A Survey of Artificial Cognitive Systems: Implications for the Autonomous Development of Mental Capabilities in Computational Agents

David Vernon, Giorgio Metta, and Giulio Sandini,

IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, VOL. 11, NO. 2, APRIL 2007

SungJae Yun

Outline

- Motivation
- Different paradigms of cognition
  - Cognitive models
  - Emergent approaches
  - Hybrid models
  - Relative strengths
- Cognitive architectures
- Comparison
- Conclusions

Motivation

- Over 60 years of artificial systems, many paradigms were suggested
- The term ‘Mental(Cognition)’ implies
  - Understand things might possibly be. and Determining how to act.
  - Entails robust behavior, perception, action, deliberation, and motivation
- Motivation is also important
  - Motivation drive perceptual attention, action selection, and system development, resulting in the long-term robust behavior
- From these perspective, writer overview of the various paradigms of cognition.

Different paradigms of cognition

- Two broad classes, Cognitivist vs Emergent systems
  - Cognitivist approach based on symbolic information processing representational systems
  - Emergent system approach, argue against the information processing view, on principles of self-organization

The Cognitivist *vs.* Emergent Paradigms of Cognition

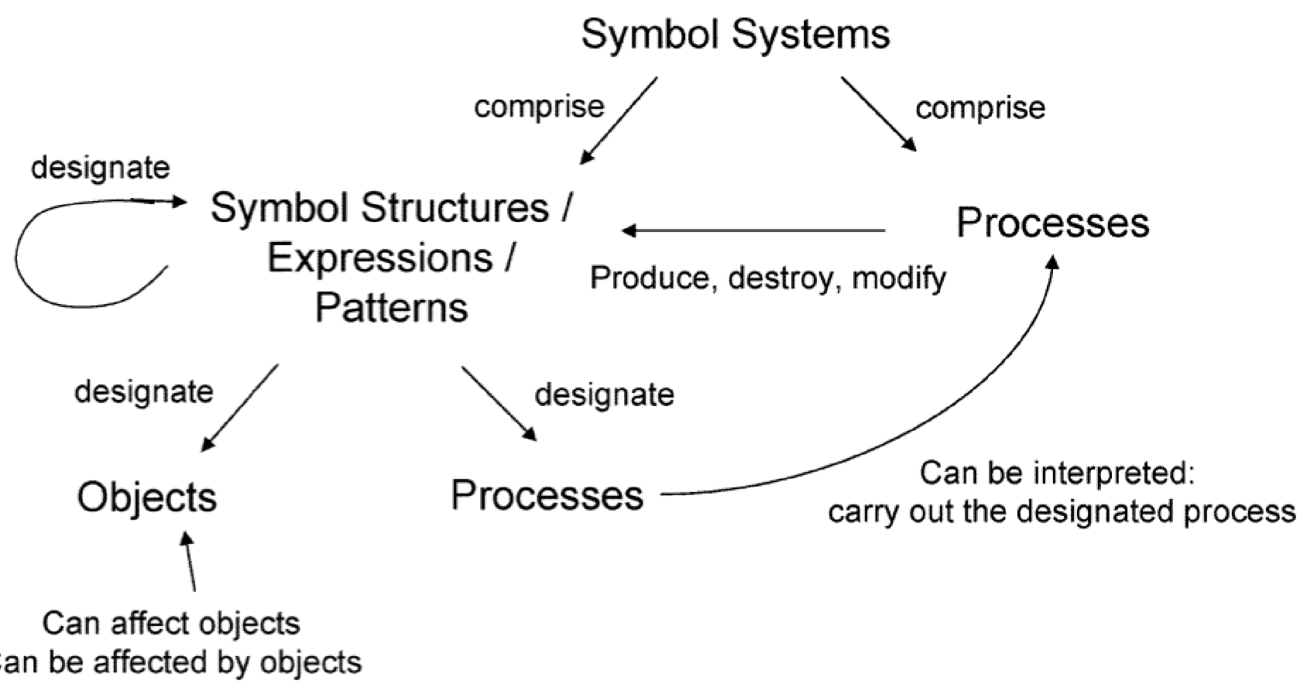
Characteristic	Cognitivist	Emergent
Computational Operation	Syntactic manipulation of symbols	Concurrent self-organization of a network
Representational Framework	Patterns of symbol tokens	Global system states
Semantic Grounding	Percept-symbol association	Skill construction
Temporal Constraints	Not entrained	Synchronous real-time entrainment
Inter-agent epistemology	Agent-independent	Agent-dependent
Embodiment	Not implied	Cognition implies embodiment
Perception	Abstract symbolic representations	Response to perturbation
Action	Causal consequence of symbol manipulation	Perturbation of the environment by the system
Anticipation	Procedural or probabilistic reasoning typically using *a priori* models	Self-effected traverse of perception-action state space
Adaptation	Learn new knowledge	Develop new dynamics
Motivation	Resolve impasse	Increase space of interaction
Relevance of Autonomy	Not necessarily implied	Cognition implies autonomy

Cognitivist models

- Origins in cybernetics
  - Create a science of mind, based on logic.
  - Undoubtedly predominant approach to cognition
- For cognitivist systems,
  - Cognition is manipulation of explicit symbolic representations
  - Perception is spatio-temporal representations of the external world from sensory data
  - Reasoning itself is symbolic
- Symbolic representations are descriptive product of a human designer
  - Can be directly accessed and interpreted by humans
  - This is also key limiting factor for the system
    - Programmer-dependent representations effectively bias the system

Cognitivist models

- Two hypothesis of Simon & Newell's Physical symbol system (1976)
  - The Physical Symbol System Hypothesis
    - System that exhibits general intelligence is a physical symbol system
  - Heuristic Search Hypothesis
    - Symbol systems solve problems by heuristic search.

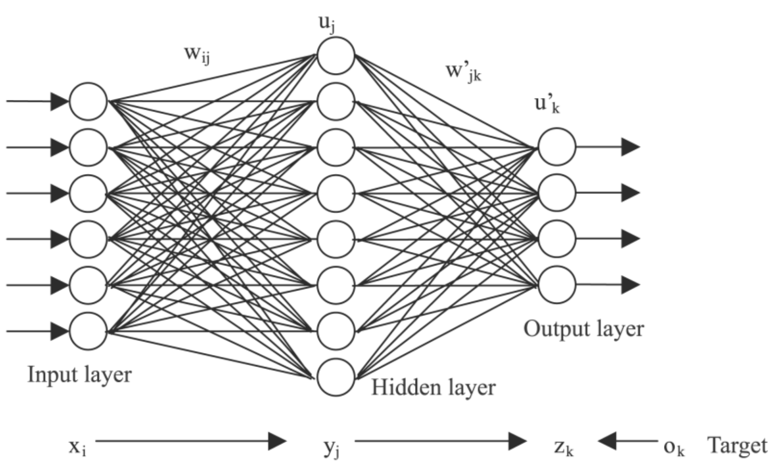


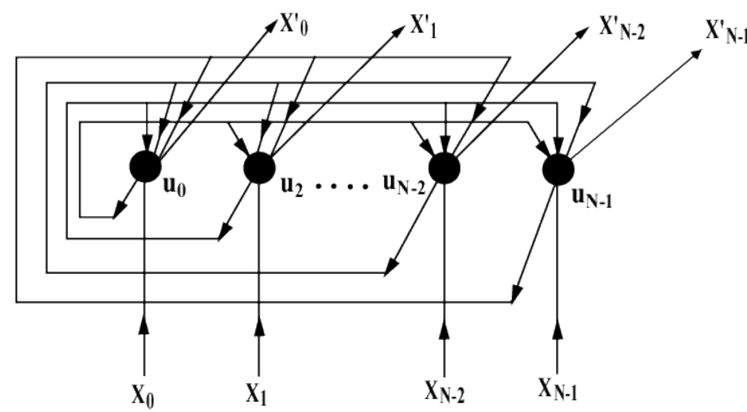
Emergent approaches

- Cognition in emergent system is,
  - Autonomous system becomes viable in environment
  - System Builds its own understanding as it develops
    - Through moderation of mutual system-environment interaction and co-determination, agent constructs its reality as a result of its operation in that world
- Perception is 'Acquisition of sensory data' in order to enable effective action
  - Cognitive learning is anticipative skill construction

Emergent approaches

- Connectionist models
  - Parallel processing of nonsymbolic distributed activation patterns
  - Using statistical properties, rather than logical rules
  - System’s connectivity plays important role
    - Meaning relates to the global state of the system
  - Perceptron, ART, SOMs, Hopfield net, Elman nets, Boltzman machines





Emergent approaches

- Dynamical systems models

- Enactive systems models
  - Concept of autopoiesis, enaction involves different degrees of autopoeisis and three orders of system
    - 1st order - Cellular entities,
      - Achieve a physical identity through structural coupling with their environment (i.e. instinctive behaviors)
    - 2nd order - Metacellular systems
      - Engage in structural coupling with environment, nervous system that enables the association of many internal states with the different interactions in which the organism is involved (i.e. ontogenetic behaviors)
    - 3rd order - Coupling between second-order systems

    (i.e. Communicative behaviors)

Hybrid models

- Combine aspects of the both systems
  - Use of explicit programmer-based knowledge + the development of active "animate" perceptual systems
- Perception-action behaviors become the focus
  - Flexibly interact with external world and interaction
- Recent results in biologically motivated system
  - Modeling brain function and cortical pathways and exploiting optical flow as its primary visual stimulus
  - Extent to the ability to learn a simple object affordance and use it to mimic the actions of human

Relative strength

- Cognitivist systems are actually most advanced. But, quite brittle.
  - Three problems of cognitivist systems
    - Symbol grounding, frame, combinatorial problems
  - Cognitivist models have difficulties in exhibit robust sensorimotor interactions in complex, noisy, dynamic environments
- Enactive and dynamical systems
  - Much less brittle. In theory
    - Emerge through mutual specification and co-development with the environment,
  - But, ability to build cognitive systems based on these principles is very limited at present.
- Hybrid approaches seems offer the best of both worlds:
  - Adaptability of emergent systems + Advanced starting point of cognitivist systems
  - Unclear how well one can combine what are ultimately highly antagonistic underlying philosophies

Cognitive architectures

Cognitivist	Emergent	Hybrid
Soar	AAR	HUMANOID
EPIC	Global Workspace	Cerebus
ACT-R	I-C SDAL	Cog: Theory of Mind
ICARUS	SASE	Kismet
ADAPT	DARWIN	

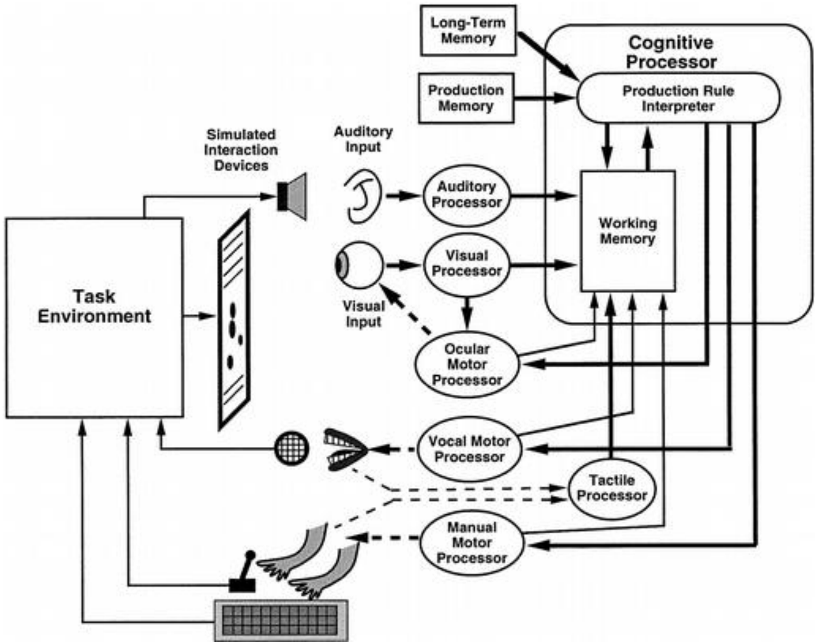
Soar Cognitive architecture

- Newell's candidate for a Unified Theory of Cognition
- Operates in a cyclic manner : Production cycle and Decision cycle
  1. Production fires if the contents of working memory matches
     - Fires alter the state of working memory and cause other productions to fire.
  2. If no more productions fire -> Decision cycle begins in which a single action from several possible actions is selected.
     - Selection based on action(stored) preferences.
- Universal subgoaling
  1. When decision cycle goes 'Impasse',
  2. Soar sets up an new state in a new problem space, subgoaling, with the goal of resolving the impasse
  3. After impasse resolved, Soar summarizes the process and create a new production rule.

     \- Resolving an impasse alters the system superstate. -> Learning in Soar architecture

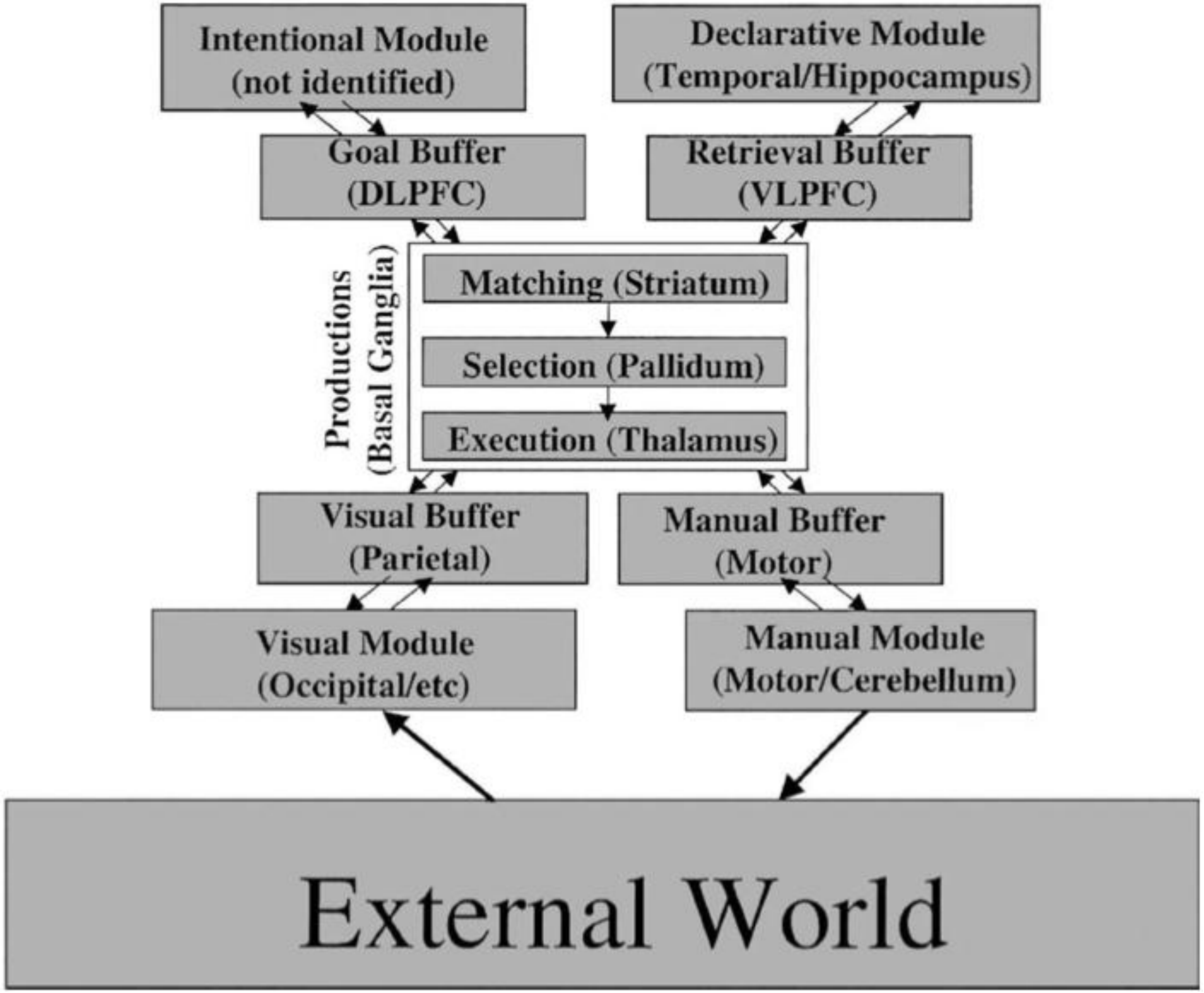
Executive Process Interactive Control (EPIC)

- All processors run in parallel.
- Perceptual processors simply model the temporal aspects of perception.
  - Only concerned with the time it takes for motor output to be produced after the cognitive processor has requested.
- Two phase movements: Preparation and Execution phase.
  - Preparation phase
    - Independent timing of the features that need to be prepared to effect the movement.
  - Execution phase
    - Concerned with the timing for the implementation of a movement
- Does not have any learning mechanism



ACT-R

- Adaptive Control of Thought–Rational



ICARUS

- Follows cognitivist architecture. such as ACT-R, Soar, and EPIC
  – Symbolic representations of knowledge
  – Use of pattern matching to select relevant elements
- Symbolic physical system hypothesis.
  – Mental states are always grounded in physical states
  – Symbolic operators always expand to actions that can be effected or executed.
- Distinguishes between concepts and skills, two different types of representation and memory
  – Conceptual memory encodes knowledge about general classes
  – Skill memory encodes knowledge about ways to act and achieve goals

ADAPT

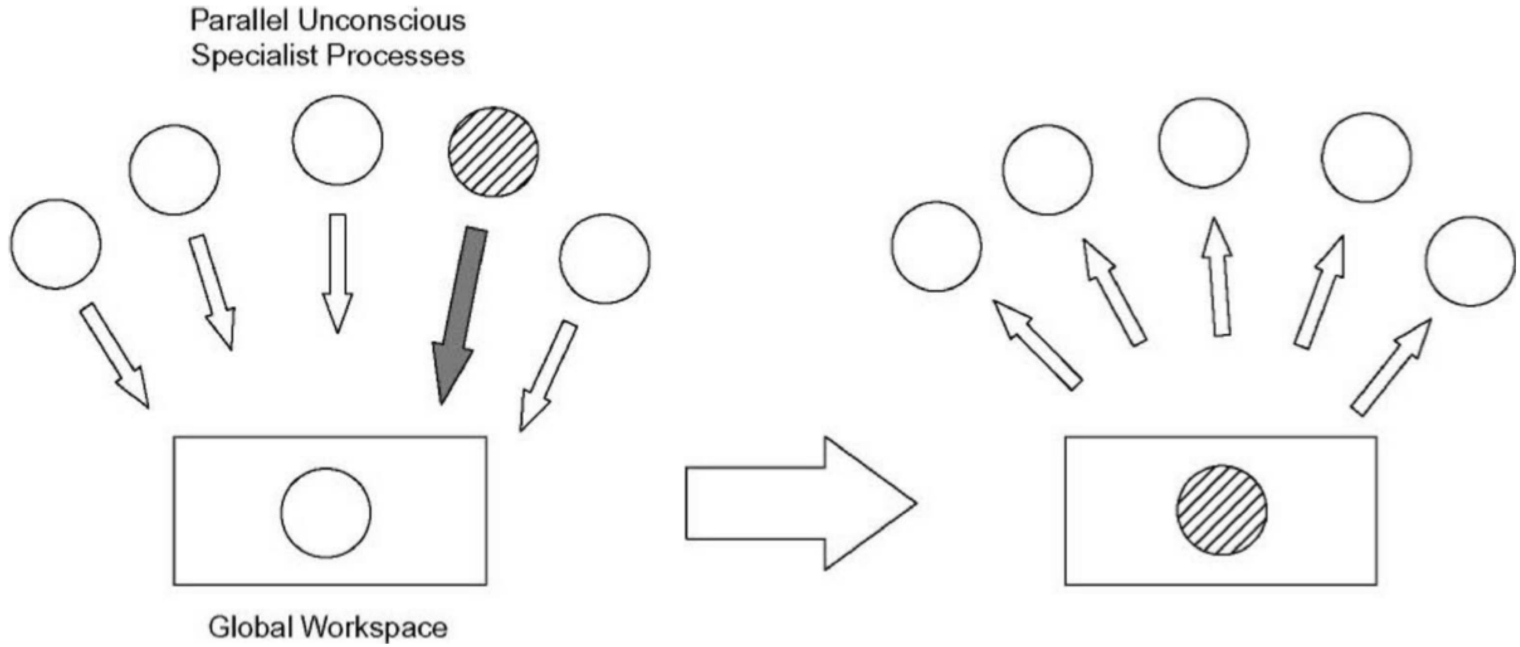
- Cognitivist architectures do not easily support robotics paradigms.
  – Since their “Sequential search and selection” paradigm
- ADAPT consists of,
  – Based on Soar’s architecture
  – ACT-R’s
    - Long-term declarative memory: Sensorimotor schemas to control perception and action are stored
  – EPIC’s
    - Perceptual processes fire in parallel
  – Low-level sensory data is placed in short-term working memory where it is processed by the cognitive mechanism
- Two distinctive parts of process: Task & Architecture
  – Architectural part : Restricted to only one goal or action at any one time, represented procedurally.
  – Task part has no restrictions, represented declaratively

Autonomous Agent Robotics (AARs)

- Architecture is based on interacting whole systems
  - Instead of a cognitivist system that is based on a decomposition into functional components (e.g., representation, concept formation, reasoning)

  1. Simple whole systems that act in simple circumstances,
  2. Layers of more sophisticated systems are added incrementally, each layer subsuming the layers beneath it
- Limitations also exist
  - Explosion of systems states results from the incremental integration of subsystems brings the difficulty

Global Workspace Cognitive Architecture



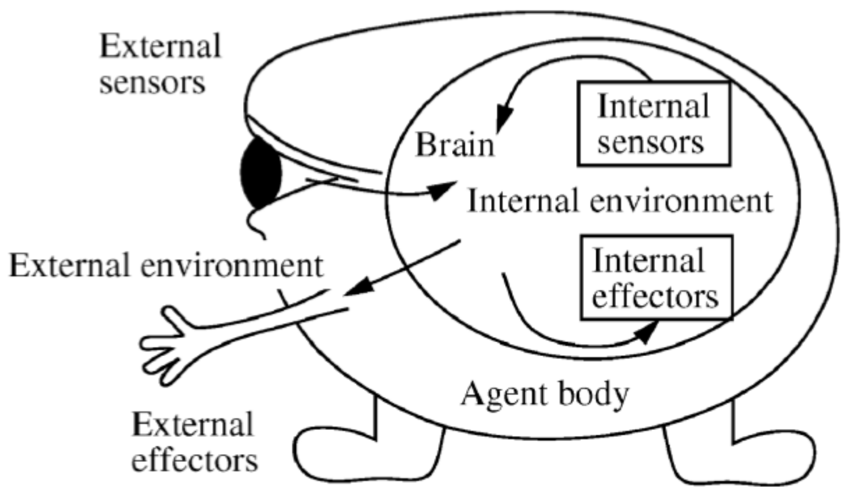
- Brain-inspired neural-level cognitive architecture
- Cognitive functions and planning are realized through internal simulation
- Winner takes all competition
  - Processes compete and co-operate for access to a global workspace.
  - Winner gain access to the global access and then broadcast information back to the competing specialist processes

Self-Directed Anticipative Learning

- Primary model for anticipative skill construction
  - Processes that both guide action and improve the capacity to guide action
  - Intelligence as a continuous management process
- Explicit norm signals : Low-order and high-order
  - Signals that system uses to differentiate an appropriate context performing an action
  - Conditions for the of the system's autonomy

Self-Aware Self-Effecting (SASE)

- Not specific to tasks
  - Capable of adapting and developing to learn both the tasks required of it and the manner in which to achieve the tasks
  - Emergent paradigms of cognition, especially the enactive approach
- Distinction between the sensors and effectors.
- Multiple pathway between the sensory and the motor mapping
  - One encapsulates innate behaviors while the other encapsulates learned behaviors
  - Pathways are mediated by a subsumption-based motor mapping which accords higher priority to the ontogenetically developed pathway.



Darwin

- Neuro-mimetic Robotic Brain-Based Devices
  - Exploit a simulated nervous system
  - Synaptic plasticity, reward (or value) system, reentrant connectivity, dynamic synchronization of neuronal activity, and neuronal units with spatiotemporal response properties.
  - Interaction of these mechanisms make adaptive behavior

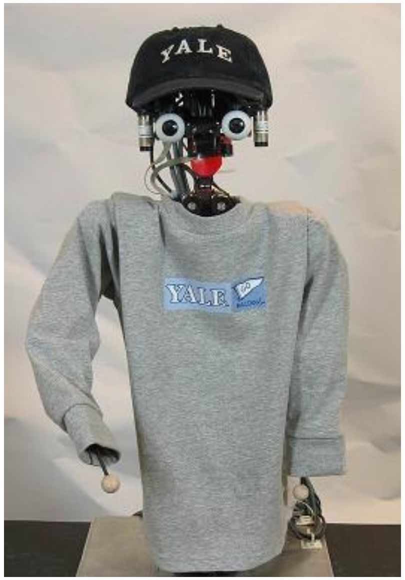
- Series of Darwin
  - Darwin VIII is capable of discriminating reasonably simple visual targets
  - Darwin VIII to center its gaze on a visual object
  - Darwin IX can navigate and categorize textures using artificial whiskers
  - Darwin X is capable of developing spatial and episodic memory based on a model of the hippocampus and surrounding regions

Humanoid Robot Cognitive Architecture

- Interacting parallel behavior-based components
  - Three-levels of each hierarchical perception, hierarchical task handling system
    - Perception subsystem comprises a three-level hierarchy with low, mid, and high-level perception modules
    - Task handling subsystem comprises a three-level hierarchy with task planning, task coordination, and task execution levels
  - Long-term memory subsystem based on a global knowledge database
  - Dialogue manager which mediates between perception and task planning
  - Execution supervisor, all levels of perception and task management

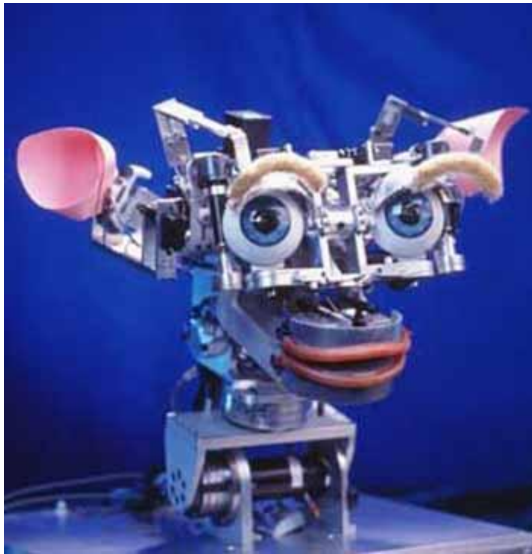
Cog : Theory of mind

- Upper-torso humanoid robot platform
- Learning from an observer using normal social signals and would be capable of expressing its internal state (emotions, desires, goals)
  - Interpretation of perceptual stimuli (auditory, and tactile) associated with self-propelled motion,
  - Interpretation of visual stimuli associated with eye-like shapes.
  - Both stimuli are feed in Theory of Mind module represents intentional knowledge or “epistemic mental states” of other agents.

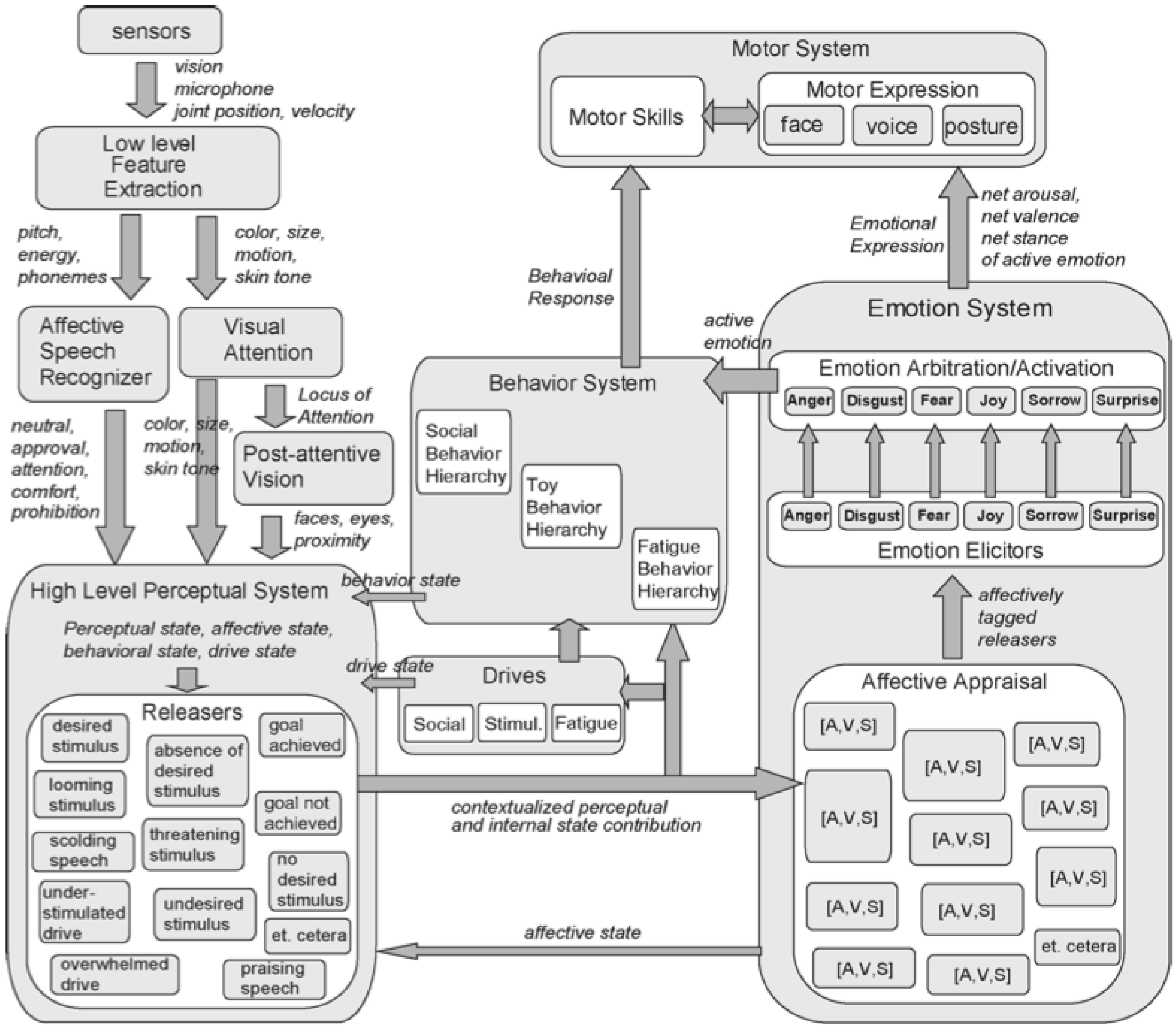


Kismet

- Designed to engage people in natural and expressive face-to-face interaction
  - Perceiving natural social cues and responding through gaze direction, facial expression, body posture, and vocal babbling.
  - Ultimate goal is to learn from people though social engagement,
- Two types of motivations: drives and emotions
  - Drives establish the top-level goals of the robot:
  - Emotions elicit specific behavioral responses and effect tending to cause the robot to come into contact with things that promote its “wellbeing”
- Five distinct modules
  - Perceptual system, Emotion system, Behavior system, Drive system, and Motor system



Kismet



sensors
vision
microphone
joint position, velocity
Low level Feature Extraction
pitch, energy, phonemes
color, size, motion, skin tone
Affective Speech Recognizer
Visual Attention
Locus of Attention
Post-attentive Vision
neutral, approval, attention, comfort, prohibition
color, size, motion, skin tone
faces, eyes, proximity
High Level Perceptual System
Perceptual state, affective state, behavioral state, drive state
Releasers
desired stimulus
absence of desired stimulus
goal achieved
looming stimulus
goal not achieved
scolding speech
threatening stimulus
no desired stimulus
under-stimulated drive
undesired stimulus
et. cetera
overwhelmed drive
praising speech
behavior state
drive state
Behavior System
Social Behavior Hierarchy
Toy Behavior Hierarchy
Fatigue Behavior Hierarchy
Drives
Social
Stimul.
Fatigue
Behavioal Response
Motor System
Motor Skills
Motor Expression
face
voice
posture
Emotional Expression
net arousal, net valence net stance of active emotion
active emotion
Emotion System
Emotion Arbitration/Activation
Anger
Disgust
Fear
Joy
Sorrow
Surprise
Emotion Elicitors
affectively tagged releasers
Affective Appraisal
[A,V,S]
et. cetera
contextualized perceptual and internal state contribution
affective state

Comparison

Architecture	Paradigm	Embodiment	Perception	Action	Anticipation	Adaptation	Motivation	Autonomy
Soar	C				+	+		
Epic	C		+	+	+			
ACT-R	C		+	+	+	+		
ICARUS	C		+	+	+	+		
ADAPT	C	×	×	×	+	+		
AAR	E	×	×	×			+	×
Global Workspace	E	+	+	+	×		×	×
I-C SDAL	E	+	+	+	+	+	×	×
SASE	E	×	×	×	+	×	×	×
Darwin	E	×	×	+		×	×	×
HUMANOID	H	×	×	×	×	+	+	
Cerebus	H	×	×	×	+	+		
Cog: Theory of Mind	H	×	×	×	+			
Kismet	H	×	×	×			×	

X
Characteristic is strongly addressed in the architecture,

+

Characteristic is weakly addressed

Space

Indicates that it is not addressed

Conclusions

- IMPLICATIONS FOR THE AUTONOMOUS DEVELOPMENT OF MENTAL CAPABILITIES IN COMPUTATIONAL SYSTEMS

1. Network of competing and cooperating distributed multifunctional subsystems
2. Developmental cognitive architecture must be capable of adaptation and self-modification
3. Cognitive systems are not only adaptive but also anticipatory and prospective
4. Embodied systems

- Additionally, for both cognitivist and emergent paradigms, development is dependent on the system’s phylogenetic configuration