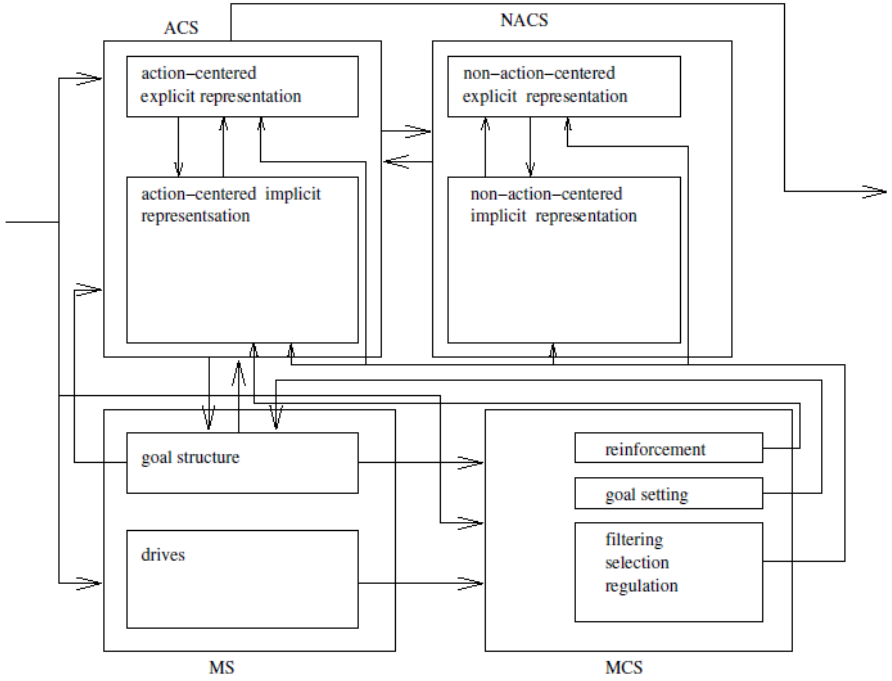
- Able to learn continuously from on-going experience



CLARION

Action-centred Subsystem (ACS)

- Controls actions
  - External physical movements
  - Internal mental operations



CLARION

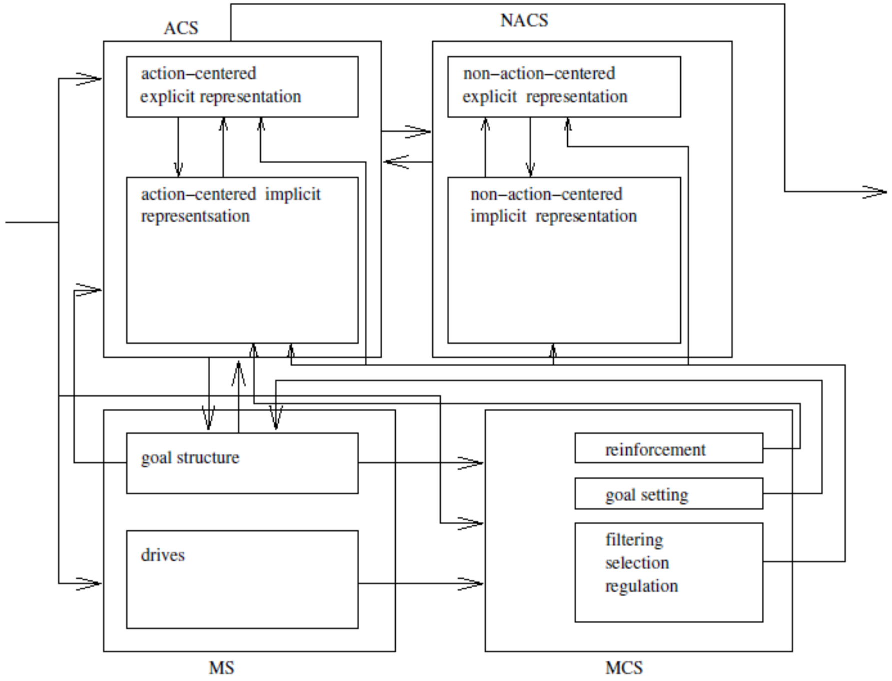
Action-centred Subsystem (ACS)

- Given some observational state, i.e. a set of sensory features x
  - The bottom level evaluates the desirability (“quality”) of all possible actions

    $Q(x, a_1), Q(x, a_2), \ldots , Q(x, a_n)$

  - The top level identifies possible actions from a rule network
    based on the input x sent up from the bottom level

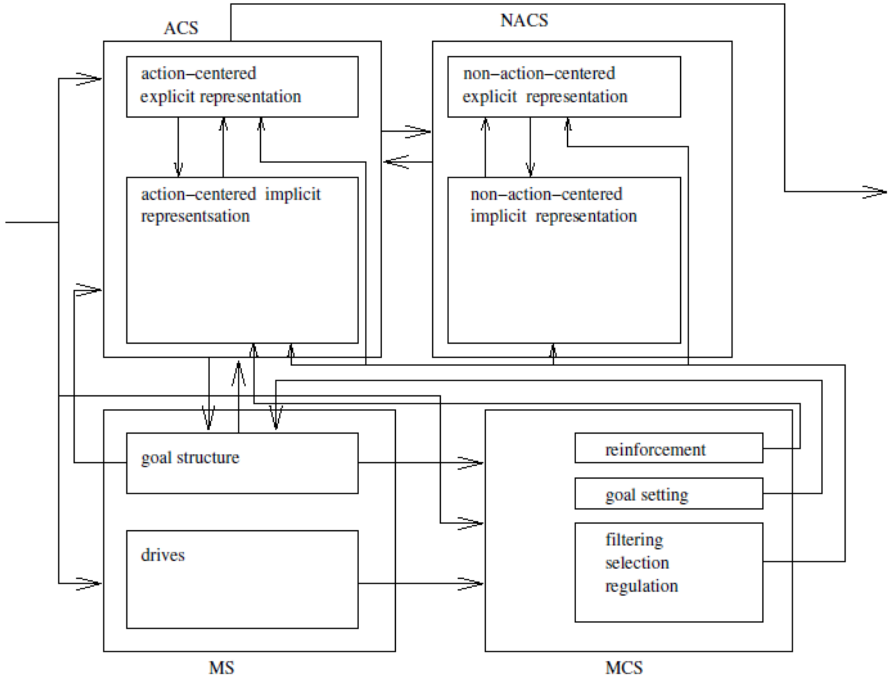
    $(b_1, b_2, \ldots , b_m)$



CLARION

Action-centred Subsystem (ACS)

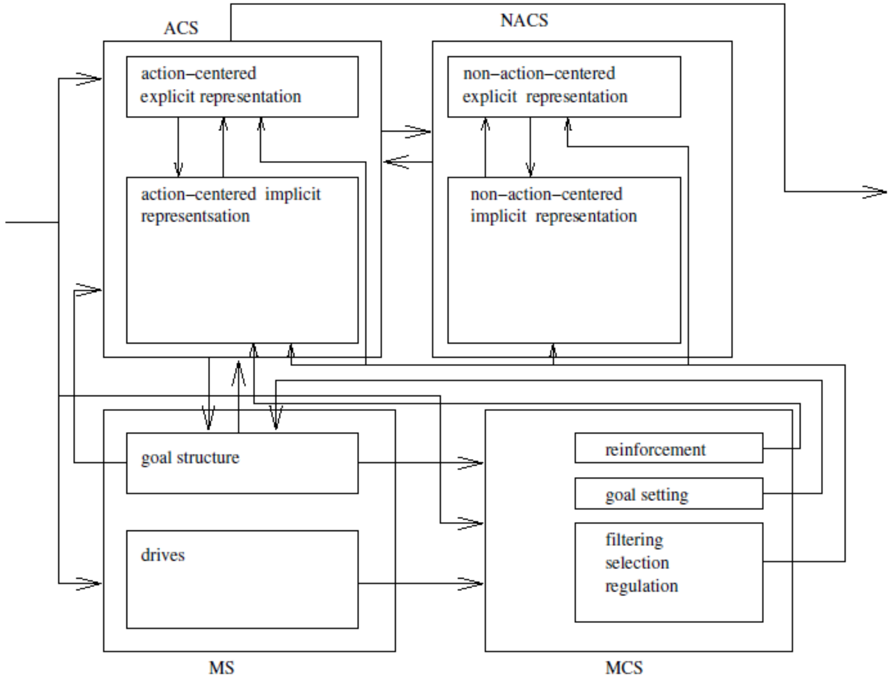
- The bottom-level actions a_i and top-level actions b_j are compared and the most appropriate top-level action b is selected
- Action b is performed and the outcome is observed
  - The next state y and (possibly) a reinforcement r are determined
  - The Q values at the bottom level are updated using the Q-Learning-Backpropagation algorithm
  - The top-level rules are also updated using the Rule-Extraction-Refinement algorithm
- This process continues indefinitely



CLARION

Action-centred Subsystem (ACS)

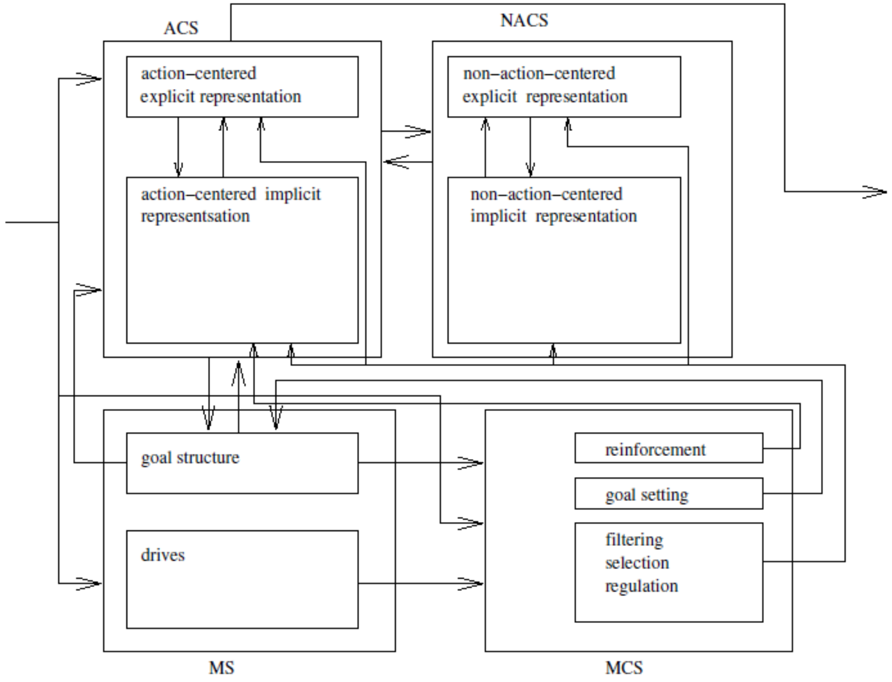
- The bottom level comprises several modules of small neural networks
  - Each adapted to a distinct sensory modality or task
  - These modules can be developed by the system
    - based on experience (i.e. though **ontogenesis**) through trial-and-error exploration
    - or they can be specified a priori and hard-wired into the cognitive architecture (i.e. as the system **phylogeny**)



CLARION

Action-centred Subsystem (ACS)

- In the top level, explicit symbolic conceptual knowledge is captured in the form of symbolic rules
- Explicit knowledge can be learned in several ways
  - Independent experiential hypothesis-testing learning
  - Mediation of implicit knowledge: bottom-up learning … Autonomous Generation of Explicit Conceptual Structures

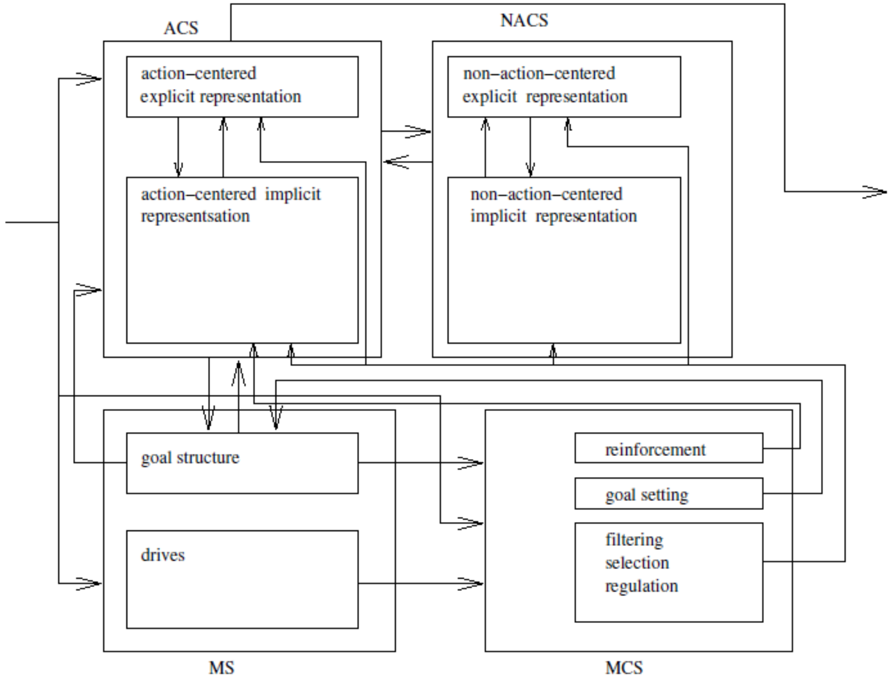


CLARION

Action-centred Subsystem (ACS)

- The implicit bottom level & the explicit top level representations
  interact to effect bottom-up learning

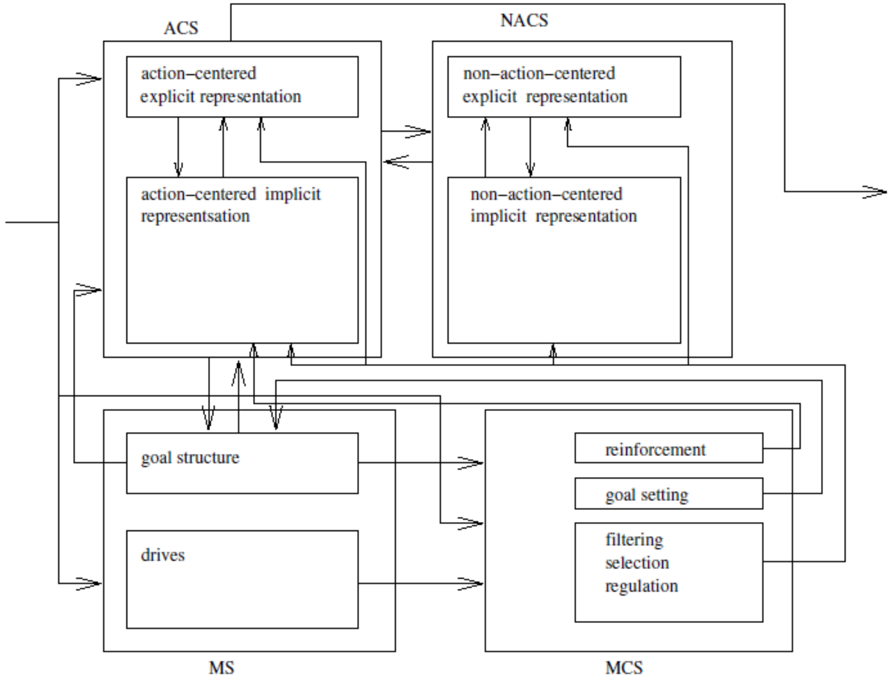
- If an action selected by the bottom level is successful
  - the system extracts an explicit rule that corresponds to the sensory features and the selected action
  - adds the rule to its top level rule network



CLARION

Action-centred Subsystem (ACS)

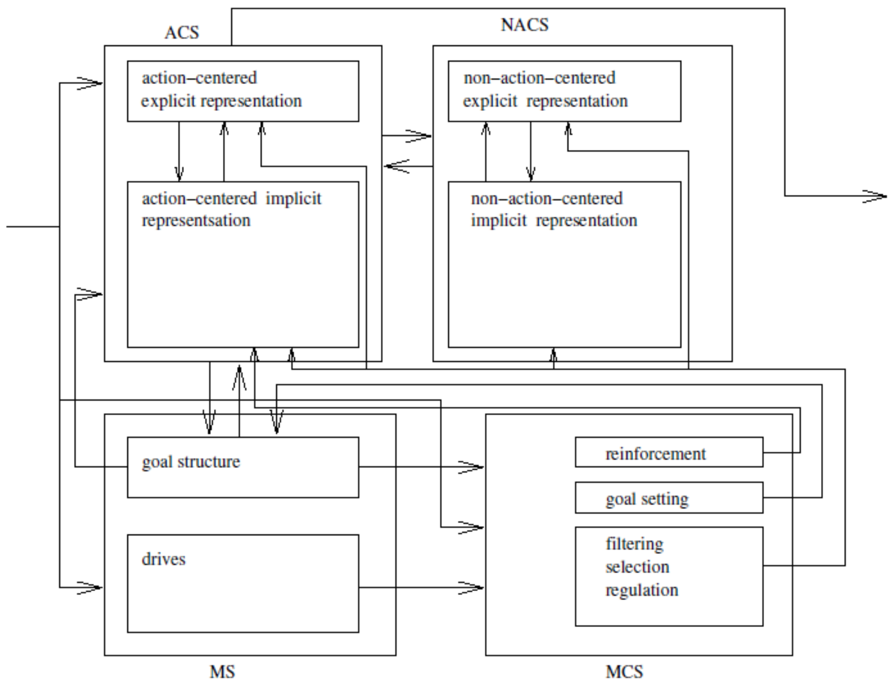
- The system subsequently verifies the extracted rule by considering the outcome of applying the rule
  - If the outcome is successful, the rule is generalized (made more universal and applicable to other situations)
  - If the outcome is unsuccessful, the rule is refined (made more specific and exclusive of the current situation)
- i.e. autonomous generation of explicit conceptual structures by exploiting implicit knowledge acquired by trial-and-error learning



CLARION

Action-centred Subsystem (ACS)

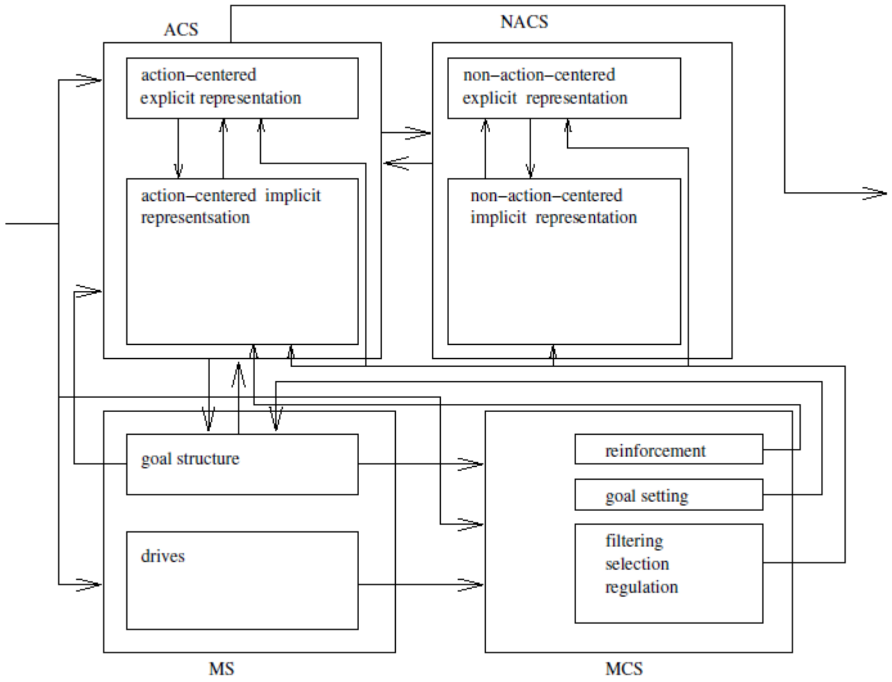
- Assimilation of externally-given conceptual structures
  - **Internalizing** externally-provided knowledge in the form of explicit rule-based conceptual structures with existing conceptual structures at the top-level
  - **Assimilating** these into the bottom level implicit representation … top-down learning



CLARION

Non-Action-centred Subsystem (NACS)

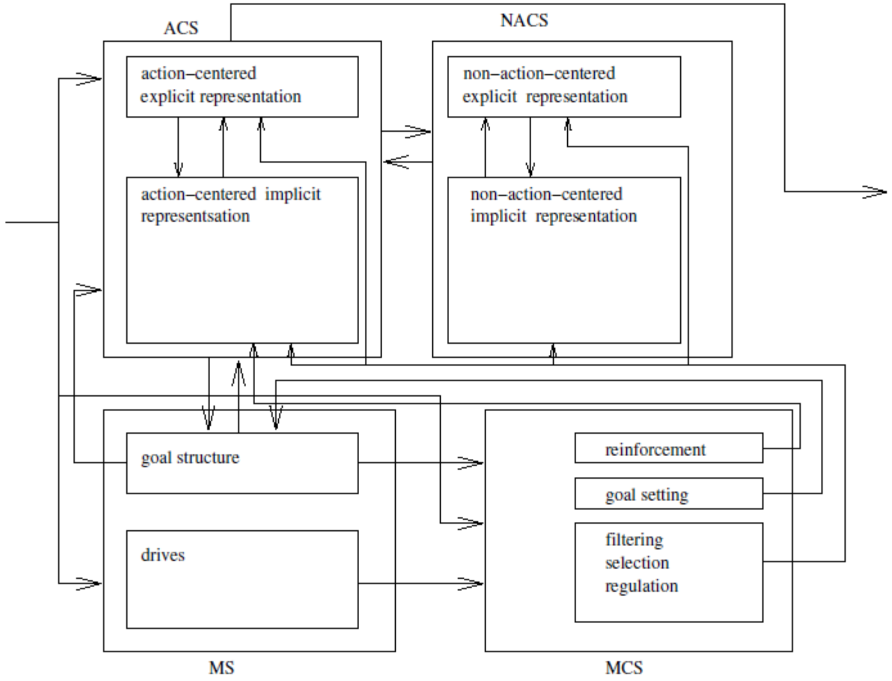
- Maintains the system's general knowledge
  - Implicit knowledge in connectionist form
    - Associative memory networks (mapping input to output)
  - Explicit knowledge in symbolic form
    - A network of nodes
    - Each node corresponds to an entity-specific chunk comprising
      - » an entity identifier (e.g. table_1)
      - » a vector of feature dimensions / feature value pairs (e.g. (size, large) … (colour, white), (number_of_legs, 4))



CLARION

Non-Action-centred Subsystem (NACS)

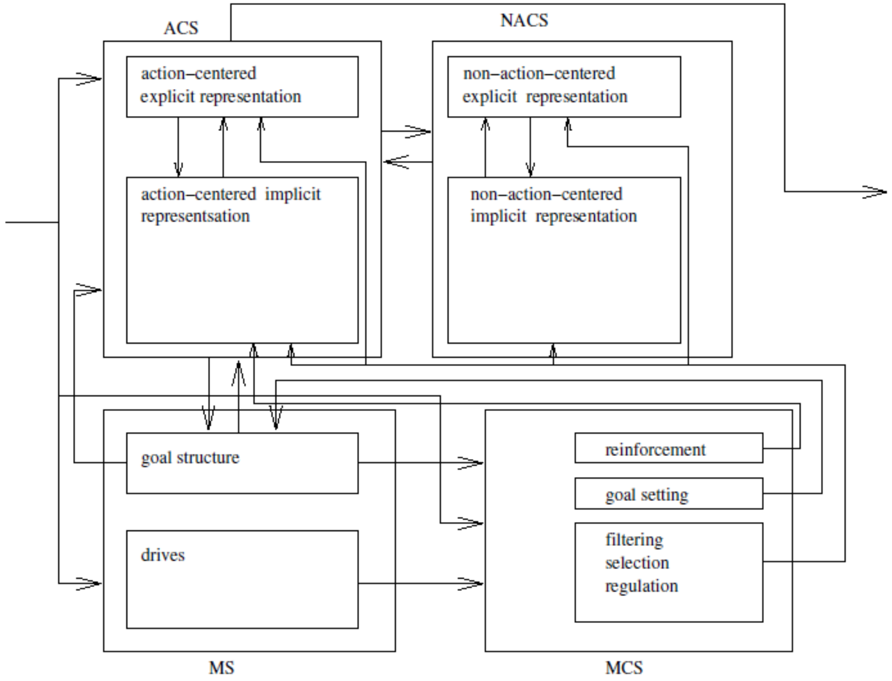
- Maintains the system's general knowledge
  - The feature values are represented by nodes in the bottom level associative memory
  - Chunks are linked through association rules
- Both bottom-up and top-down learning can take place
  - Extract explicit knowledge in the top level from the implicit knowledge in the bottom level
  - Assimilate explicit knowledge of the top level into implicit knowledge in the bottom level



CLARION

Motivational Subsystems (MS)

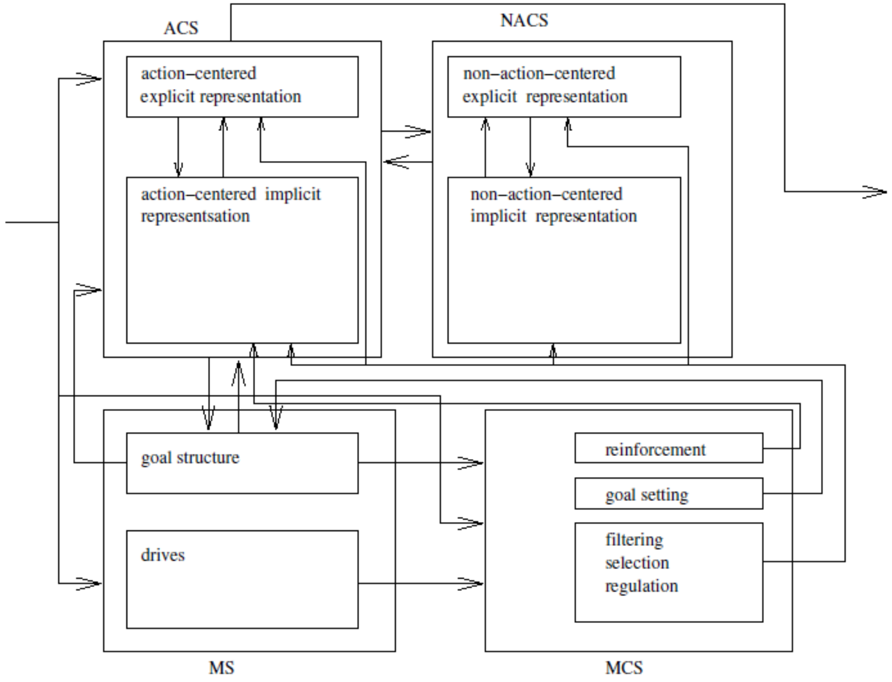
- Provides
  - The drives that determines what the agent does
  - Evaluates the feedback
    (were the outcomes of an action satisfactory or not)



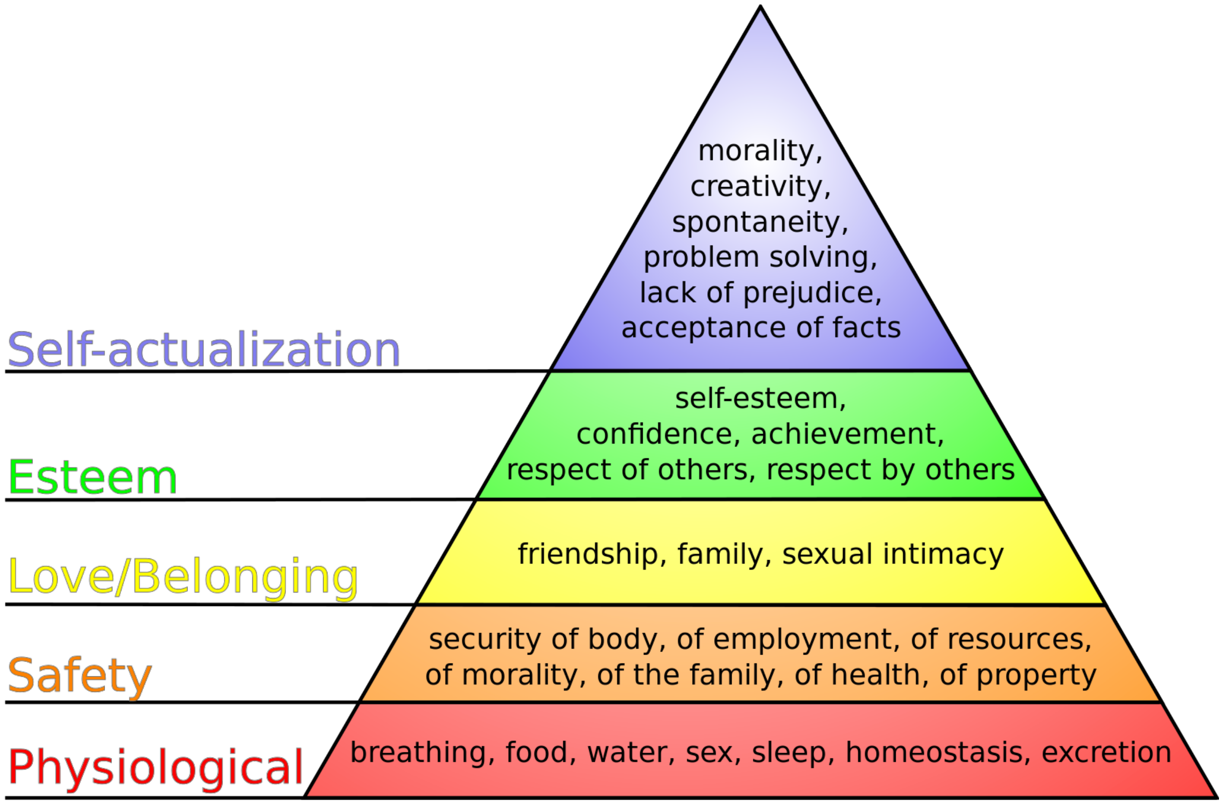
CLARION

Motivational Subsystems (MS)

- Provides the ACS with goals derived from
  - Low-level drives concerning physiological needs (e.g. need for food, need for water, need to avoid danger, need to avoid boredom, …)
  - High-level drives (e.g., desire for social approval, desire for following social norms, desire for reciprocation, desire for imitation of other people, …)
    - Primary hard-wired drives (cf. Maslow's hierarchy of needs)
    - Secondary derived drives (changeable, acquired mostly in the process of satisfying primary drives)



Maslow's Hierarchy of Needs

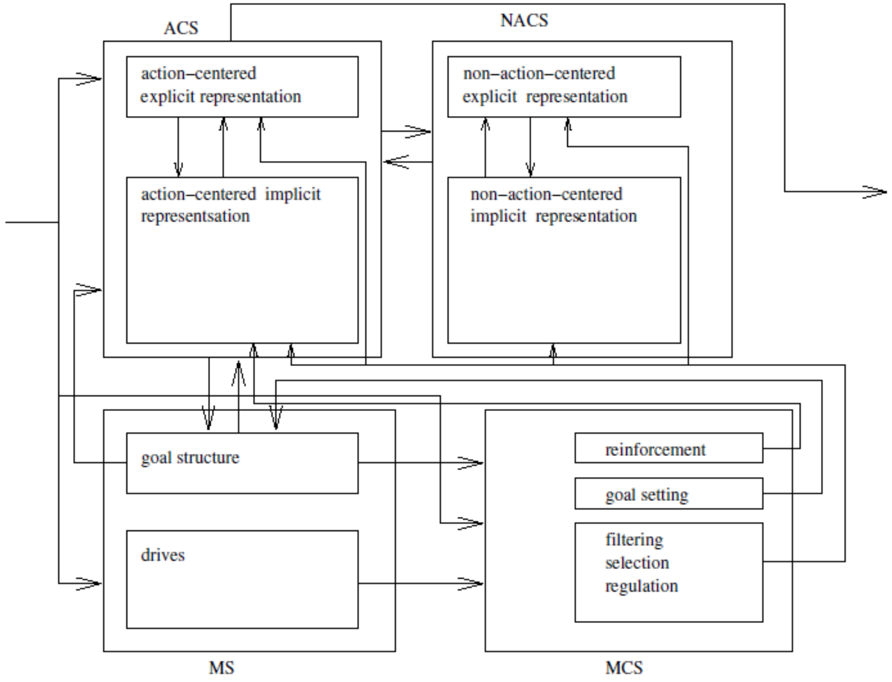


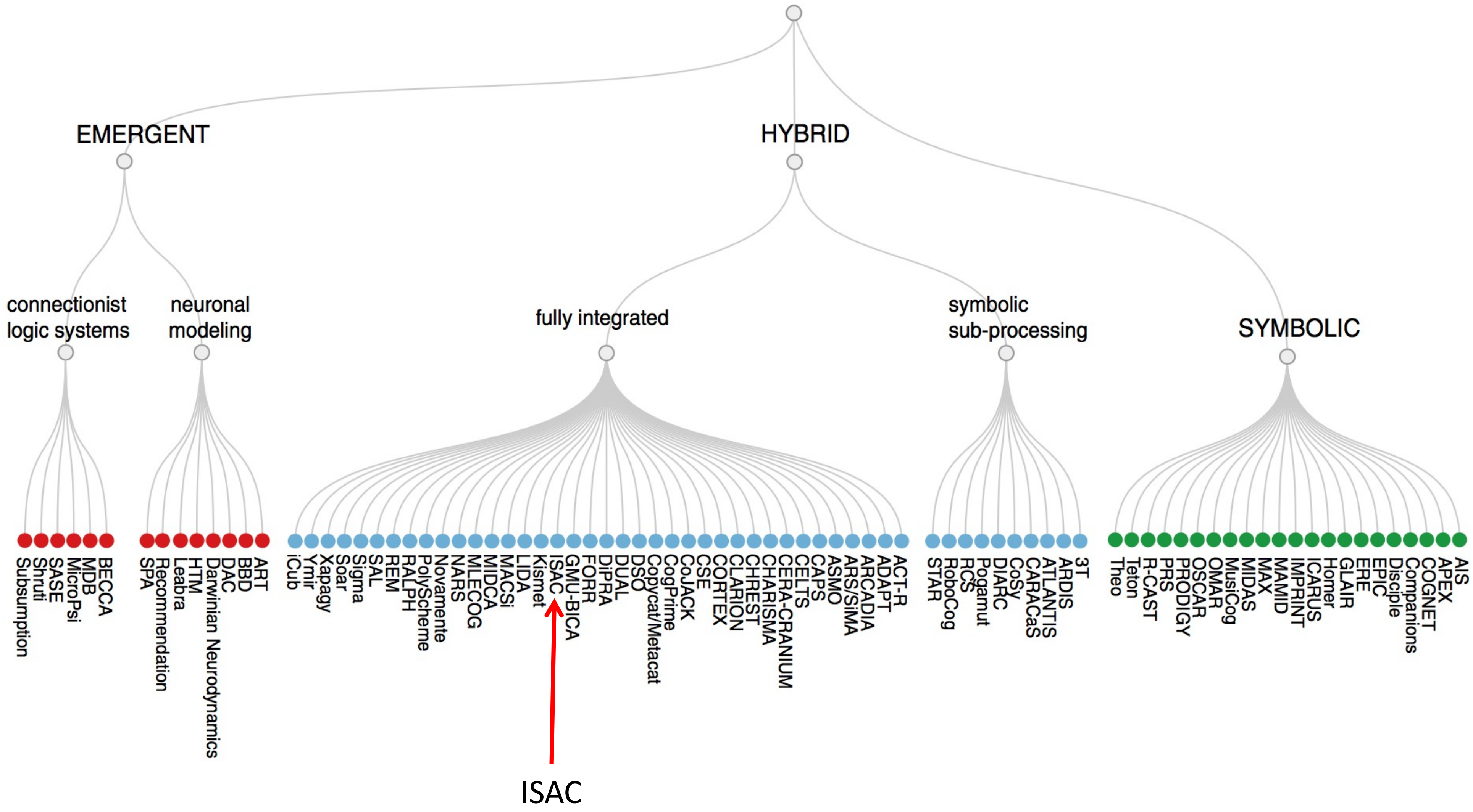
https://commons.wikimedia.org/wiki/File:Maslow%27s_hierarchy_of_needs.svg

CLARION

Meta-cognitive Subsystem (MCS)

- Monitors, regulates, and modify the overall behaviour of the cognitive system to improve cognitive performance
  - By setting goals for the action-centred subsystem
  - By setting essential parameter values the action-centred and non-action-centred subsystems
  - For example, setting reinforcement functions
  - Can be achieved by setting drive states in the motivational subsystem
- Also comprises a top level (explicit) and bottom level (implicit)

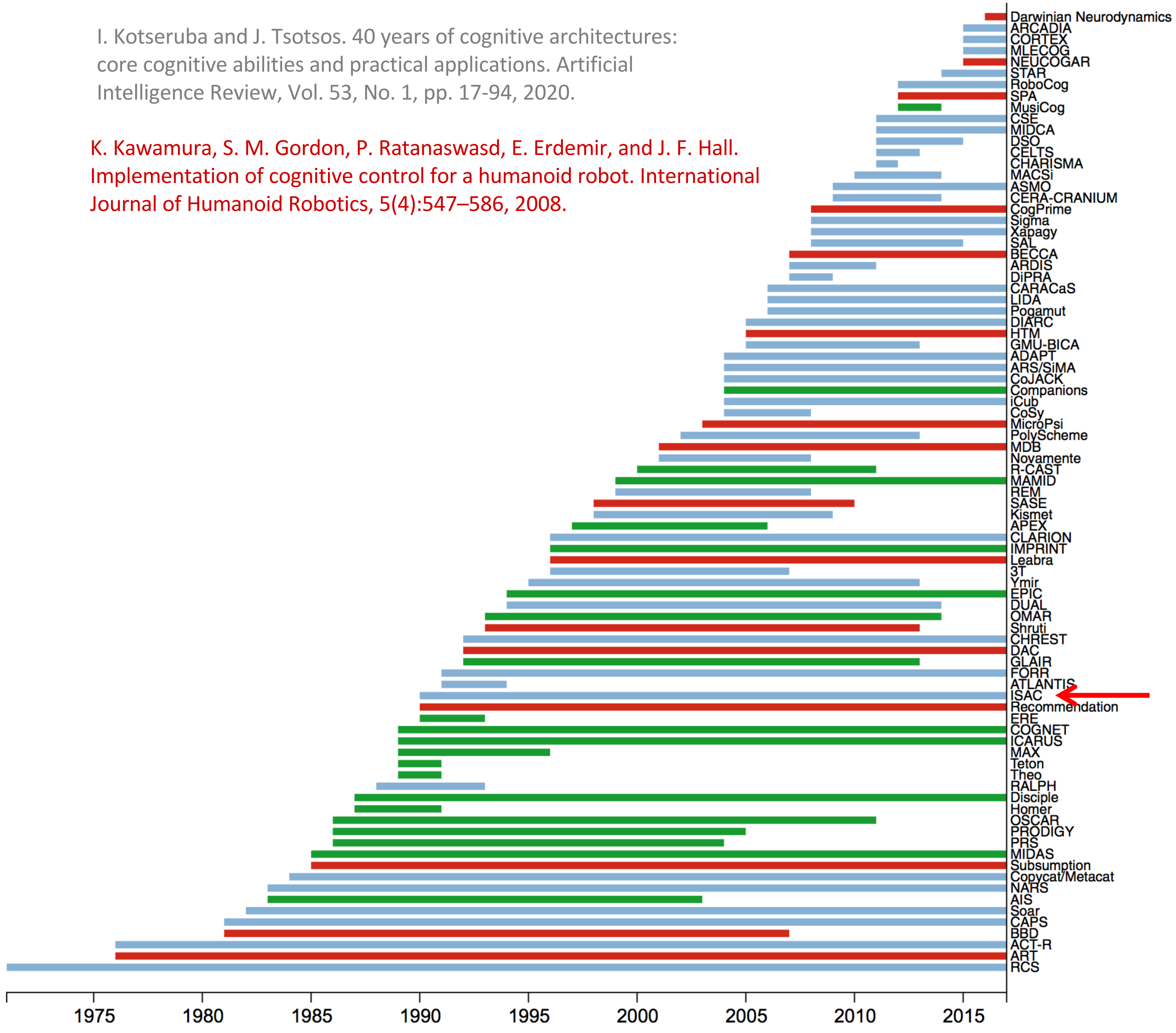




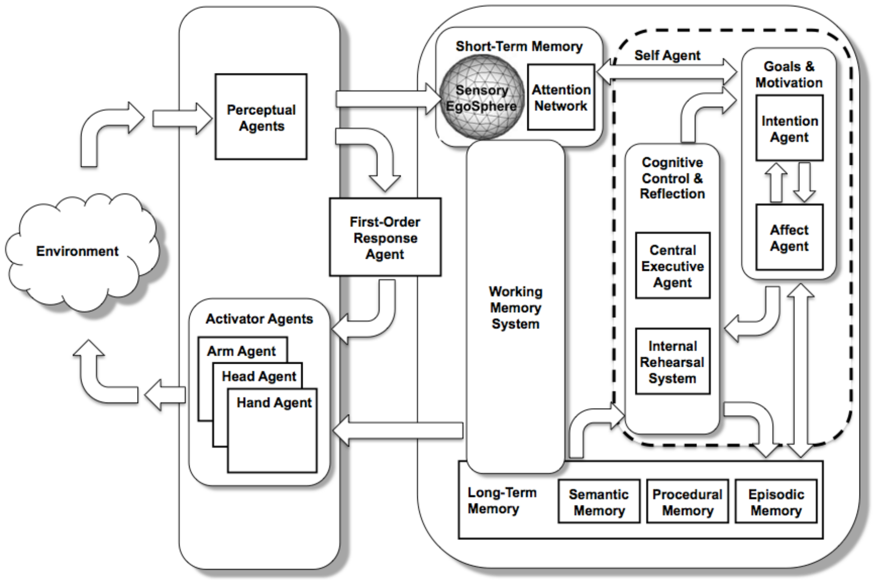
We will now study one of these cognitive architectures in a little more detail

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.

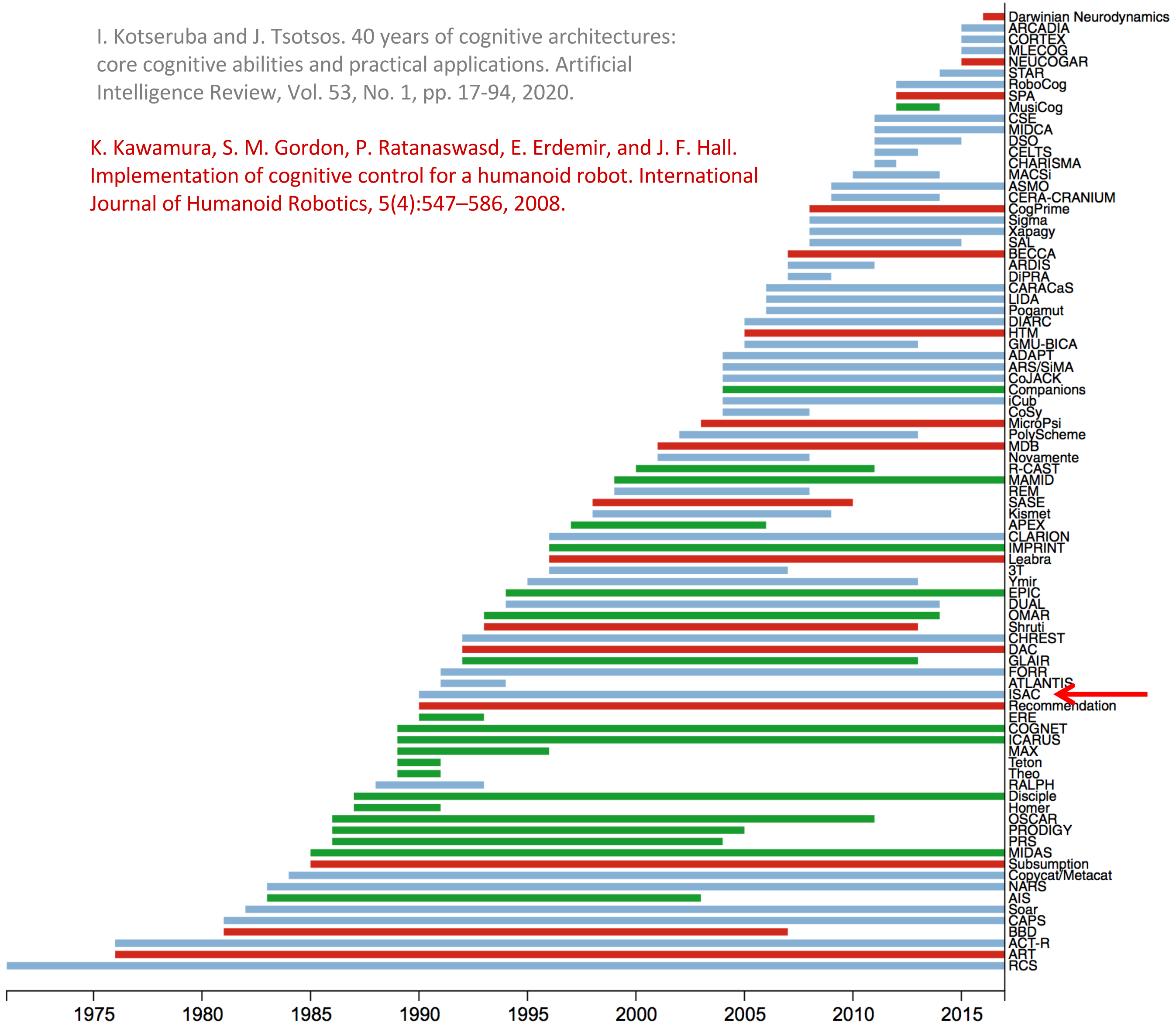


ISAC

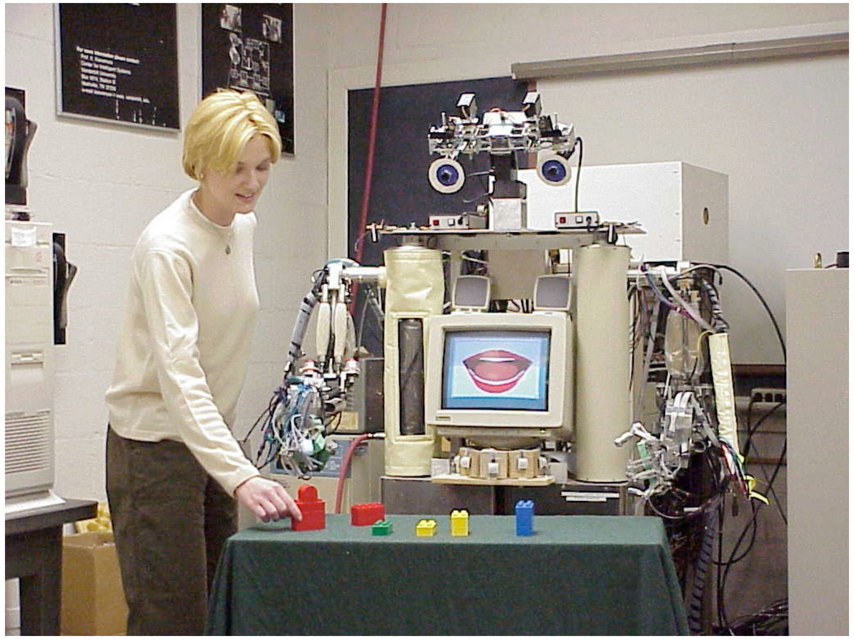


I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

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.



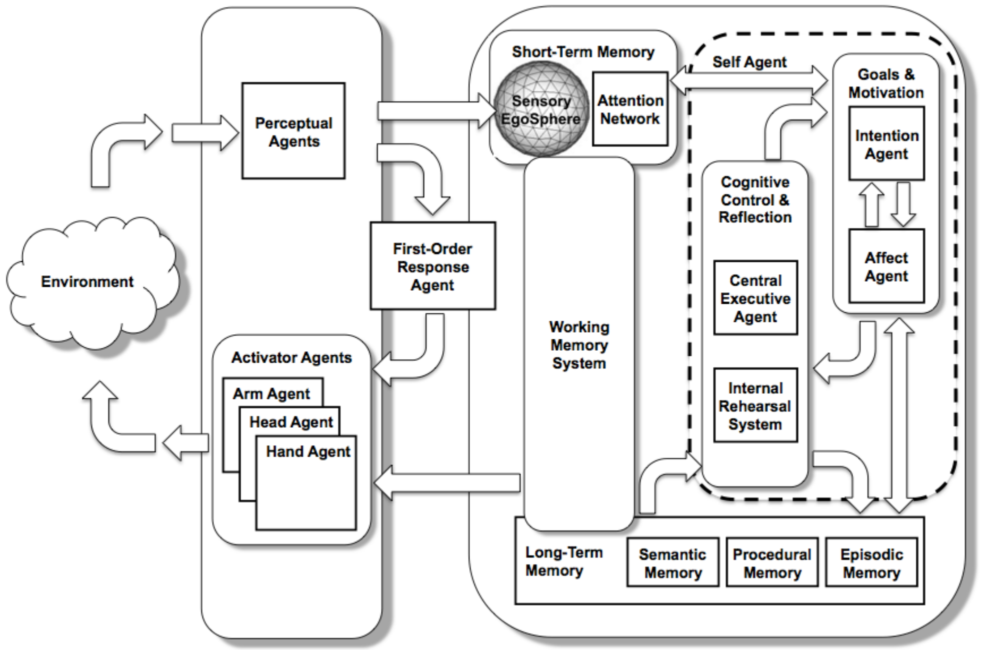
ISAC



ISAC

ISAC – Intelligent Soft Arm Control

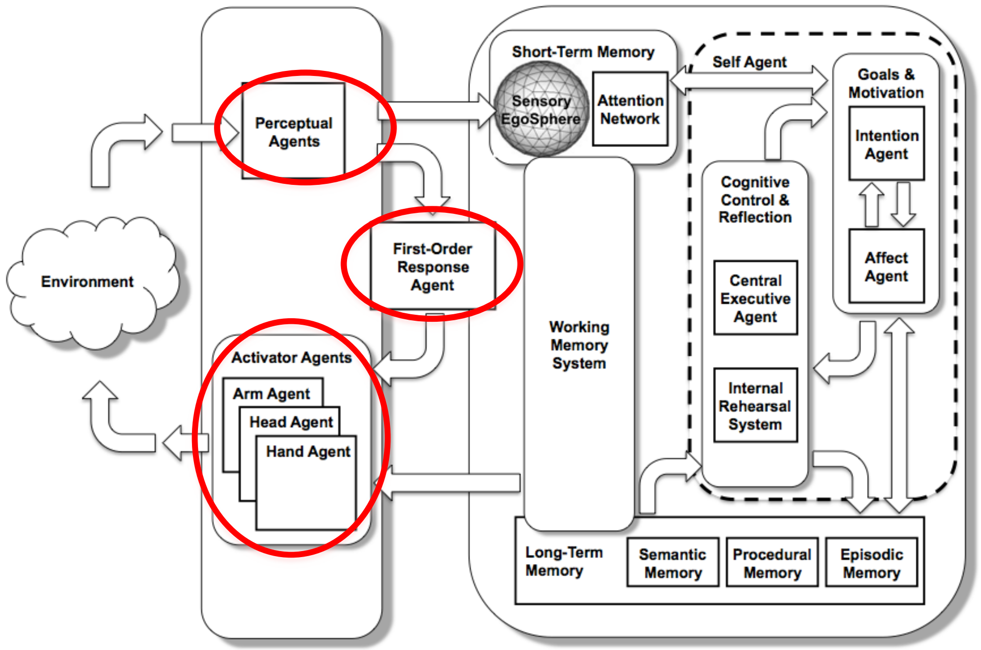
- Hybrid cognitive architecture for an upper torso humanoid robot (also called ISAC)
- Comprises an integrated collection of software agents and associated memories
- Agents operate asynchonously and communicate with each other by message passing



ISAC

Comprises activator agents

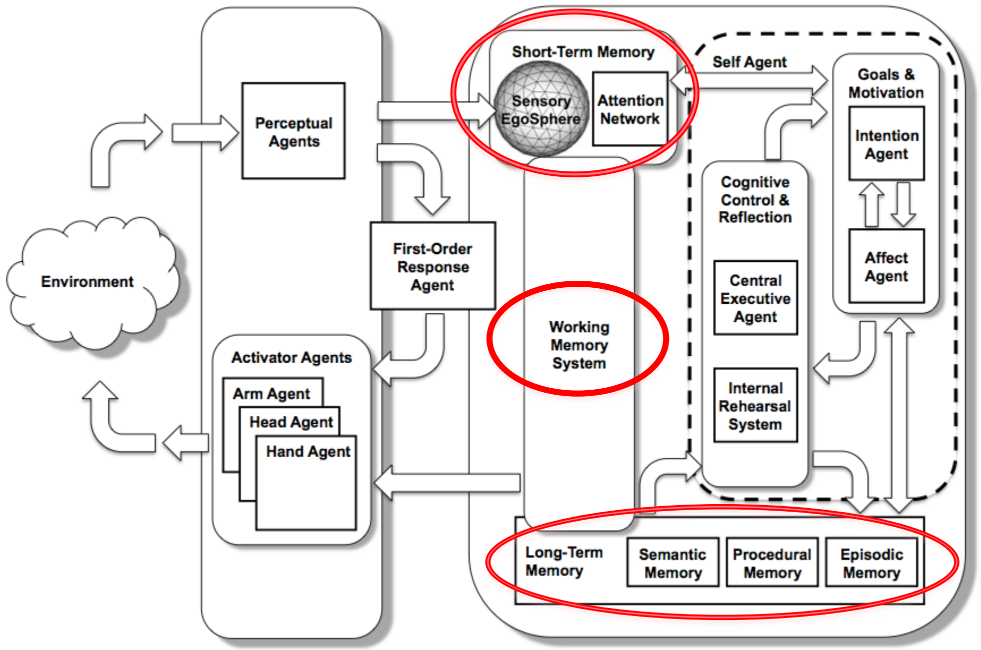
- Activator agents for motion control
- Perceptual agents
- First-order Response Agent (FRA)
  to effect reactive perception-action control



ISAC

Three memory systems

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



ISAC

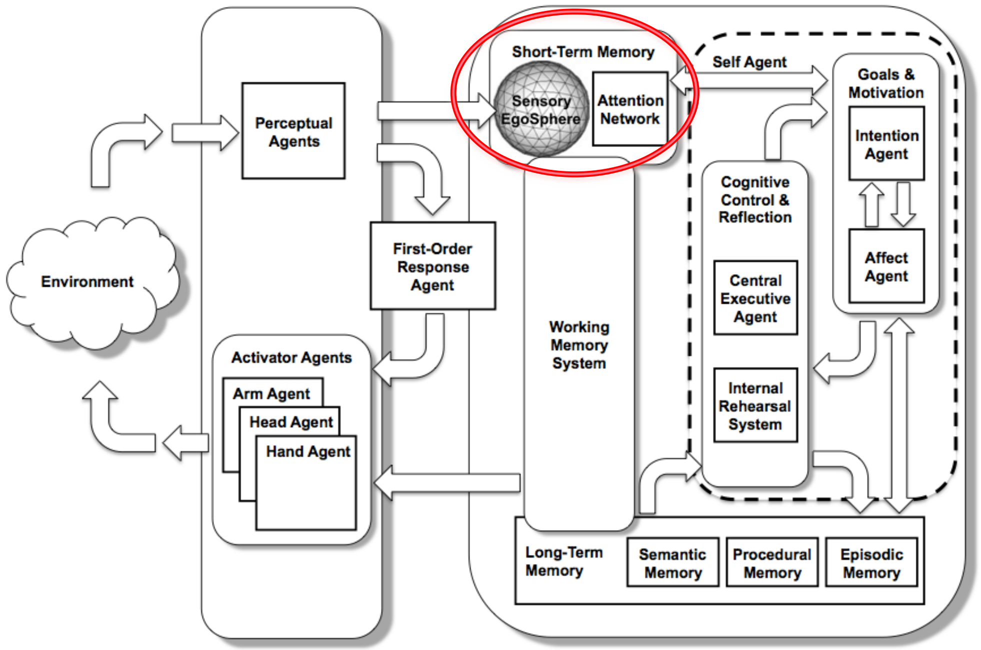
Short-term Memory

- Robot-centred spatio-temporal memory of the current perceptual events

- This is called a Sensory EgoSphere (SES)

    - Discrete representation of what is happening around the robot
    - Represented by a geodesic sphere indexed by two angles

- STM also has an attentional network
    - Determines the perceptual events that are most relevant



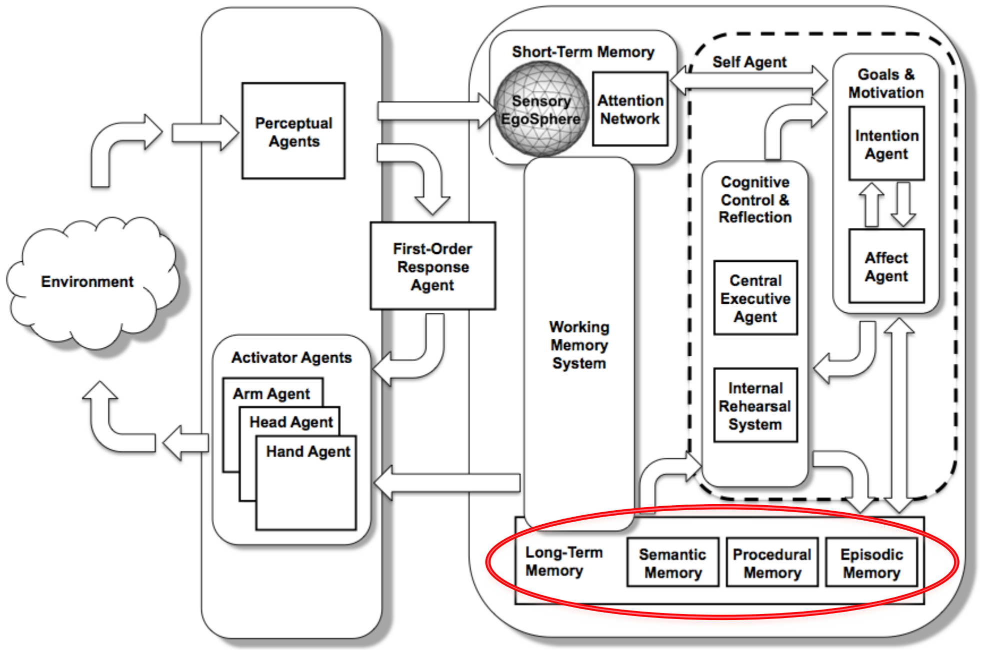
ISAC

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

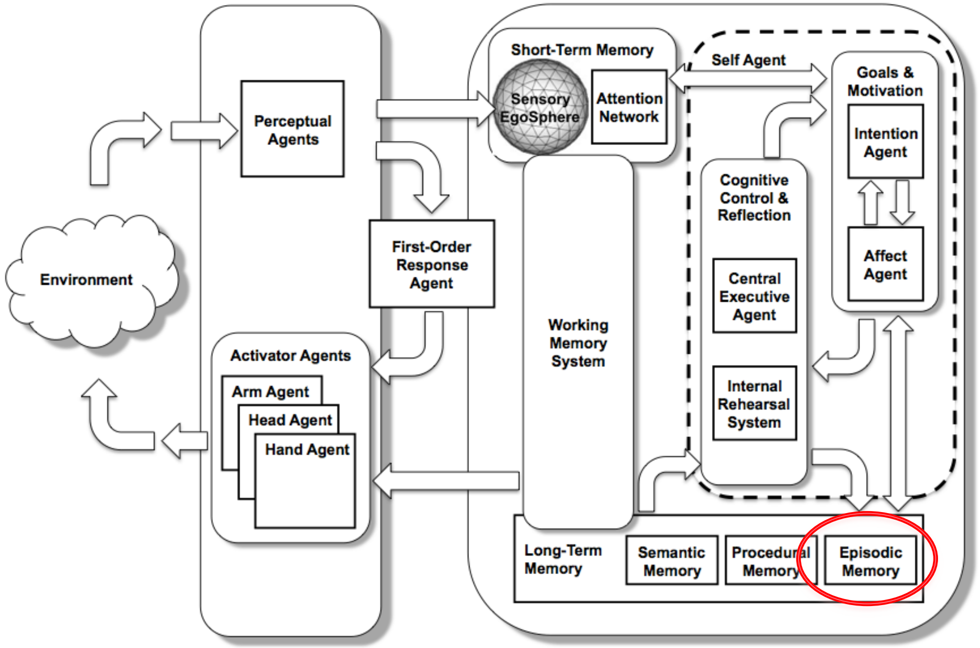
Abstracts past experiences & creates links or associati
between them

i.e. task-relevant percepts from the SES

- External situation
- Goals
- Emotions
- Actions
- Outcomes that arise from actions
- Valuations of these outcomes

i.e. internal evaluation of the perceived situation

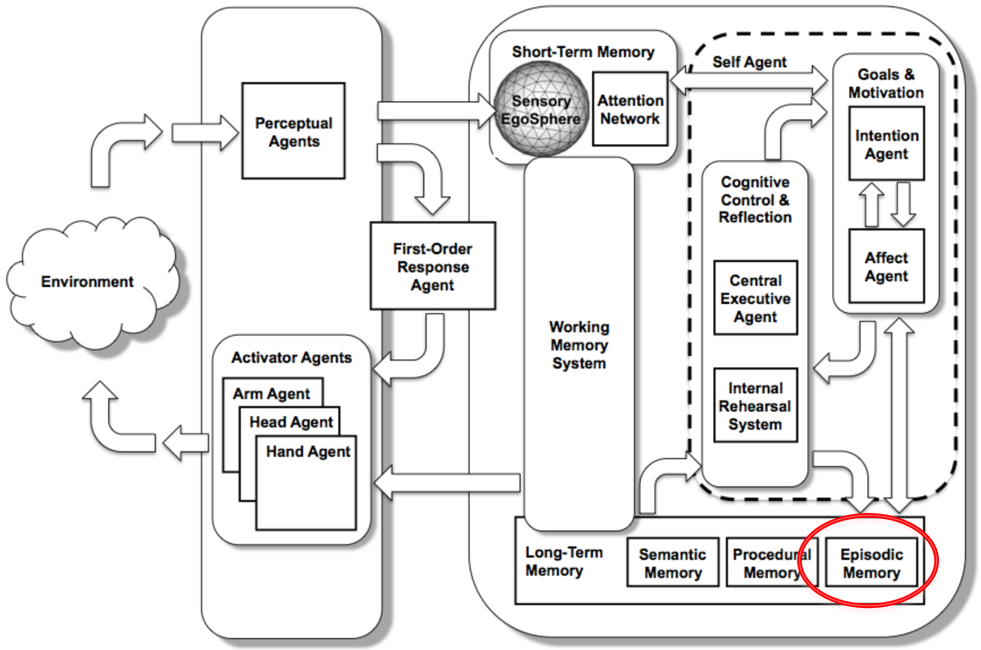
e.g. how close they are to the desired goal state and any reward received at a result



ISAC

Episodic memory

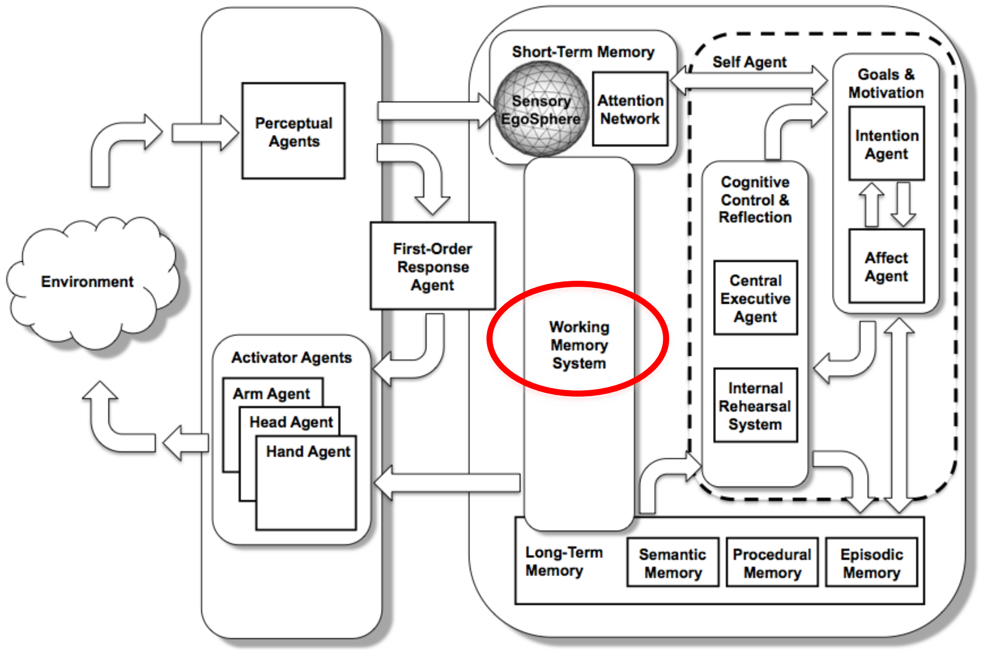
- Episodes are **connected by links** that encapsulate behaviours as transitions from one episode to another
- Multi-layered



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)



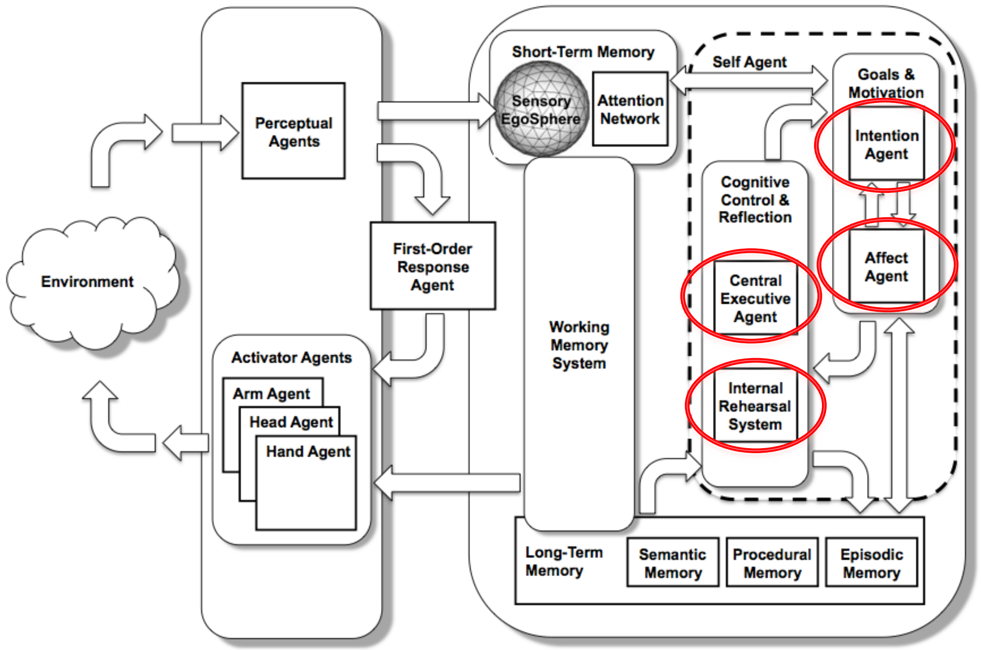
ISAC

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

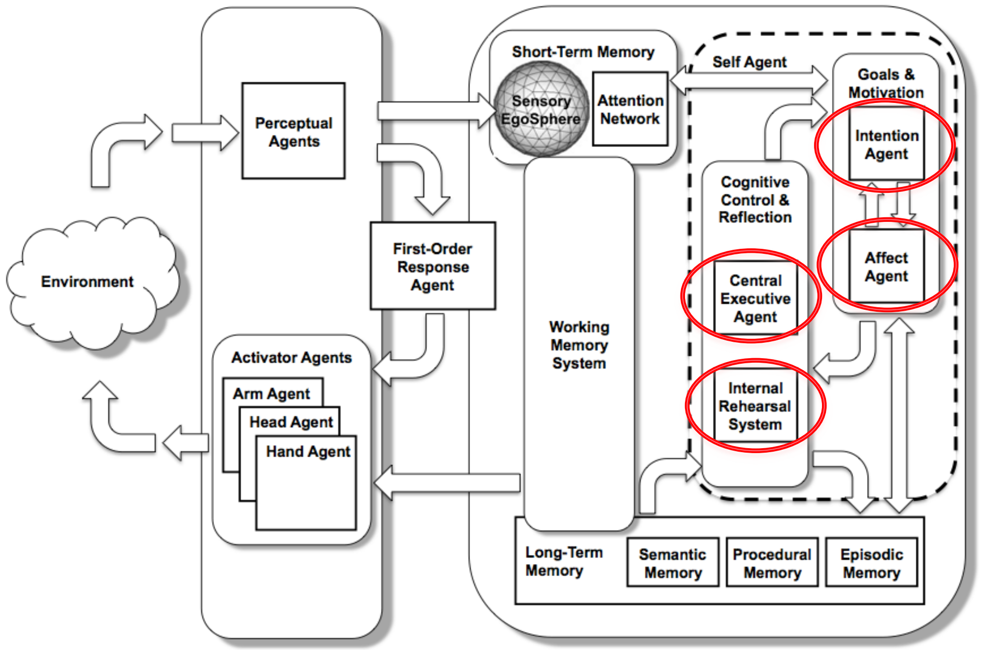
- Goals & Motivation sub-system
  - **Intention Agent**
  - **Affect Agent**



ISAC

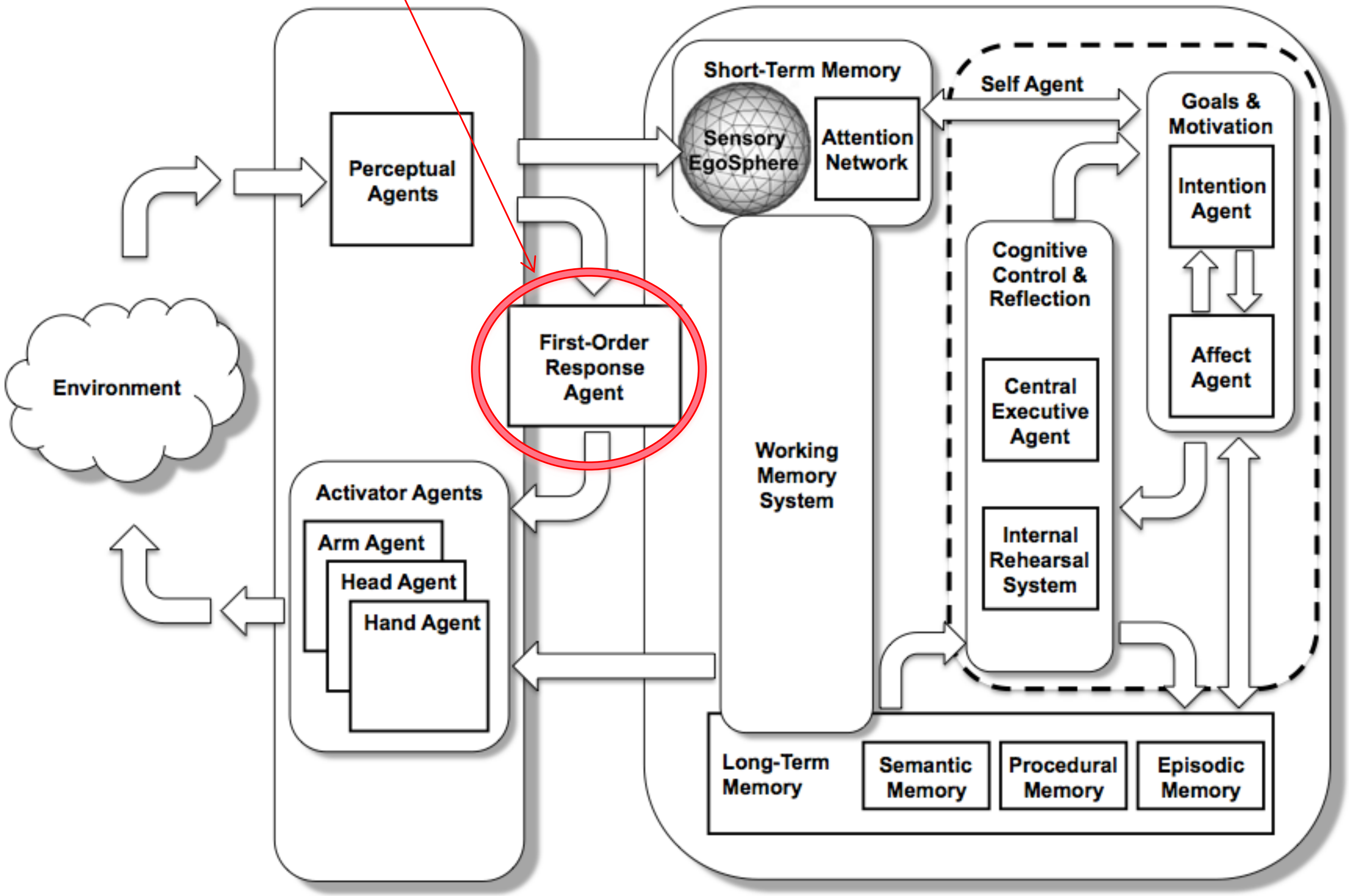
Cognitive behaviour is achieved through the interaction of several agents

- 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



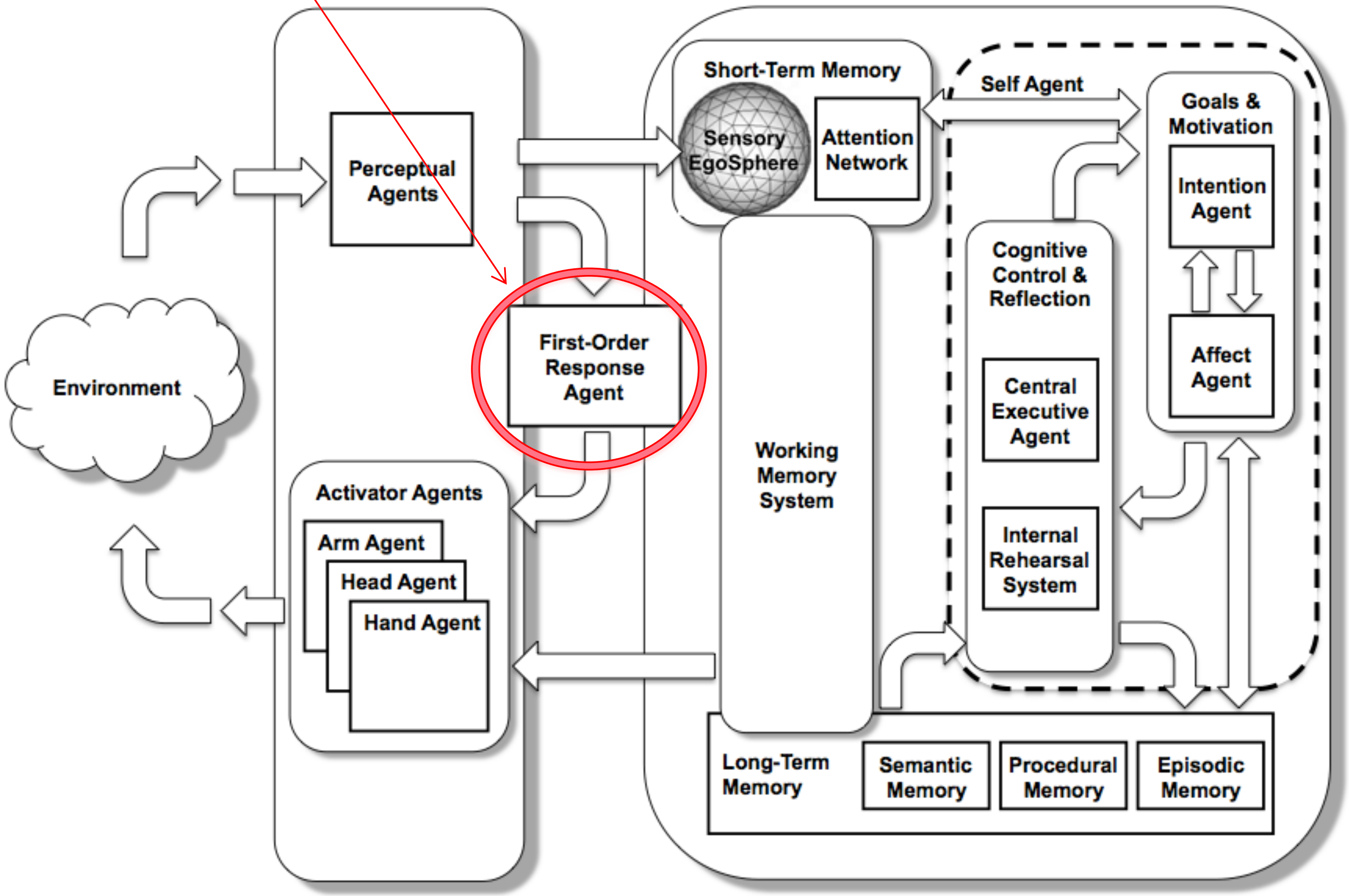
Normally, the First-order Response Agent (FRA) produces reactive responses to sensory triggers

ISAC



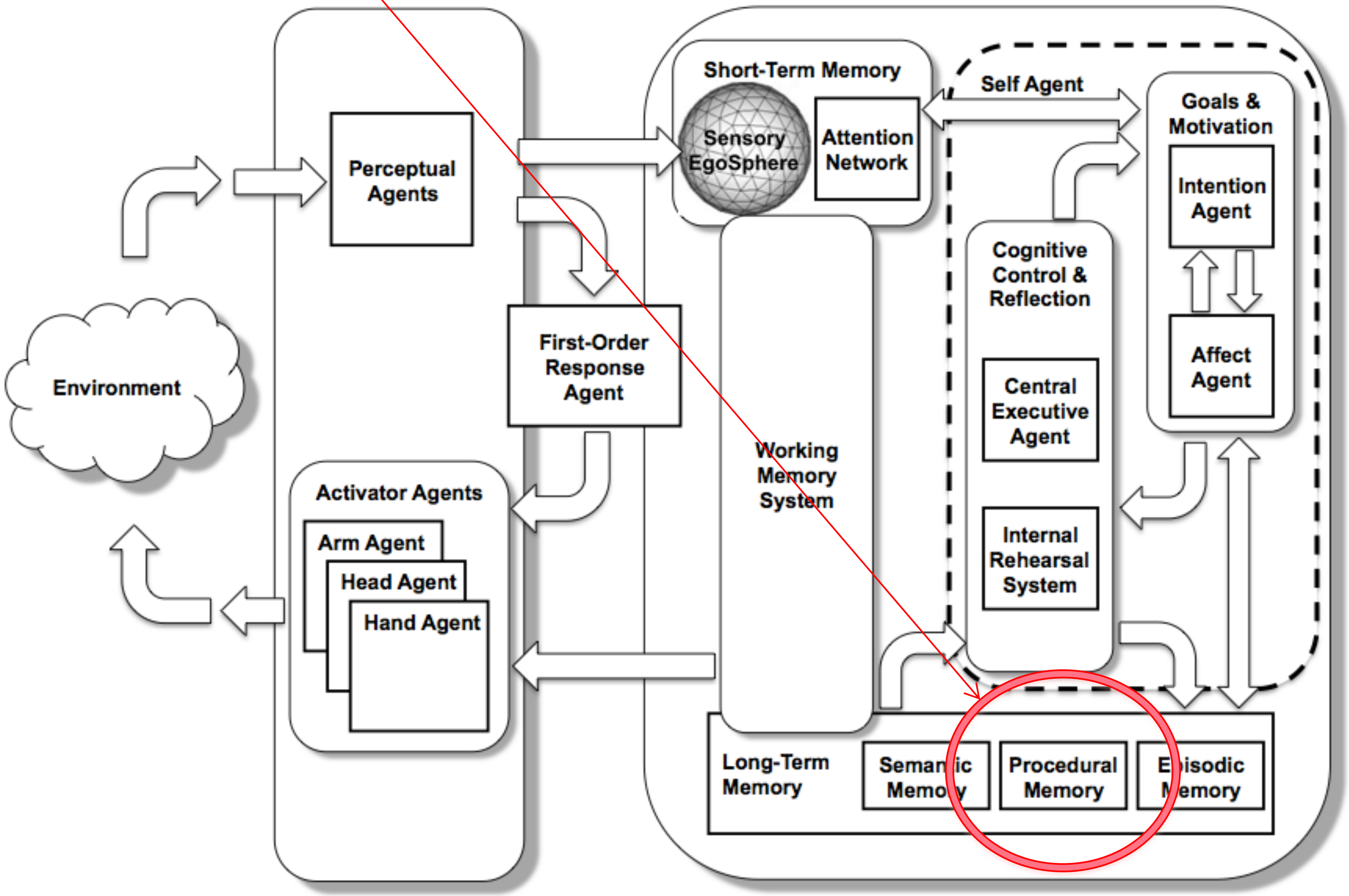
First-order Response Agent (FRA)
is also responsible for executing tasks

ISAC



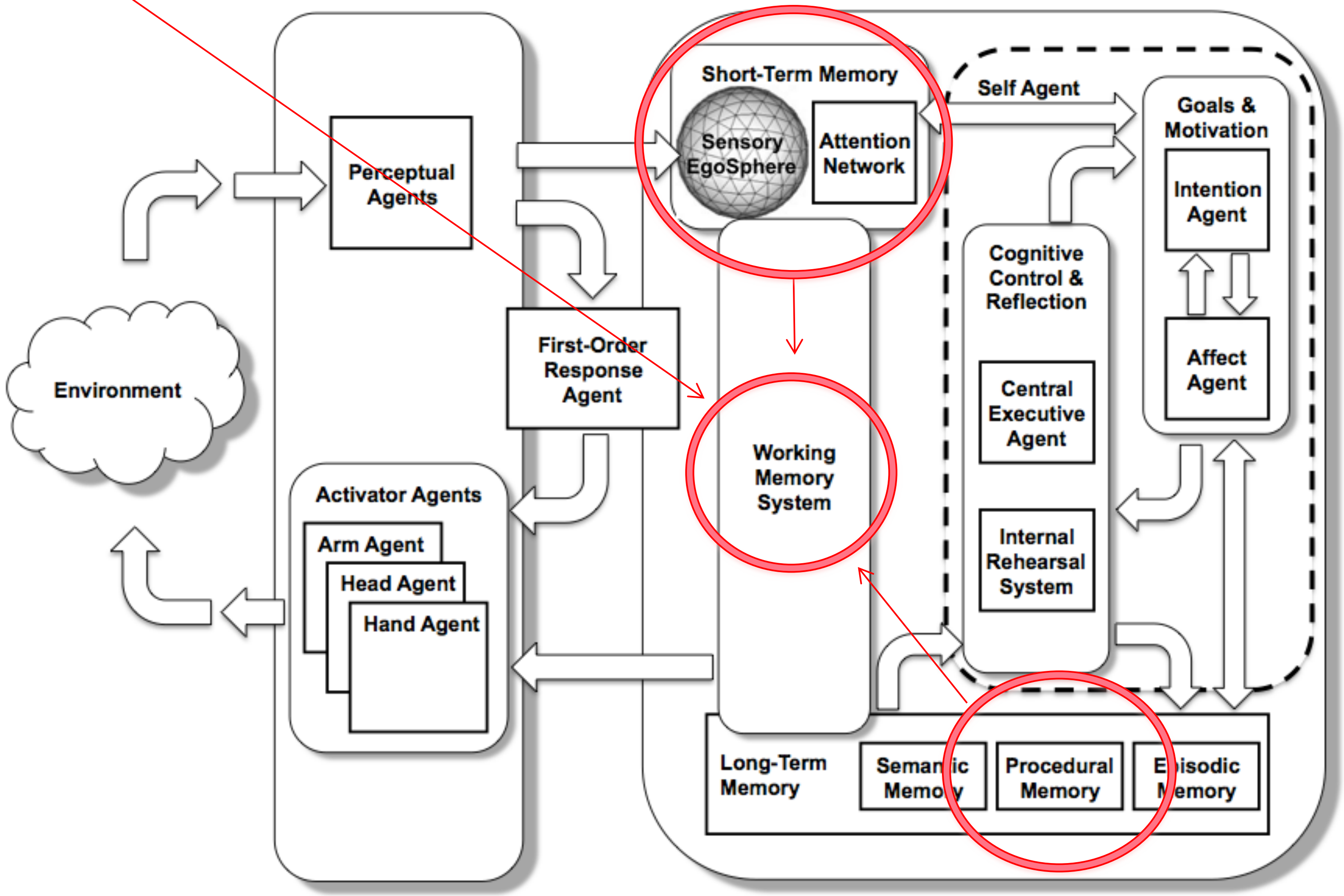
When a task is assigned by a human, the FRA retrieves the skill from procedural memory in LTM that corresponds to the skill described in the task information

ISAC



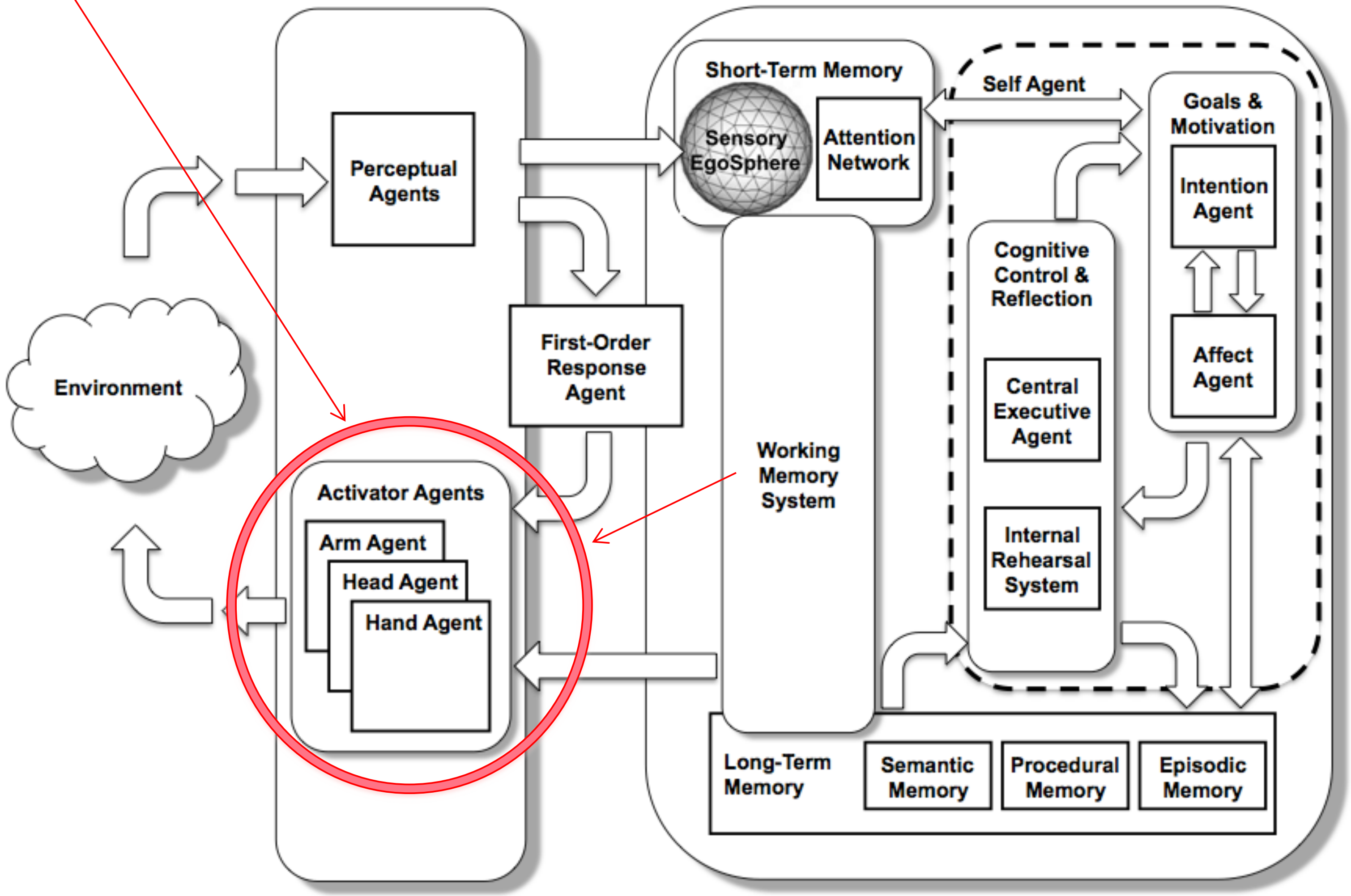
It then places it in the WMS as chunks along with the current percept

ISAC



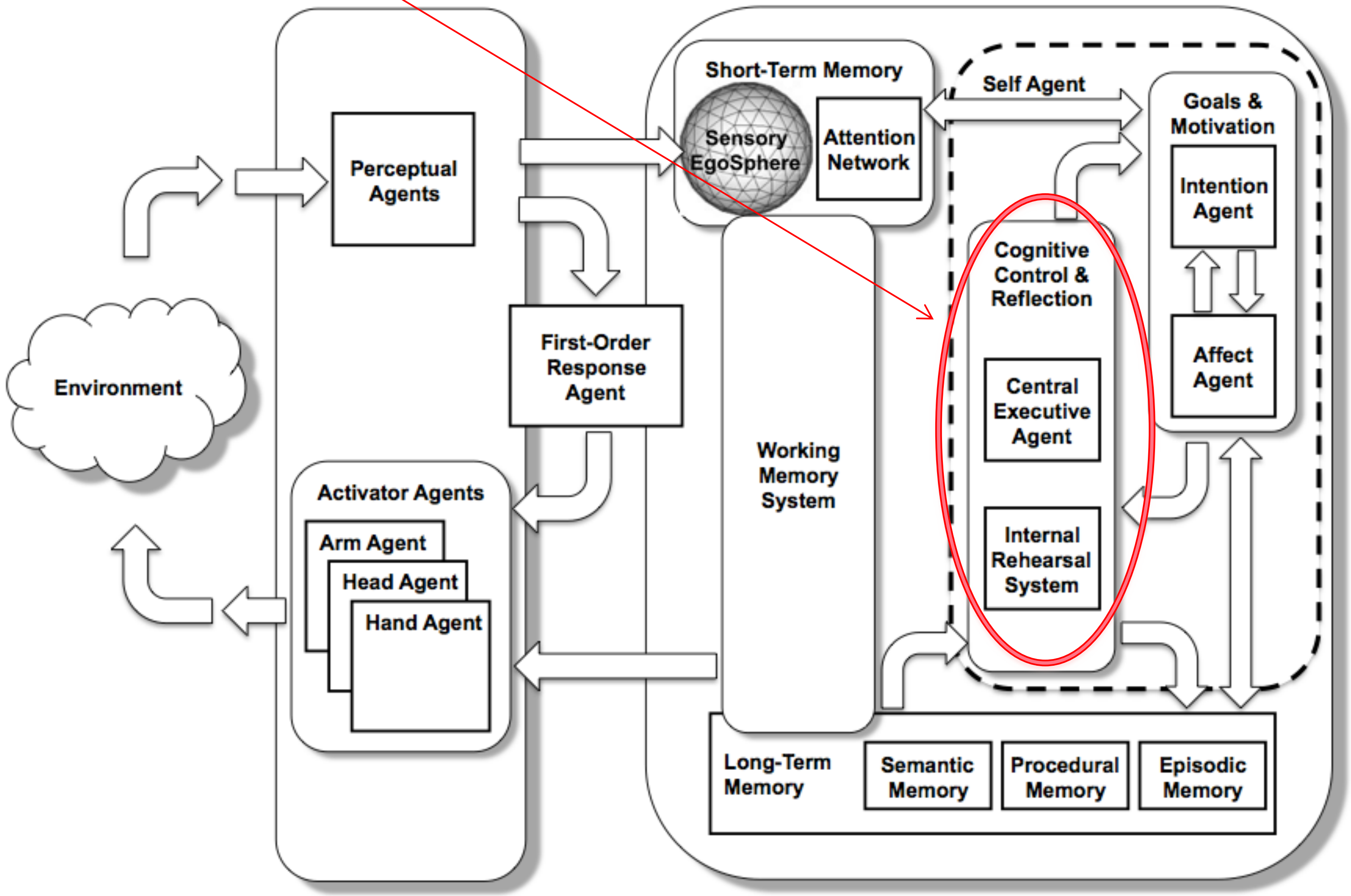
The Activator Agent then executes it, suspending execution whenever a reactive response is required

ISAC



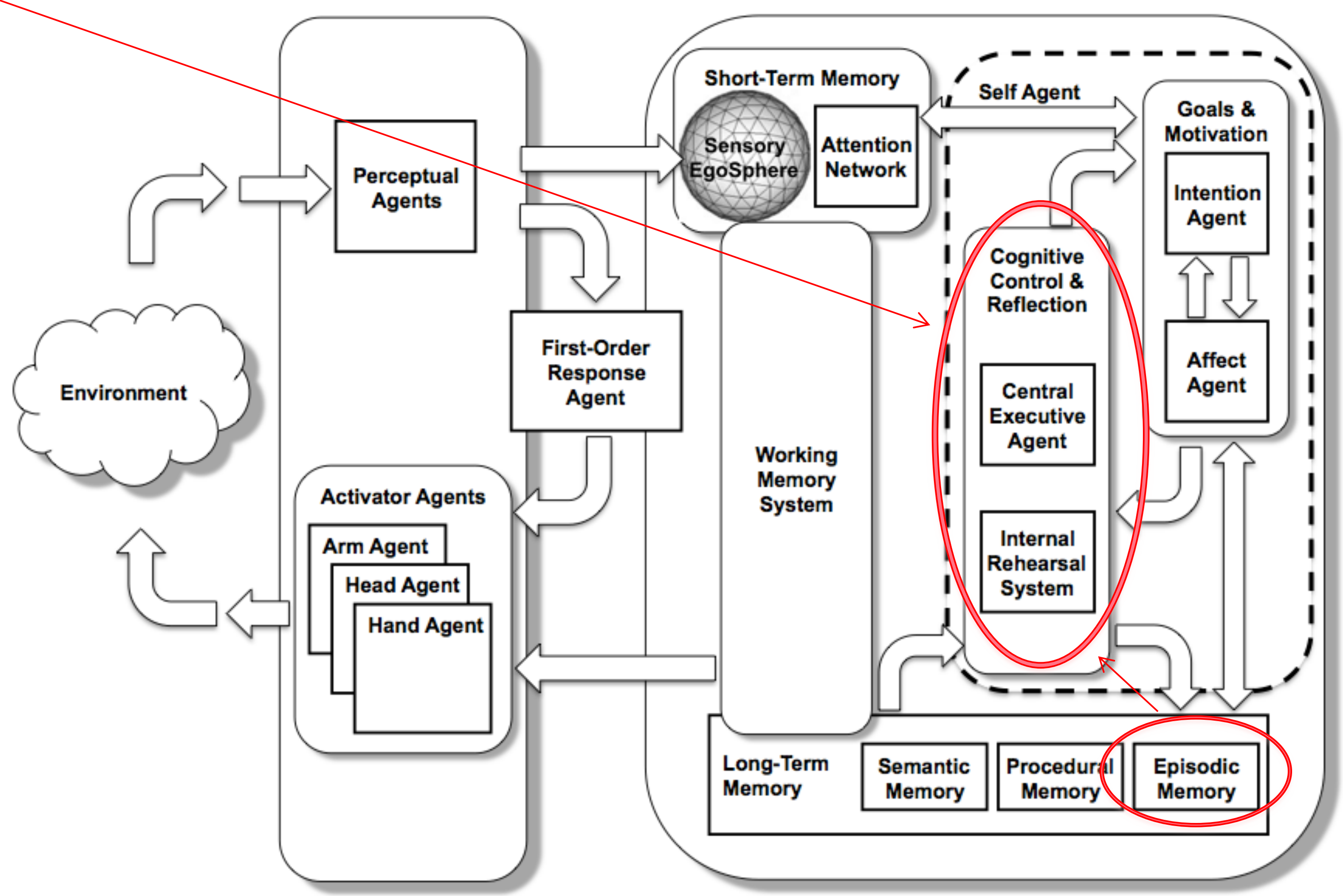
If the FRA finds no matching skill for the task, the Central Executive Agent takes over

ISAC



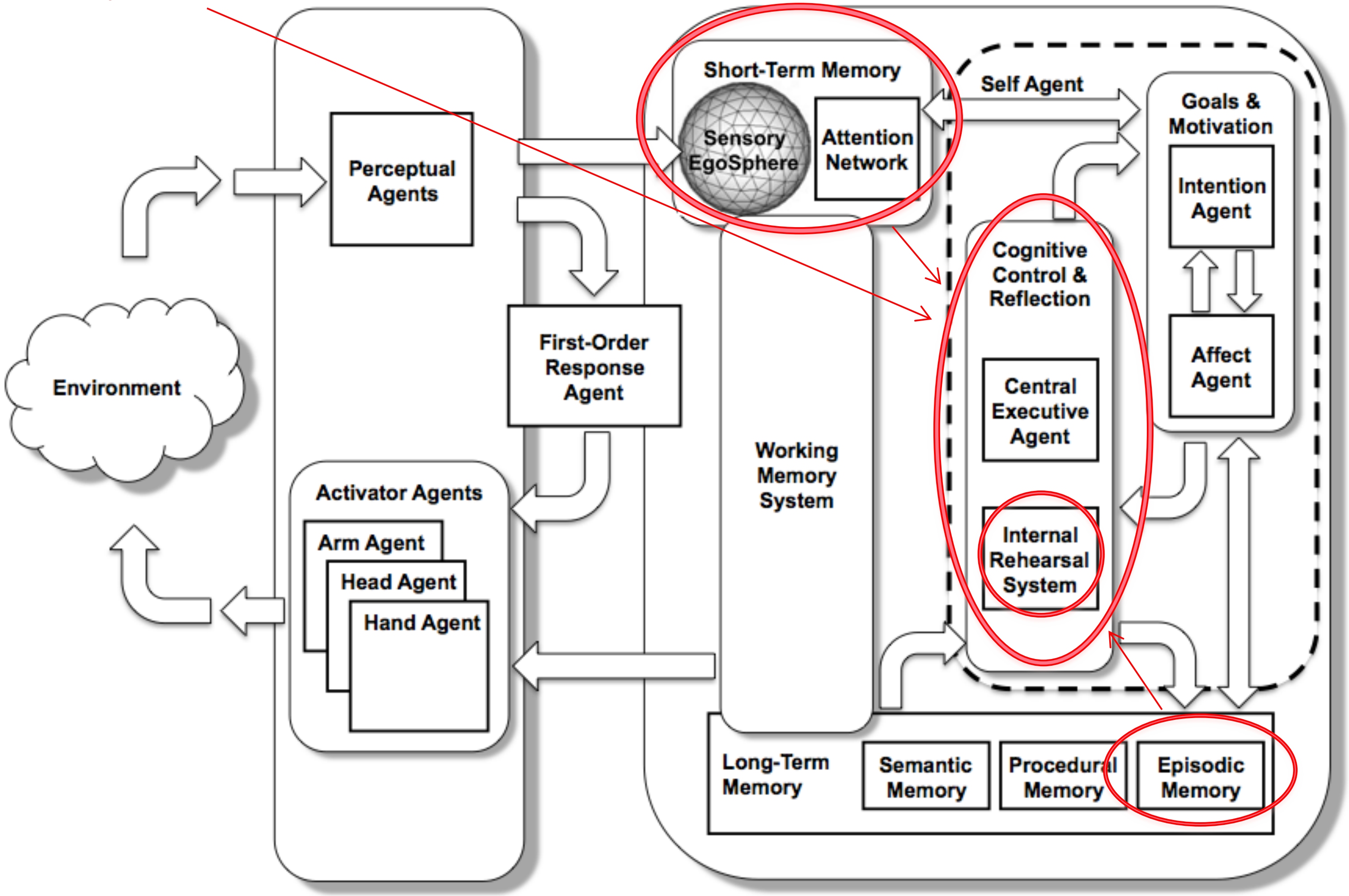
Recalls from episodic memory past experiences and behaviours that contain information similar to the current task

ISAC



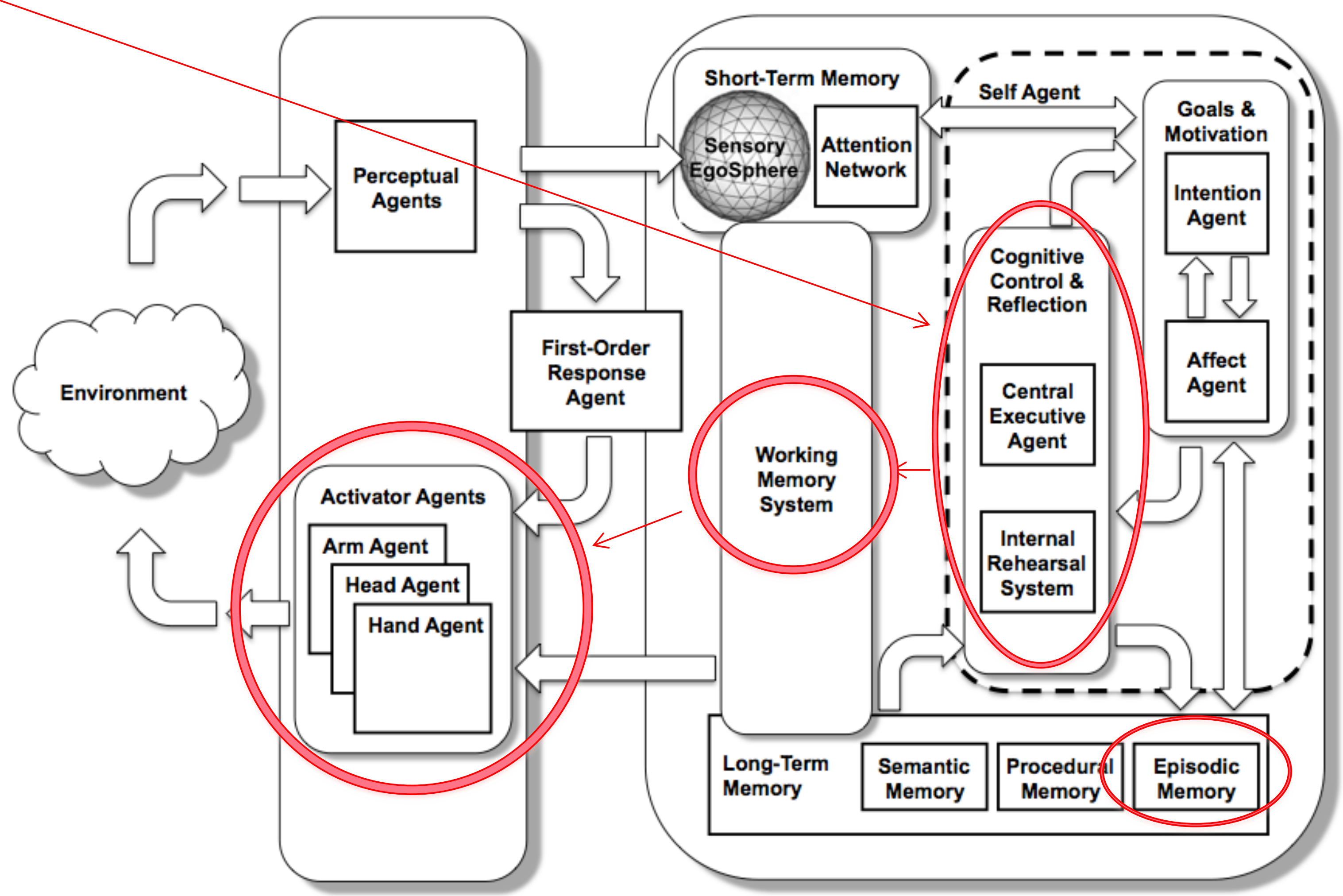
Select a behaviour-percept pair,
based on the current percept in the SES,
its relevance, and the likelihood of successful
execution as determined by internal simulation

ISAC



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

ISAC



Reading

D. Vernon, Artificial Cognitive Systems – A Primer, MIT Press, 2014; Chapter 3, Sections 3.4, 3.5, pp. 75-83.

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

A.3.6 (CLARION)

D. Vernon, "Cognitive Architectures", in Cognitive Robotics, A. Cangelosi and M. Asada (Eds.), MIT Press, Chapter 10, 2022, Section 10.6.2.

Further Reading

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.

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

R. Sun. The importance of cognitive architectures: an analysis based on CLARION. Journal of Experimental & Theoretical Artificial Intelligence 19(2), 159–193, 2007.

Recommended Videos

These and other short videos on cognitive architectures can be found at the 2021 TransAIR Workshop on Cognitive Architectures for Robot Agents

https://transair-bridge.org/workshop-2021/