Artificial Cognitive Systems

Module 3: Cognitive Architectures

Lecture 2: Example cognitive architectures: Soar, ACT-R, Global Workspace, LIDA, BBD

David Vernon
Carnegie Mellon University Africa

www.vernon.eu

Example Cognitive Architectures

Surveys:

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

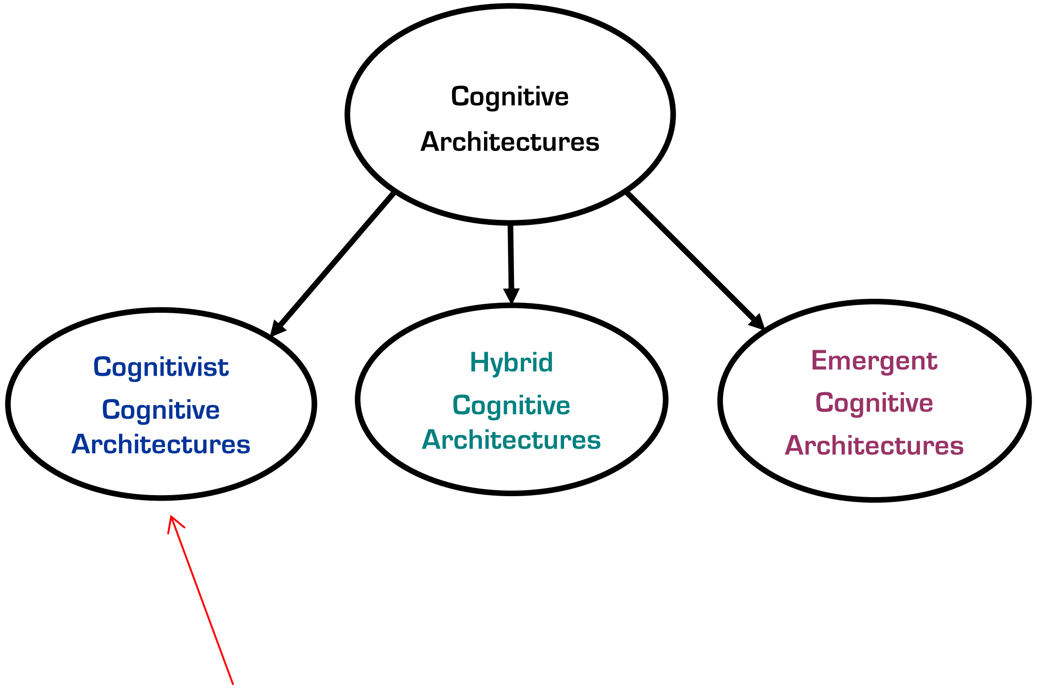
A Survey of Cognitive and Agent Architectures, University of Michigan, http://ai.eecs.umich.edu/cogarch0/ (12 cognitive architectures)

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)

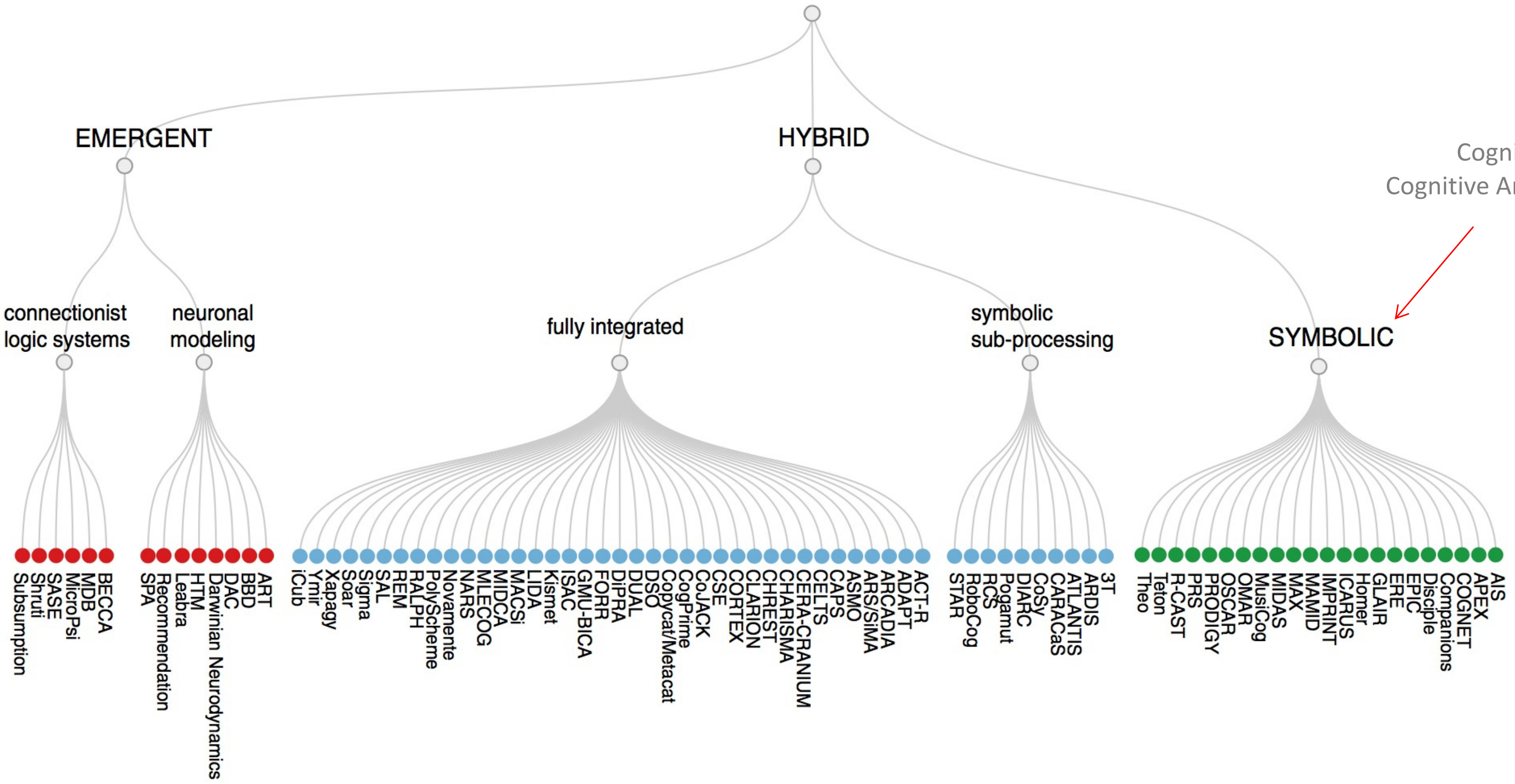
D. Vernon, C. von Hofsten, and L. Fadiga. "A Roadmap for Cognitive Development in Humanoid Robots", Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2011. Chapter 5 and Appendix I (20 cognitive architectures)

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)

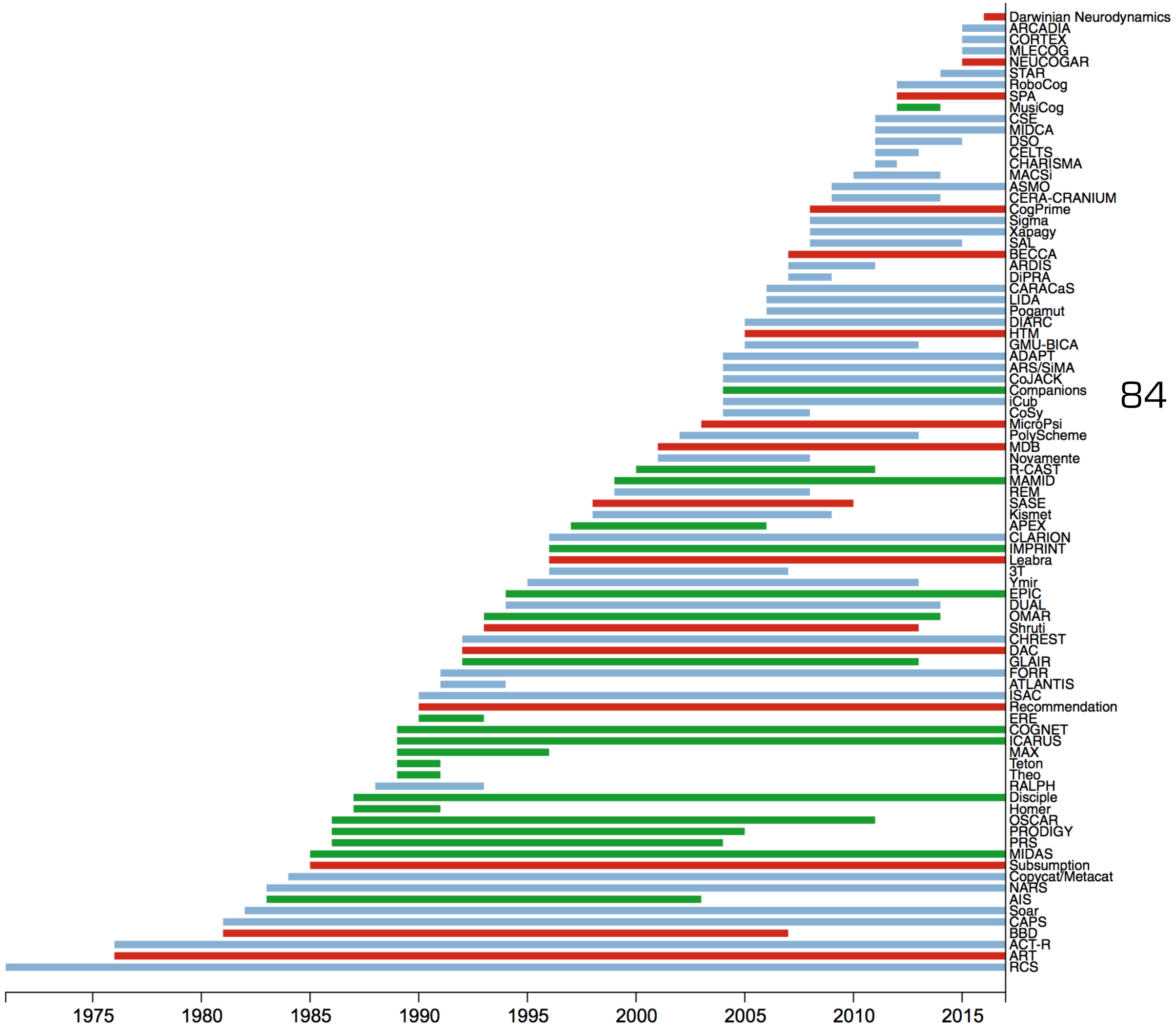


Kotseruba and Tsotsos refer to these as Symbolic Architectures

Cognitivist
Cognitive Architectures

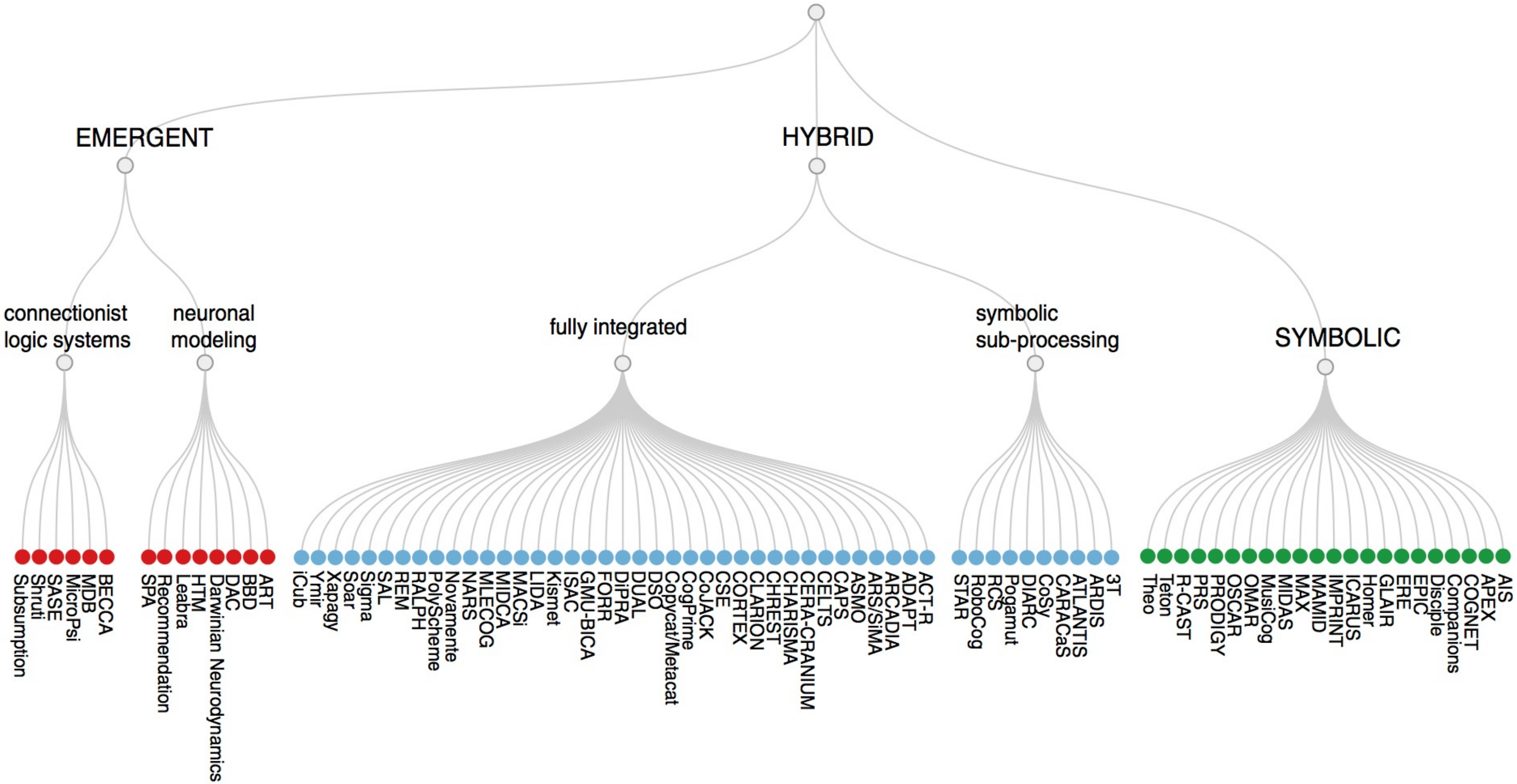


I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

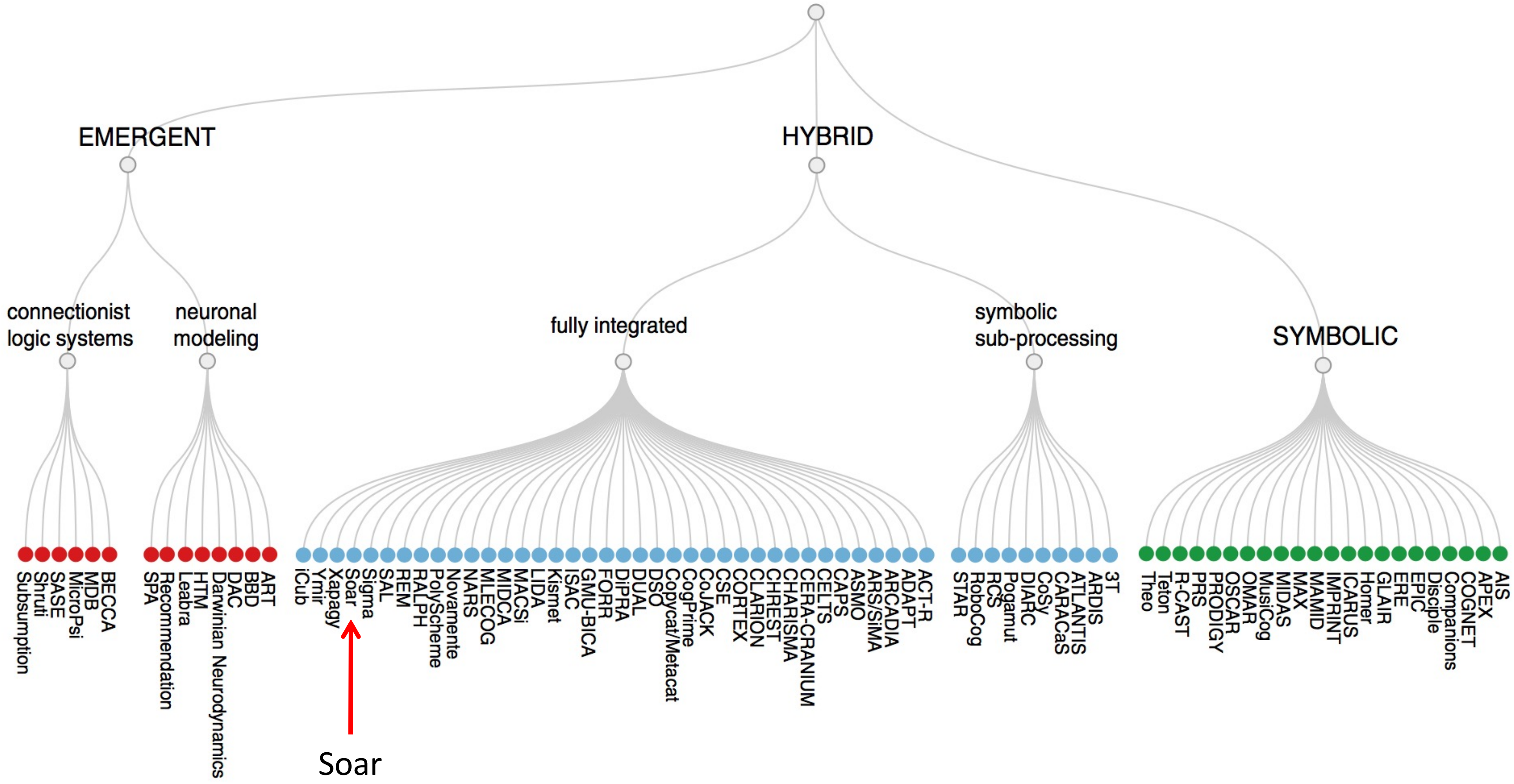


84 Cognitive Architectures

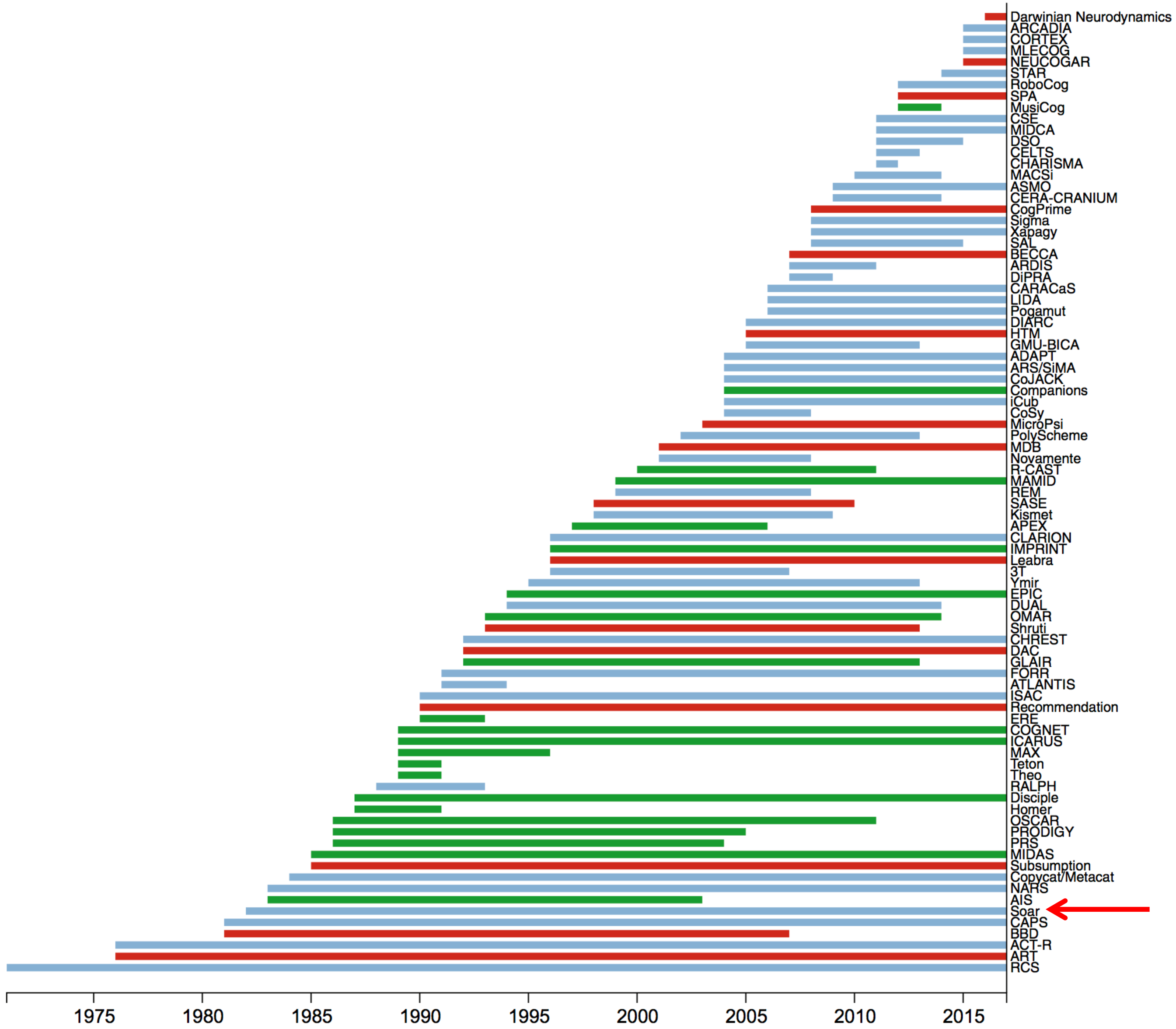
I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.



We will briefly sample some 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



160 A Catalogue of Cognitive Architectures

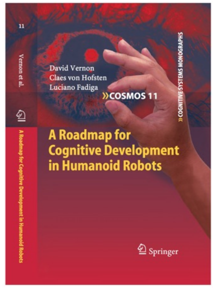
A.1 Cognitivist Cognitive Architectures

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.

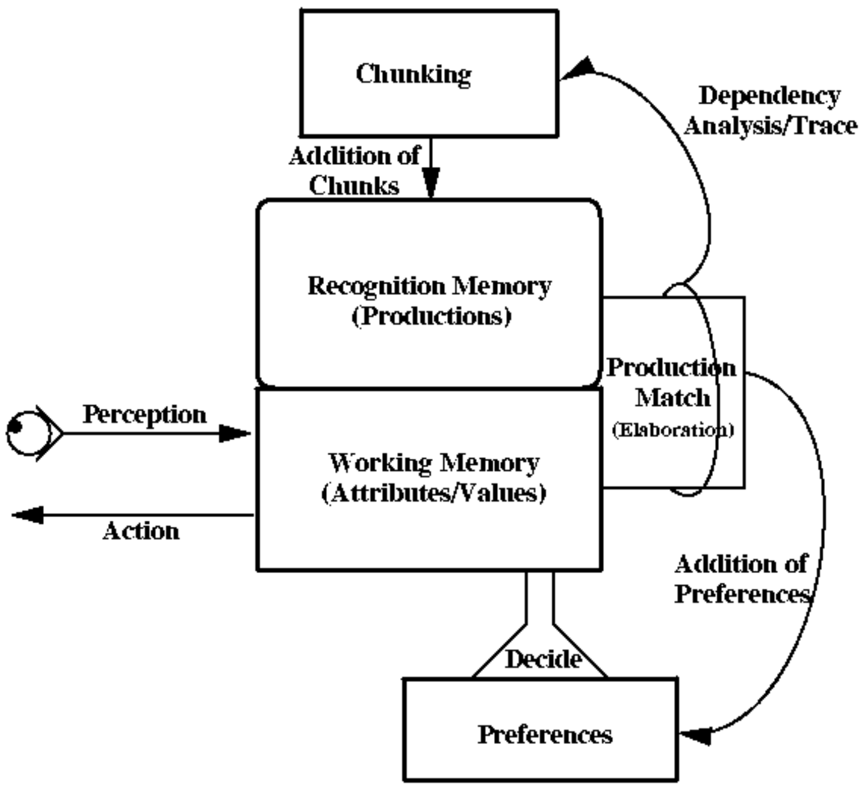


D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.

Soar

Soar [Laird et al. 1987]

- Newell's candidate for a Unified Theory of Cognition
- 1983 –
- v. 9.6
- Production system (i.e. rule-based)
- Cyclic operation
  - Production firing (all)
  - Decision (cf. preferences)
- Fine-grained knowledge representation
- Universal sub-goaling (dealing with impasse)
- General-purpose learning (encapsulates resolution of impasse)



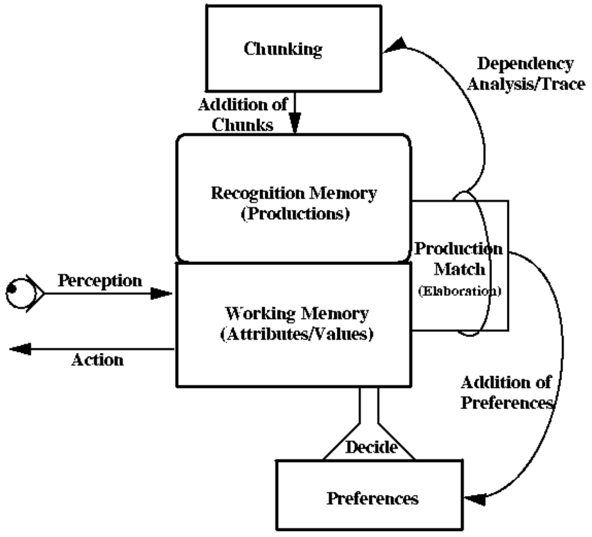
(Based on Figure 3.1, pg 20, The Soar's User Manual, Version 6)

https://soar.eecs.umich.edu/

Soar

Operates in a cyclic manner

- Production cycle
  - 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.
- Decision cycle
  - a single action from several possible actions is selected
  - The selection is based on stored action preferences.

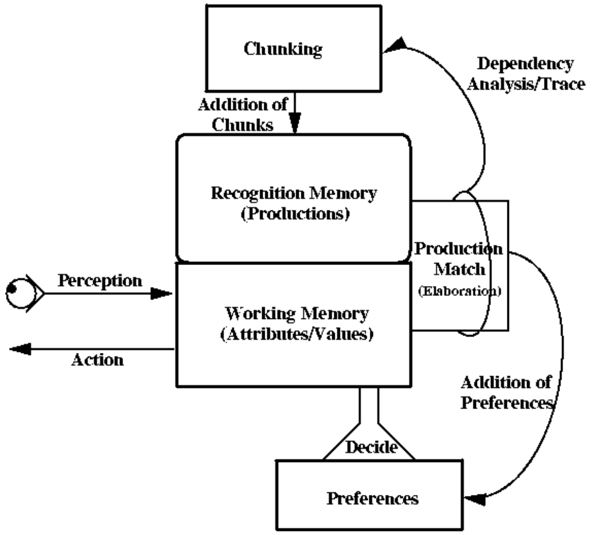


(Based on Figure 3.1, pg 20, The Soar's User Manual, Version 6)

Soar

Universal sub-goaling

- There no guarantee that the action preferences will lead to
  - a unique action or
  - any action
- In this case, the decision cycle may lead to an 'impasse'
  - 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 other impasses 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

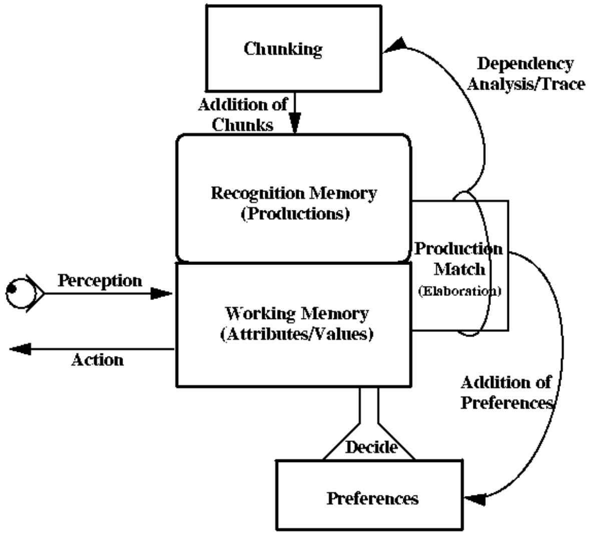


(Based on Figure 3.1, pg 20, The Soar's User Manual, Version 6)

Soar

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
- Resolving an impasse alters the system super-state
  - This change is called a result
  - It 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
  - It is the only form that occurs in Soar
  - i.e. Soar only learns new production rules: **chunks**

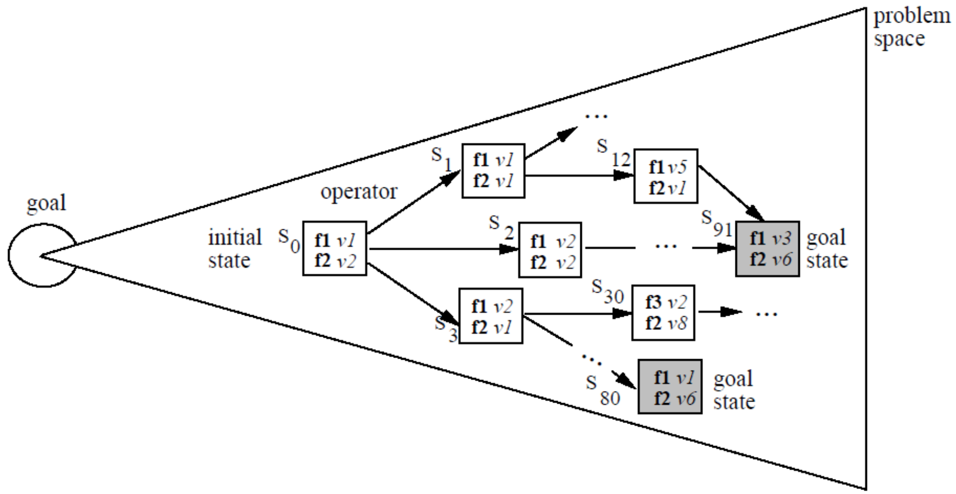


(Based on Figure 3.1, pg 20, The Soar's User Manual, Version 6)

Soar

Behaviour as movement through problem spaces

- Goal (circle)
- Problem space: expanding set of possibilities that can unfold over time (triangle)
- States (rectangles)
  - Vocabulary of features (bold)
  - Their possible values (italics) … values can also be a set of features
- State transition (arrows) … operators reflecting internal or external behaviour



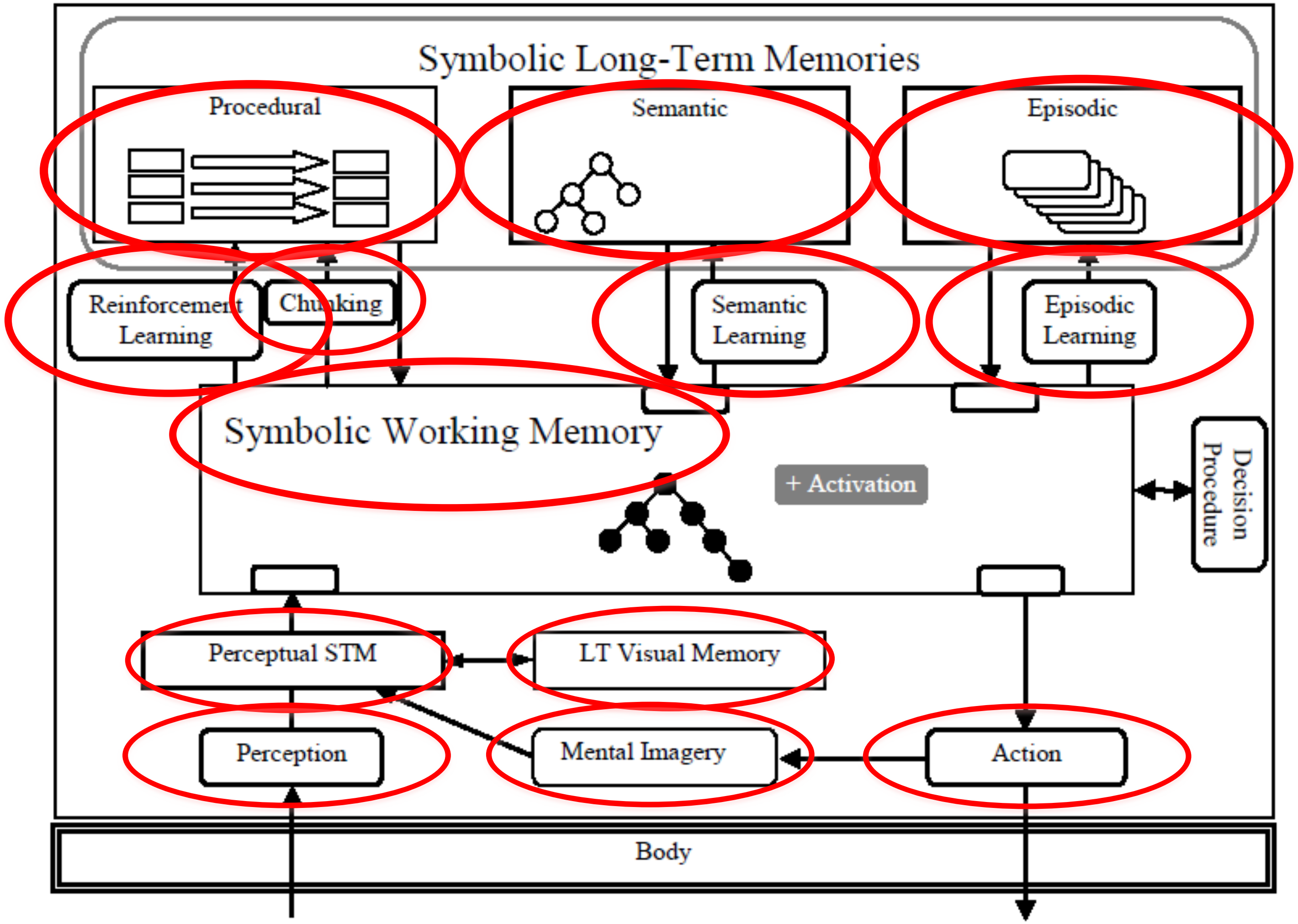
Soar

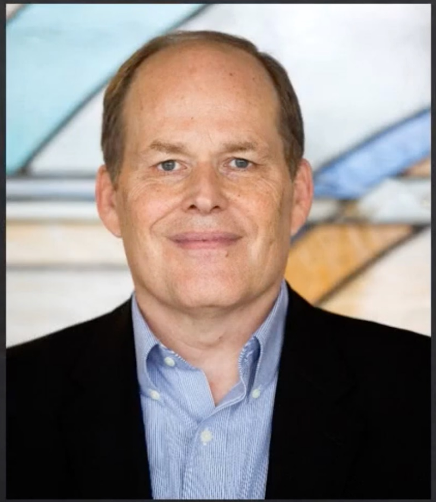
Tying the content to the architecture

Knowledge about things in the world	K1: Knowledge of the objects in the game e.g. baseball, infield, base line, inning, out, ball/strike count K2: Knowledge of abstract events and particular episodes e.g. how batters hit, how this guy batted last time he was up
Knowledge about abstract ideas	K3: Knowledge of the rules of the game e.g. number of outs, balk, infield fly K4: Knowledge of objectives e.g. get the batter out, throw strikes
Knowledge about mental actions	K5: Knowledge of actions or methods for attaining objectives e.g. use a curve ball, throw to first, walk batter
Knowledge about how to use knowledge	K6: Knowledge of when to choose actions or methods e.g. if behind in the count, throw a fast ball
Knowledge about Physical actions	K7: Knowledge of the component physical actions e.g. how to throw a curve ball, catch, run

Soar

(Laird et al. 2012)





Open Research and the Soar Cognitive Architecture

John Laird

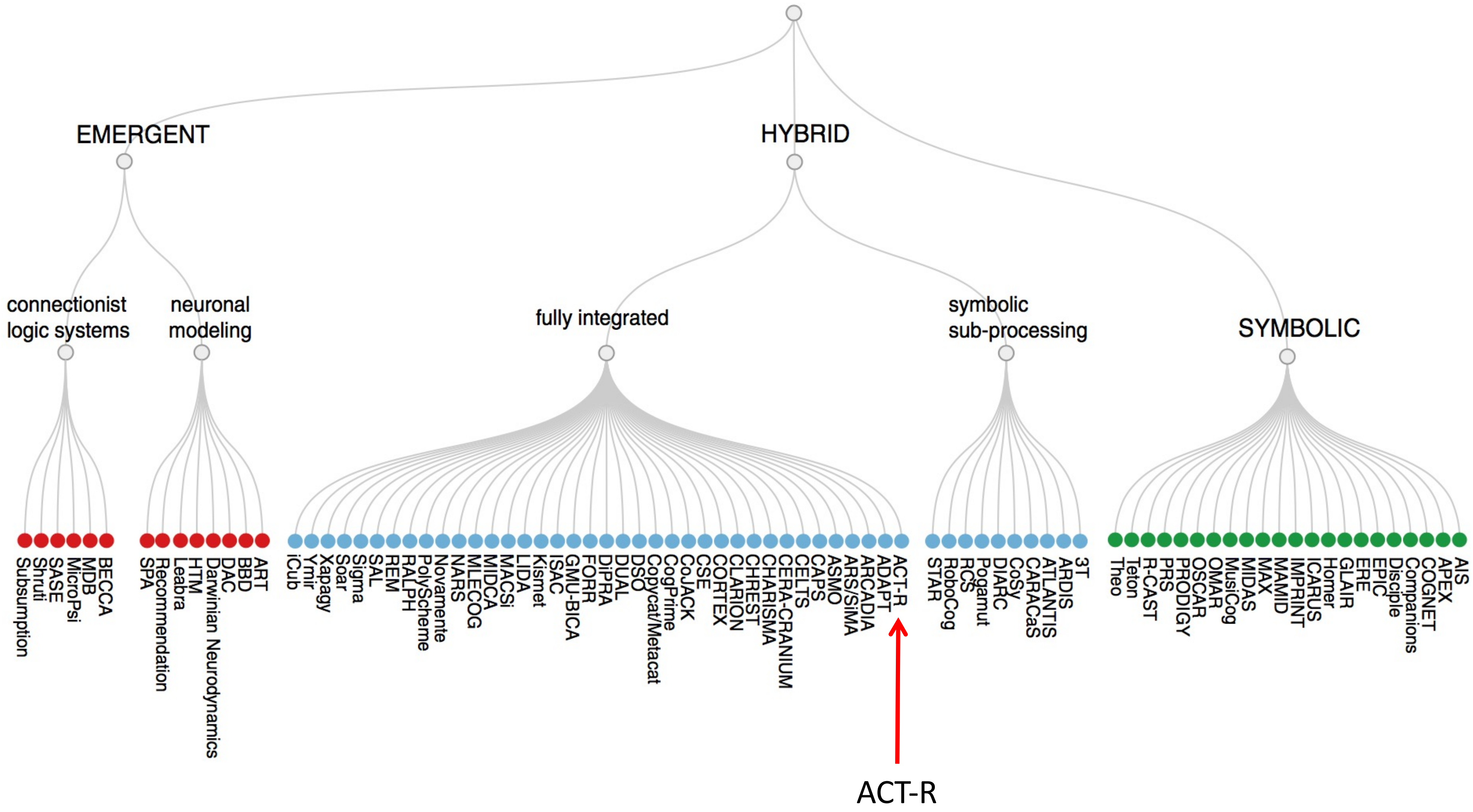
Democratizing AI

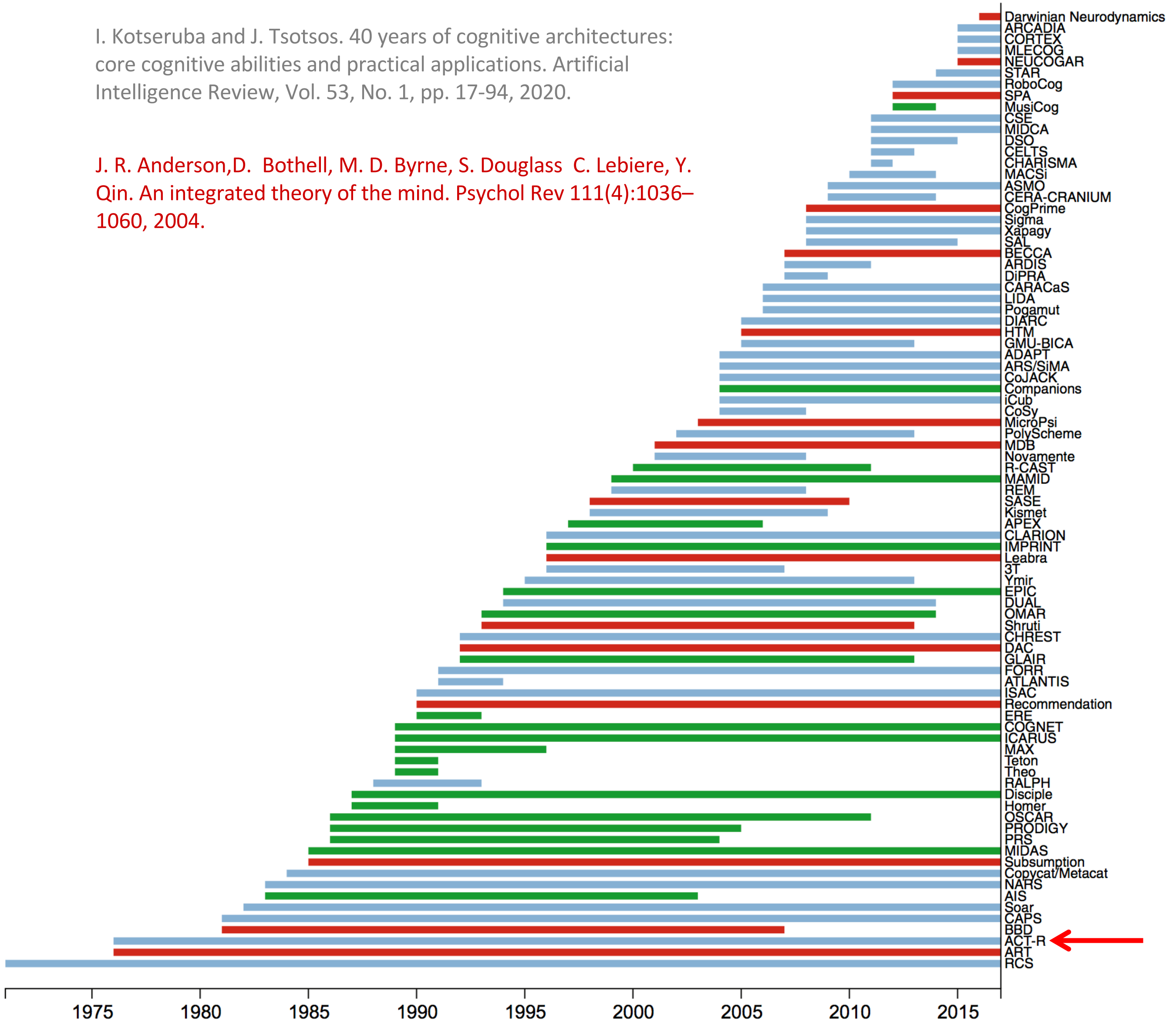
October 6, 2020

https://soar.eecs.umich.edu/

The following is a short excerpt

For the full talk see: https://www.youtube.com/watch?v=2pNsfBj7XSA&feature=youtu.be


EMERGENT
HYBRID
SYMBOLIC
connectionist logic systems
neuronal modeling
fully integrated
symbolic sub-processing
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
ACT-R



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.

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.

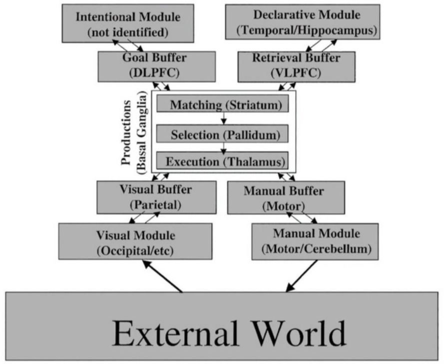
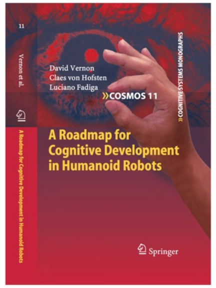


Fig. A.1 The ACT-R Cognitive Architecture (from [7])

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.

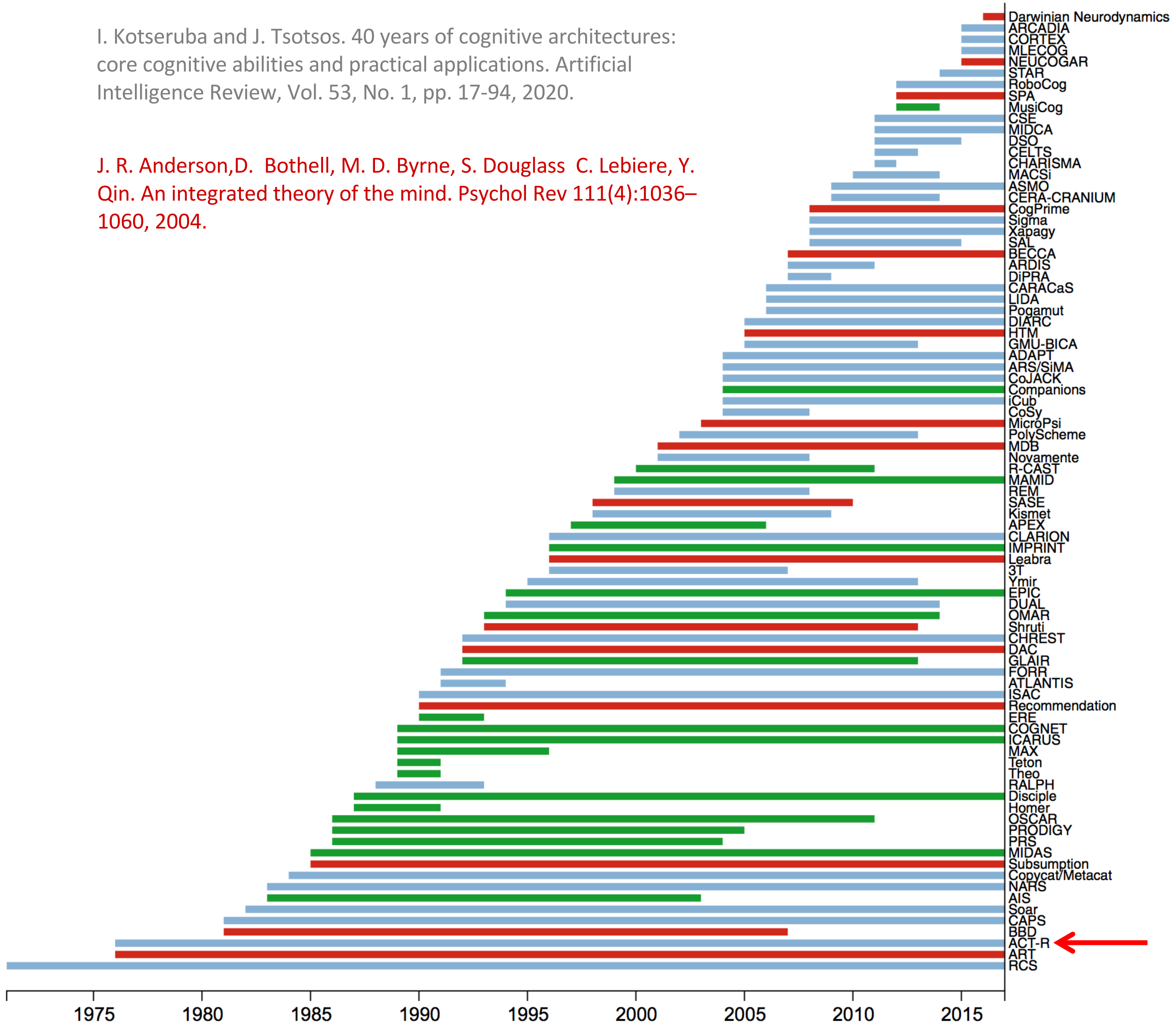
There are two serial bottle-necks in ACT-R. One is that the content of any buffer is limited to a single declarative unit of knowledge, called a 'chunk'. This implies that only one memory can be retrieved at a time and indeed that a single object can be encoded in the visual field at any one time. The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



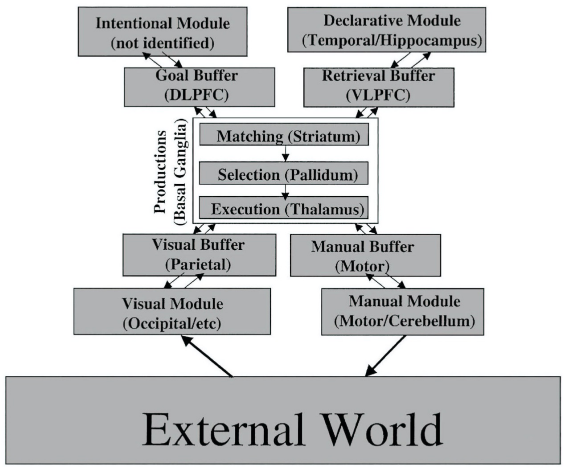
D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.

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.

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



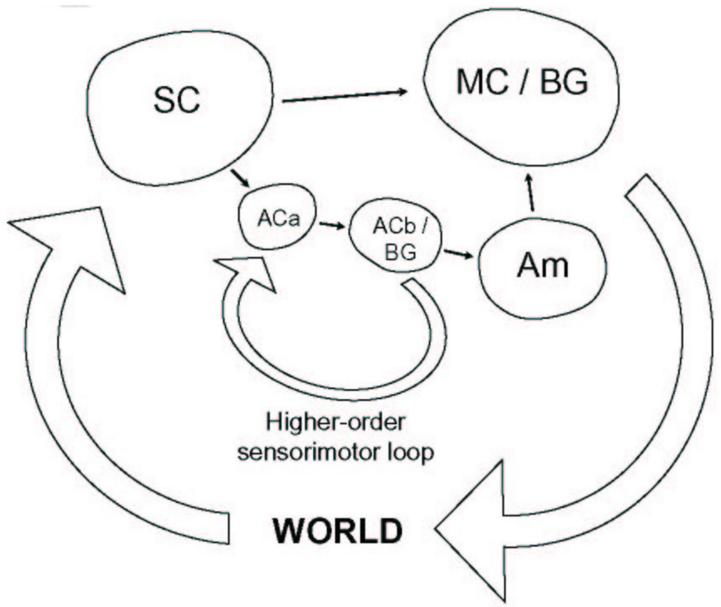
- J. Anderson's candidate for a Unified Theory of Cognition
- 1996, 2004; now version 7
- Production system with five modules: Intentional, Declarative, Visual, Manual, Production
- Cyclic operation: executes one production per cycle
  - Pattern of information in the buffers is recognized
  - A single production is selected and fires
  - The buffers are updated
- Each cycle takes approximately 50 ms.

Global Workspace

Shanahan's Global Workspace Architecture

- Anticipation and planning achieved through internal simulation
- Action selection (internal and external) mediated by affect
- Analogical representation (small semantic gap & easier grounding)
- Global workspace model: parallelism is a fundamental component of the architecture, not an implementation issue

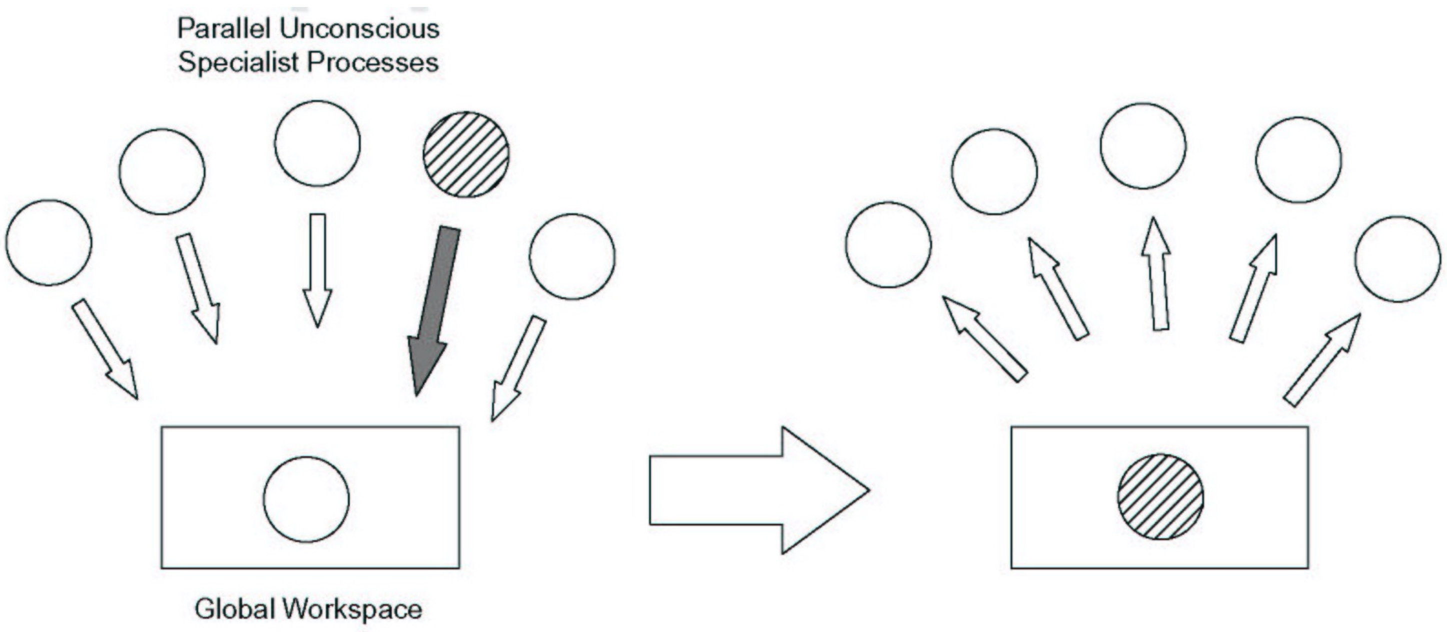
(Shanahan06,ShanahanBaars06,Shanahan05a,Shanahan05b)



SC Sensory Cortex
MC Motor Cortex
BG Basal Ganglia (action selection)
AC Association Cortex
Am Amygdala (affect)

Global Workspace

Global workspace model: sequence of states emerge from multiple competing and cooperating parallel processes



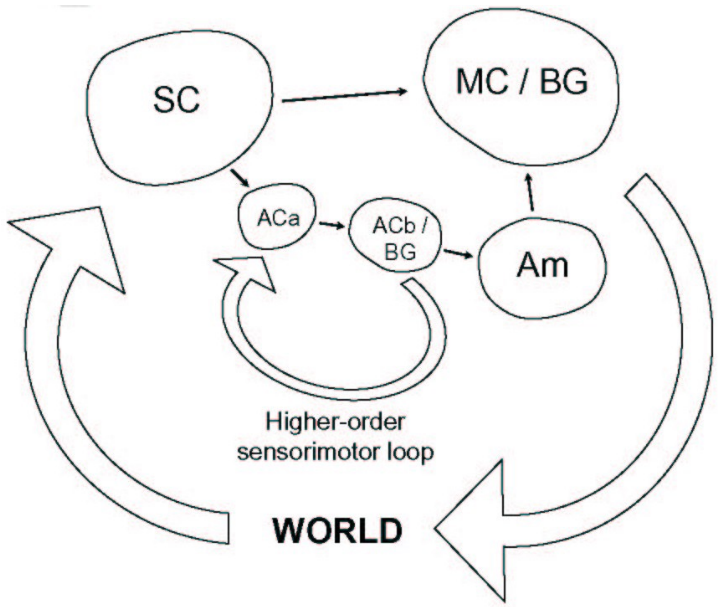
Global Workspace

Global workspace model: sequence of states emerge from multiple competing and cooperating parallel processes

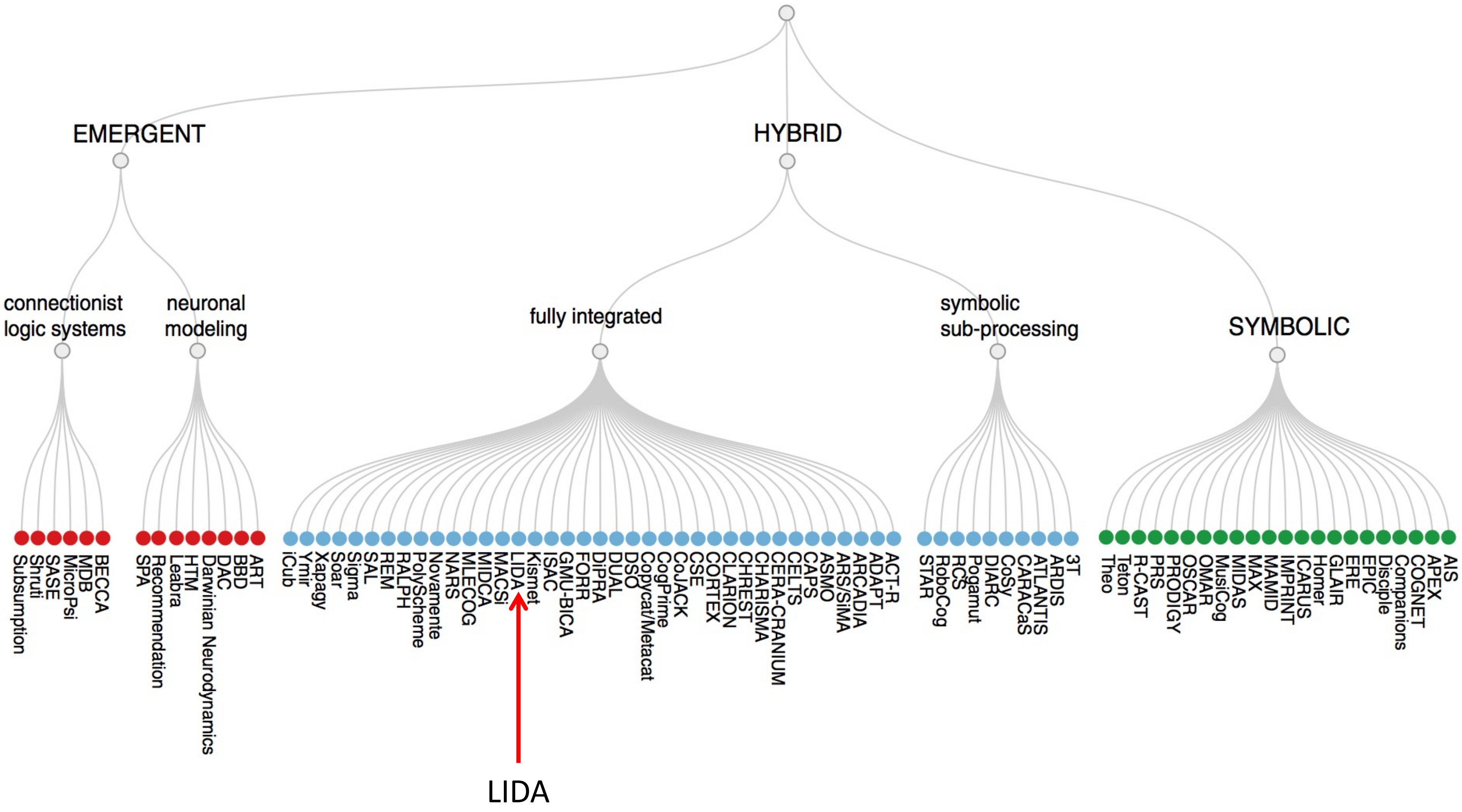
Global Workspace

Shanahan's Global Workspace Architecture

- Implemented using G-RAMS (generalized random access memories)
- Global workspace and cortical assemblies define an attractor landscape
- Perceptual categories define attractors
- Higher-order loop allows the GW to visit these attractors



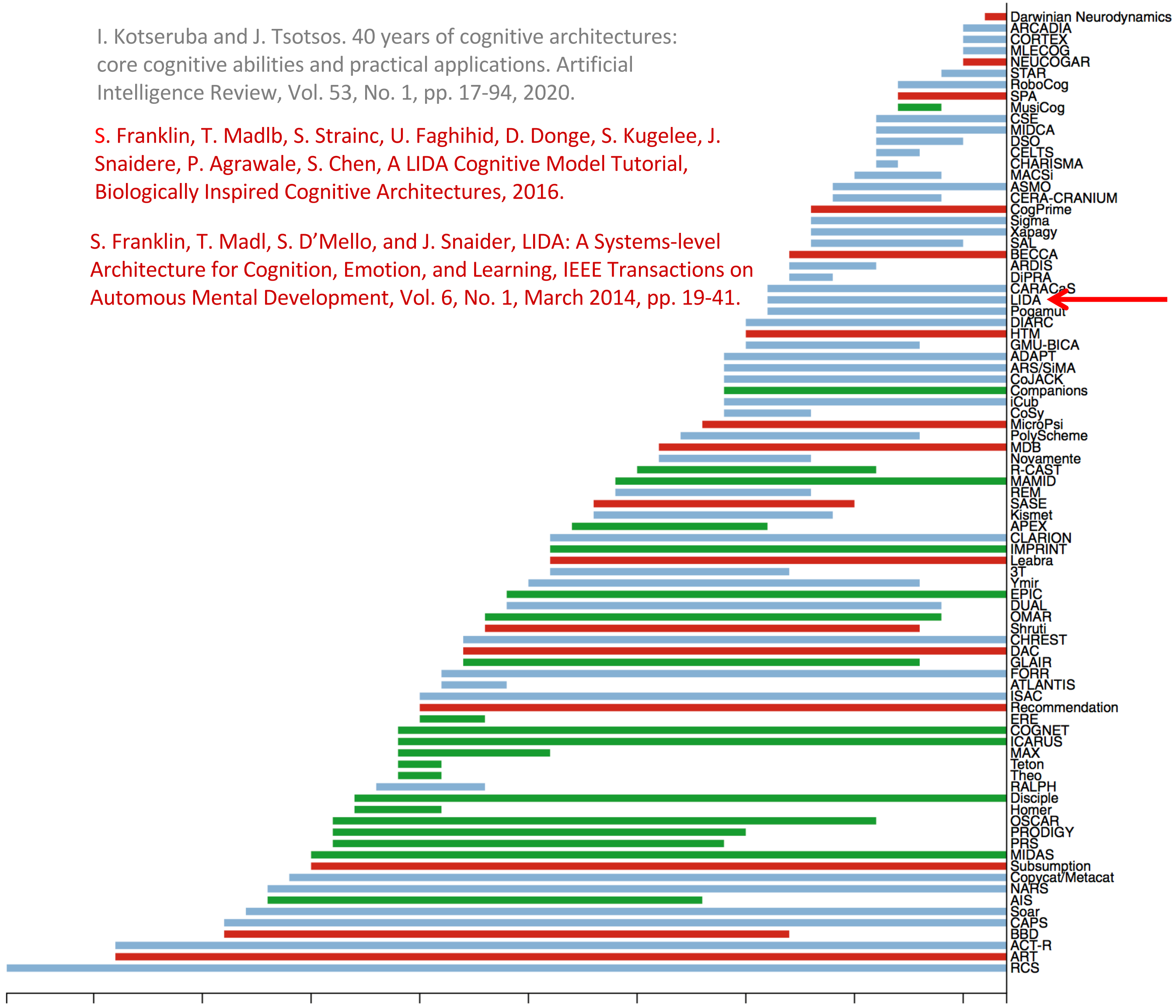
SC Sensory Cortex
MC Motor Cortex
BG Basal Ganglia (action selection)
AC Association Cortex
Am Amygdala (affect)


EMERGENT
HYBRID
SYMBOLIC
connectionist logic systems
neuronal modeling
fully integrated
symbolic sub-processing
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
LIDA

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.

S. Franklin, T. Madlb, S. Strainc, U. Faghihid, D. Donge, S. Kugelee, J. Snaidere, P. Agrawale, S. Chen, A LIDA Cognitive Model Tutorial, Biologically Inspired Cognitive Architectures, 2016.

S. Franklin, T. Madl, S. D'Mello, and J. Snaider, LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning, IEEE Transactions on Automous Mental Development, Vol. 6, No. 1, March 2014, pp. 19-41.



192 A Catalogue of Cognitive Architectures

A.3.5 *The LIDA Cognitive Architecture*

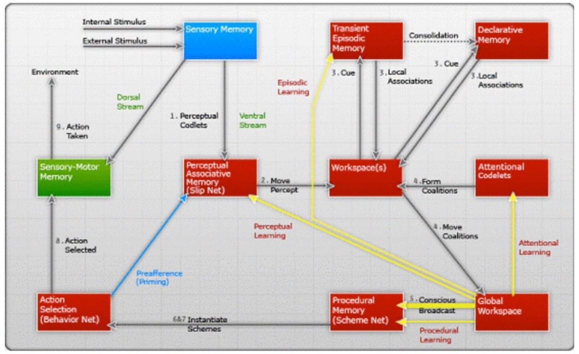


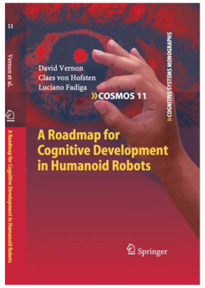
Fig. A.9 The LIDA cognitive cycle (from [17])

LIDA (Learning Intelligent Distribution Agent) is a hybrid cognitive architecture which combines features of both symbolic cognitivist and connectionist approaches [17, 103, 104, 106, 299]. It deploys several modules and processes to effect attention, action selection, and learning. The operation of LIDA is based around the concept of an atomic cognitive action-perception cycle. Each cycle comprises three phases: understanding, attending, and action selection (see Figure A.9).

The understanding phase involves sampling or sensing the environment and then it "makes sense" of its current situation by updating its representation of external sensory-derived features and internally-generated features of the agent's world comprising objects, categories, relations, events, and situations. These features are stored in a sensory memory module and a perceptual memory module, respectively.

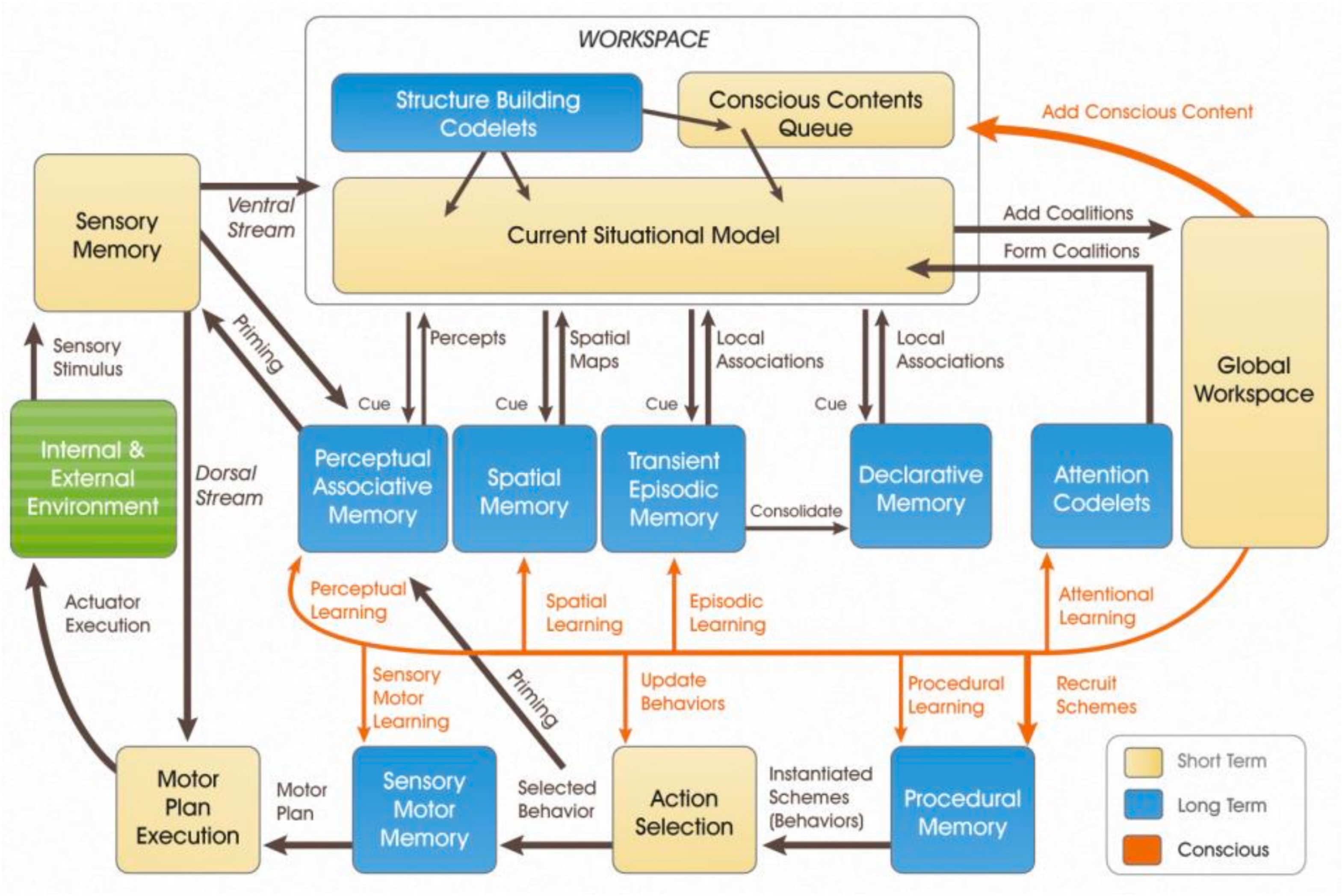
The attending phase decides what aspect of the current situation model requires attention. This attentional process uses a mechanism adapted from Global Workspace Theory [15, 16] whereby each portion of the model competes for attention by being moved to a global workspace where a single portion of this model is selected. This portion is then broadcast back to the rest of the system. The contents of the broadcast yields a set of potential actions which are then subjected to a further competition in the subsequent action selection phase.

The initial representation of the current situation resulting from the understanding phase is stored in the perceptual memory. This is used by the workspace module to access transient and declarative episodic memories of events. Both episodic memories use these inputs to recall associatively past experiences. These

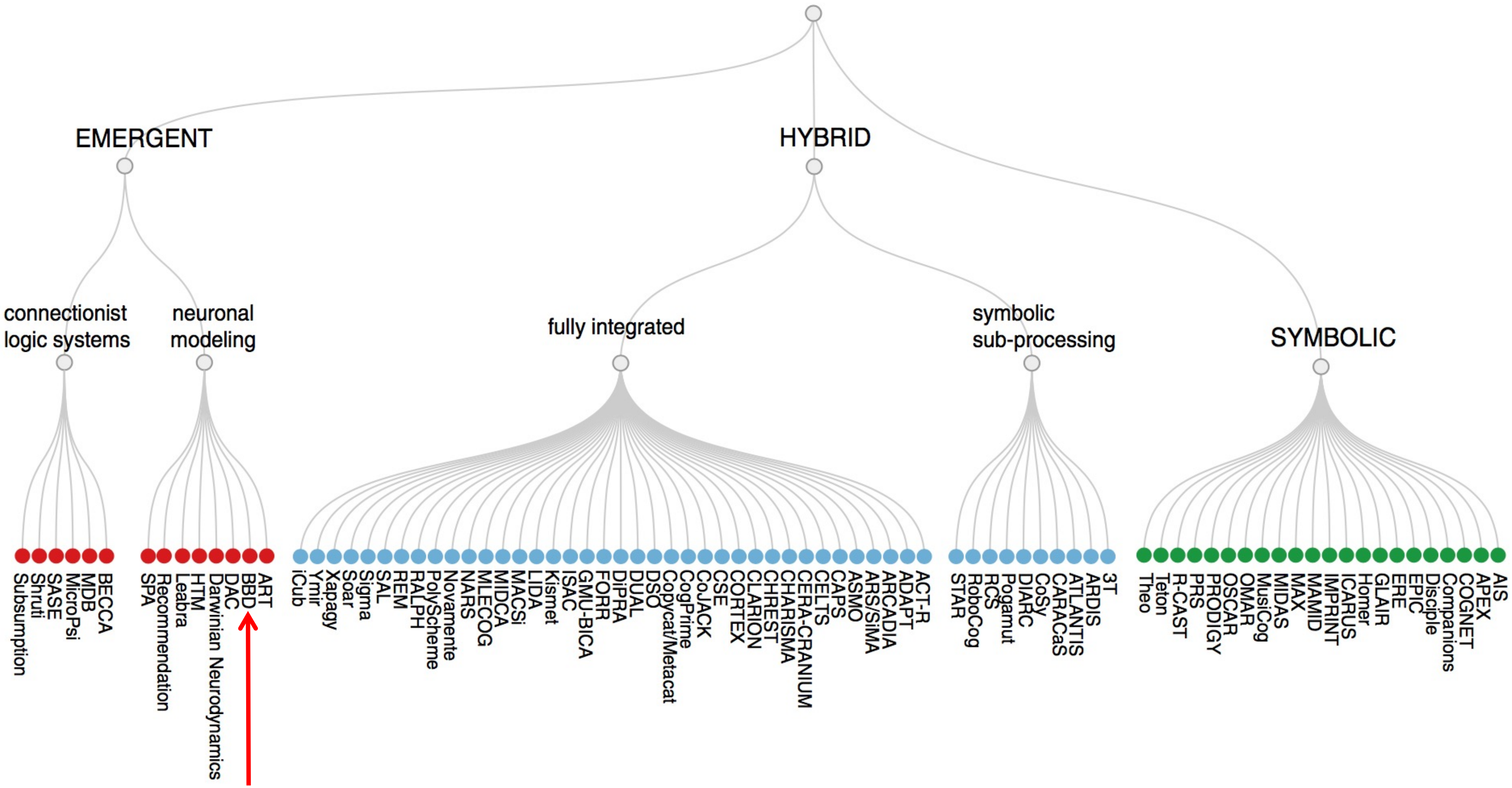


D. Vernon, C. von Hofsten, and L. Fadiga. A Roadmap for Cognitive Development in Humanoid Robots, Cognitive Systems Monographs (COSMOS), Vol. 11, Springer, 2010.

LIDA



(Franklin et al., 2014, 2016)

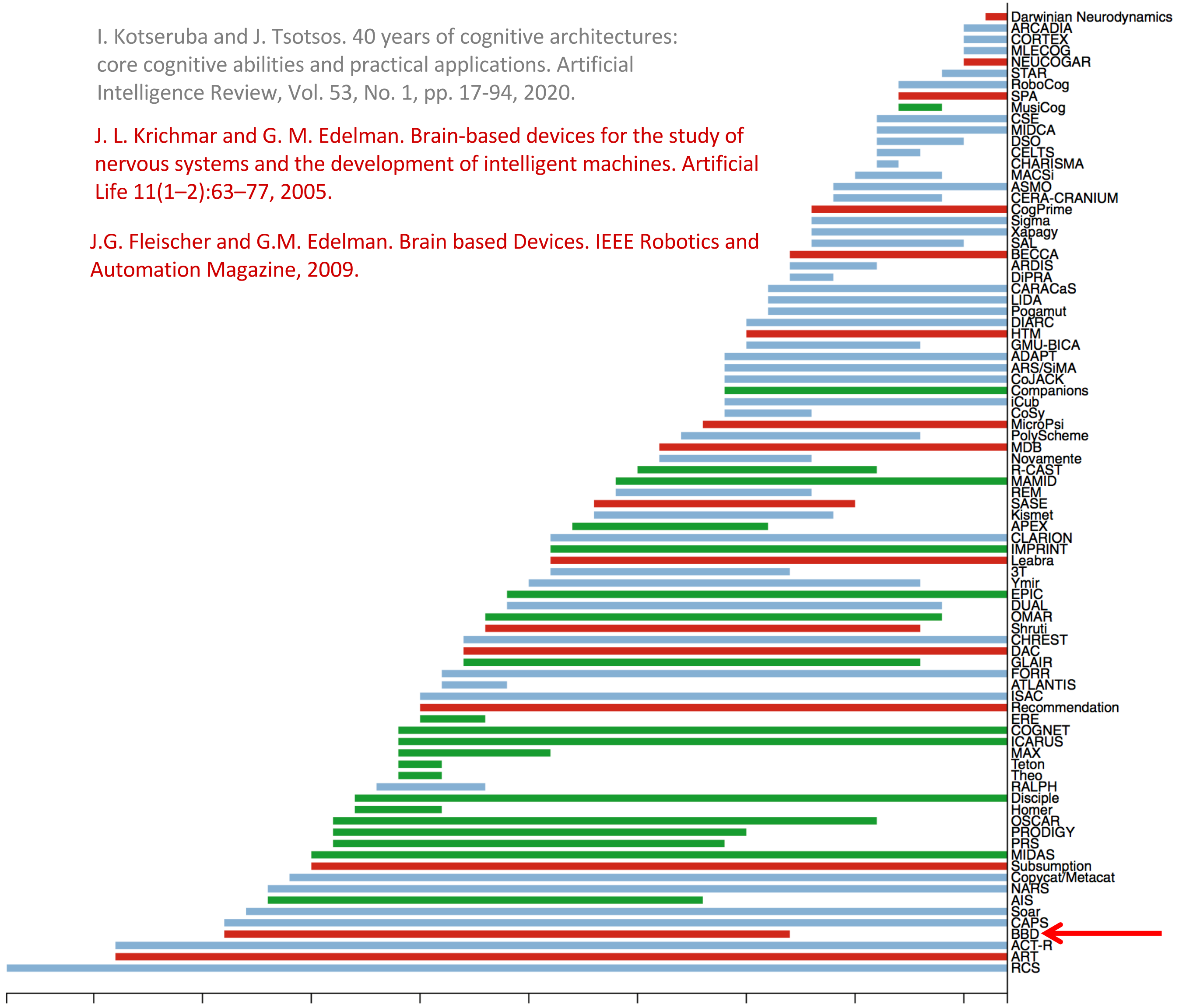


Brain-Based Devices (BBD)

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. L. Krichmar and G. M. Edelman. Brain-based devices for the study of nervous systems and the development of intelligent machines. Artificial Life 11(1–2):63–77, 2005.

J.G. Fleischer and G.M. Edelman. Brain based Devices. IEEE Robotics and Automation Magazine, 2009.



A.2 Emergent Cognitive Architectures 181

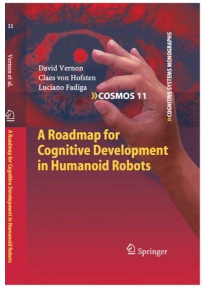
A.2.5 Darwin: Neuromimetic Robotic Brain-Based Devices

Kirchmar *et al.* [203, 204, 205, 206, 207, 334] have developed a series of robot platforms called Darwin to experiment with developmental agents. These systems are 'brain-based devices' (BBDs) which exploit a simulated nervous system that can develop spatial and episodic memory as well as recognition capabilities through autonomous experiential learning. As such, BDDs are a neuromimetic approach in the emergent paradigm that is most closely aligned with the enactive and the connectionist models. It differs from most connectionist approaches in that the architecture is much more strongly modelled on the structure and organization of the brain than are conventional artificial neural networks, *i.e.* they focus on the nervous system as a whole, its constituent parts, and their interaction, rather than on a neural implementation of some individual memory, control, or recognition function.

The principal neural mechanisms of the BDD approach are synaptic plasticity, a reward (or value) system, reentrant connectivity, dynamic synchronization of neuronal activity, and neuronal units with spatiotemporal response properties. Adaptive behaviour is achieved by the interaction of these neural mechanisms with sensorimotor correlations (or contingencies) which have been learned autonomously by active sensing and self-motion.

Darwin VIII is capable of discriminating reasonably simple visual targets (coloured geometric shapes) by associating it with an innately preferred auditory cue. Its simulated nervous system contains 28 neural areas, approximately 54,000 neuronal units, and approximately 1.7 million synaptic connections. The architecture comprises regions for vision (V1, V2, V4, IT), tracking (C), value or saliency (S), and audition (A). Gabor filtered images, with vertical, horizontal, and diagonal selectivity, and red-green colour filters with on-centre off-surround and off-centre on-surround receptive fields, are fed to V1. Sub-regions of V1 project topographically to V2 which in turn projects to V4. Both V2 and V4 have excitatory and inhibitory reentrant connections. V4 also has a non-topographical projection back to V2 as well as a non-topographical projection to IT, which itself has reentrant adaptive connections. IT also projects non-topographically back to V4. The tracking area (C) determines the gaze direction of Darwin VIII's camera based on excitatory projections from the auditory region A. This causes Darwin to orient toward a sound source. V4 also projects topographically to C causing Darwin VIII to centre its gaze on a visual object. Both IT and the value system S have adaptive connections to C which facilitates the learned target selection. Adaptation is effected using the Hebbian-like Bienenstock-Cooper-Munroe (BCM) rule [41]. From a behavioural perspective, Darwin VIII is conditioned to prefer one target over others by associating it with the innately preferred auditory cue and to demonstrate this preference by orienting towards the target.

Darwin IX can navigate and categorize textures using artificial whiskers based on a simulated neuroanatomy of the rat somatosensory system, comprising 17 areas, 1101 neuronal units, and approximately 8400 synaptic connections.



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

BBD

- Series of robot platforms focussed on developmental cognition

- Brain-based devices (BBDs)
  - Simulated nervous system
  - Develop spatial and episodic memory
  - Recognition capabilities
  - Autonomous experiential learning

- Neuromimetic: mimic the neural structure of the brain

- Differ from connectionist approaches: focus on
  - Nervous system as a whole
  - Constituent parts
  - Their interaction

BBD

- Principal neural mechanisms of a BDD
  - Synaptic plasticity
  - Reward (i.e. value) system
  - Reentrant connectivity
  - Dynamic synchronization of neuronal activity
  - Neuronal units with spatiotemporal response properties

- Adaptive behaviour
  - Interaction of these neural mechanisms with sensorimotor correlations (contingencies) which have been learned autonomously through active sensing and self-motion

BBD

Darwin VIII

- Discriminates simple visual targets (coloured geometric shapes)
- By associating them with an innately preferred auditory cue
- Its simulated nervous system contains
    - 28 neural areas
    - approximately 54,000 neuronal units
    - approximately 1.7 million synaptic connections

BBD

Darwin IX

- Navigates and categorizes textures using artificial whiskers
- Based on a simulated neuroanatomy of the rat somatosensory system
- Its simulated nervous system contains
  - 17 areas
  - 1101 neuronal units
  - approximately 8400 synaptic connections

BBD

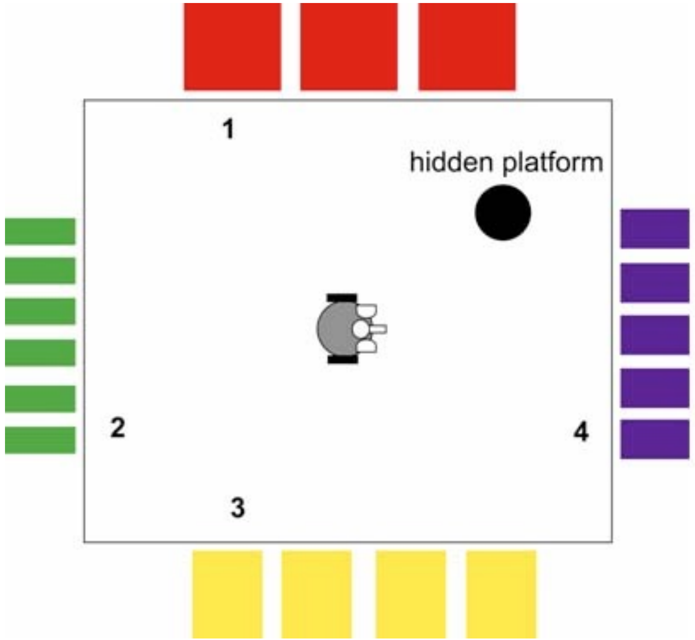
Darwin X

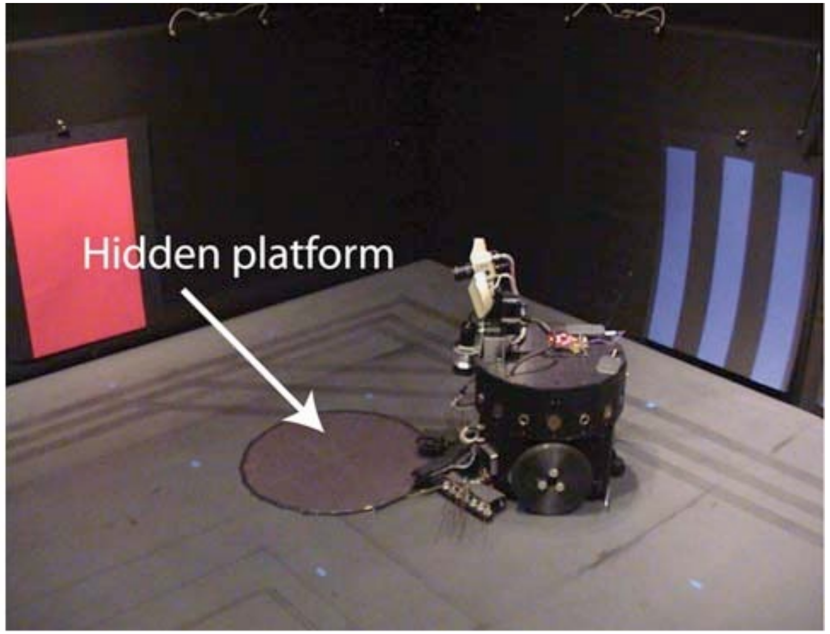
- Develops spatial and episodic memory based on a model of the hippocampus and surrounding regions

- Its simulated nervous system contains
  - 50 areas
  - 90,000 neuronal units
  - 1.4 million synaptic connections

- Systems
  - Visual system (object recognition, localization)
  - Head direction system
  - Hippocampal formation
  - Basal forebrain
  - Value/reward system based on dopaminegic function
  - Action selection system

Darwin X





J.G. Fleischer and G.M. Edelman (2009): Brain based Devices. IEEE Robotics and Automation Magazine.

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.1.1 (Soar)
- A.1.3 (ACT-R)
- A.2.2 (Global Workspace)
- A.2.4 (SASE)
- A.2.5 (DARWIN)
- A.3.5 (LIDA)

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/