Cognitive Architectures

What? Why? And How?

Punit Singh
Ph.D. Scholar
SILP Lab,
Center for Cognitive Computing,
IIIT Allahabad

What is Cognition?

- Cognition refers to a range of mental processes relating to the acquisition, storage, manipulation, and retrieval of information.
- Cognition fundamentally controls our thoughts and behaviors and these are regulated by discrete brain circuits which are underpinned by a number of neurotransmitter systems.
- Cognition is not a unitary concept and various cognitive functions, or cognitive ‘domains’, responsible for regulation of specific behaviours or actions have been identified.
- These functions are often convoluted, and operate synergistically making it challenging to measure distinct cognitive processes.

Cognitive Functions/Domains

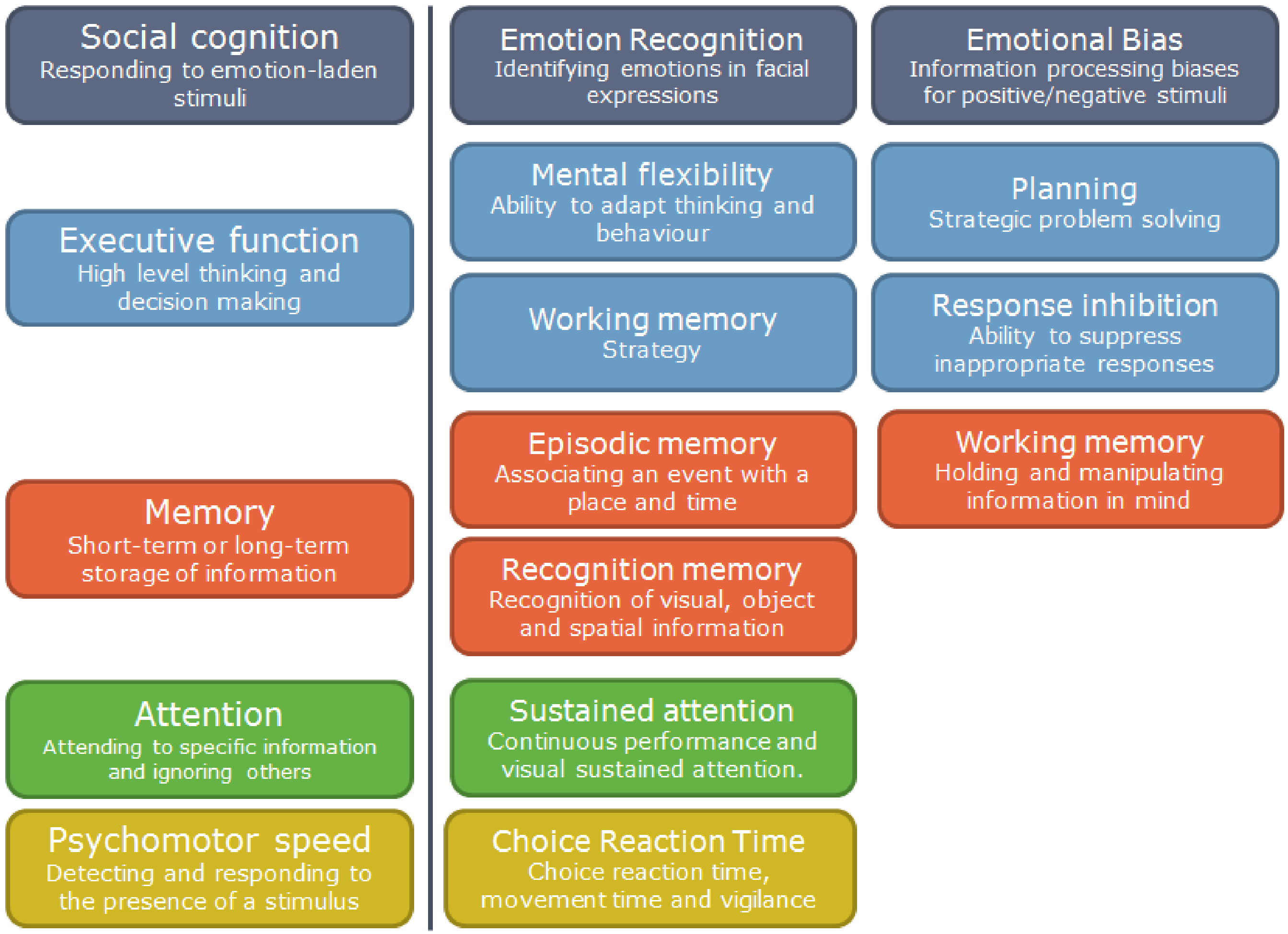

Figure 1. Domain specificity of cognition and examples of component cognitive processes underlying these mechanisms

What is a Cognitive Model?

- A cognitive model is an approximation to animal cognitive processes (predominantly human) for the purposes of comprehension and prediction.
- Cognitive models can be developed within or without a cognitive architecture, though the two are not always easily distinguishable.
- In contrast to cognitive architectures, cognitive models tend to be focused on a single cognitive phenomenon or process (e.g. list learning), how two or more processes interact (e.g. visual search and decision making), or to make behavioral predictions for a specific task or tool (e.g. how instituting a new software package will affect productivity).

What is a Cognitive Architecture?

- A cognitive architecture is a broadly-scoped, domain-generic computational cognitive model, capturing the essential structure and process of the mind, to be used for a broad, multiple-level, multiple-domain analysis of behaviour.
- Cognitive architectures tend to be focused on the structural properties of the modeled system, and help constrain the development of cognitive models within the architecture.
- An architecture includes those aspects of a system that are relatively invariant across time, domains, and individuals.
- It deals with componential processes of cognition in a structurally and mechanistically well defined way.

What is a Cognitive Architecture?

- In relation to understanding the human mind (i.e., cognitive science), a cognitive architecture provides a concrete framework for more detailed modeling of cognitive phenomena, through specifying essential structures, divisions of modules, relations between modules, and so on.
- Its function is to provide an essential framework to facilitate more detailed modeling and exploration of various components and processes of the mind.
- In relation to building intelligent systems, a cognitive architecture specifies the underlying infrastructure for intelligent systems, which includes a variety of capabilities, modules, and subsystems.
- A cognitive architecture carries also with it theories of cognition and understanding of intelligence gained from studying the human mind.

Why are Cognitive Architectures Important?

1. They shed new light on human cognition and therefore they are useful tools for advancing the understanding of cognition
2. They may (in part) serve as a foundation for understanding collective human behavior and social phenomena
3. They are “intelligent” systems that are cognitively realistic (relatively speaking) and therefore they are more human-like in many ways.

Multiple Levels of Explanations

Level	Object of analysis	Type of analysis	Model
1	inter-agent/collective processes	social/cultural	collections of agent models
2	agents	psychological	individual agent models
3	intra-agent processes	componential	modular construction of agent models
4	substrates	physiological	biological realization of modules

Capabilities
A well-defined architecture?

- Recognition & Categorization
- Decision Making & Choice
- Perception & Situation Assessment
- Prediction & Monitoring
- Problem Solving & Planning
- Reasoning & Belief Maintenance
- Execution & Action
- Interaction & Communication
- Remembering, Reflection, & Learning

Properties

Knowledge:

- Representation
- Organization
- Utilization
- Acquisition & Refinement

Evaluation Criteria for Cognitive Architectures?

- Generality, Versatility, & Taskability
- Rationality & Optimality
- Efficiency & Scalability
- Reactivity & Persistence
- Improvability
- Autonomy & Extended Operation

Three influences determine an agent's behavior

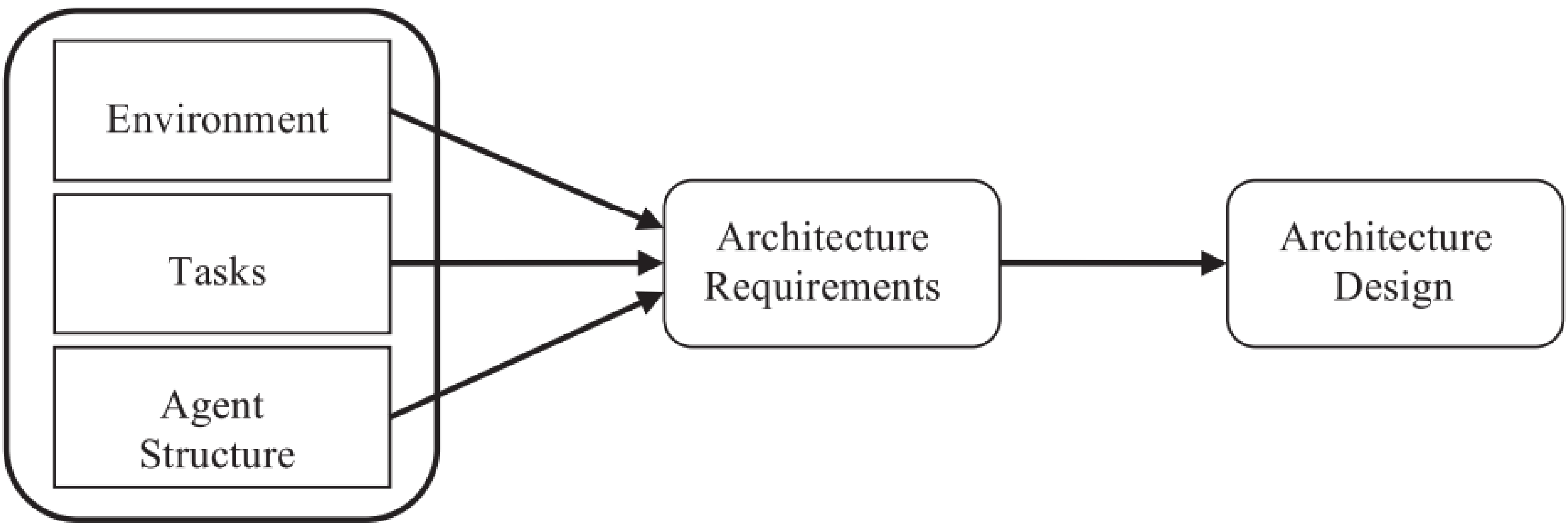

Figure 2. Influences on architecture design.

Characteristics of Environments, Tasks, and Agents

C1: The environment is complex, with diverse interacting objects.

- The world is large and complex.
- Agents can usefully interpret the environment as if it consists of independent objects, together with other materials that do not have object-like structure (e.g., air, water, and sand).
- There are many objects, and the objects interact with each other via physics.
- Objects have numerous and diverse properties.
- In order to comprehend and reason about these objects, an agent must be able to manipulate internal combinatorial representations of objects, their properties, and their relations. Because the environment is rich, there is much to be learned.

Characteristics of Environments, Tasks, and Agents

C2: The environment is dynamic.

- The agent's environment can change independent of the agent in such a way that the agent doesn't determine the state of the environment and the agent must respond to the dynamics of the world.
- Because the world can change while an agent is computing, the agent must respond relative to the dynamics of the environment (in real time).
- Moreover, the dynamics of the environment are complex enough so that the agent cannot always predict future states of the world in detail.

Characteristics of Environments, Tasks, and Agents

C3: Task-relevant regularities exist at multiple time scales.

- An environment, though it may be complex and dynamic, is not arbitrary.
- The laws of interaction that govern the environment are constant, often are predictable, and lead to recurrence and regularity that affect the agent's ability to achieve its goals.
- There are different regularities at different time scales, which makes it possible and useful to organize knowledge about tasks, actions, and the environment hierarchically.

Characteristics of Environments, Tasks, and Agents

C4: Other agents affect task performance.

- The agent is not alone, and must interact with other agents in pursuit of its goals.
- Other agents may help or hinder the agent's achievement of its tasks.
- The agent can communicate with other agents to share knowledge or to indicate intent.
- In addition, if some other agents are similar to it in structure and in capabilities (for example, have similar perception, action, and mental capabilities), the agent may be able to learn from them by observing the methods they use to solve problems.
- This characteristic is a special case of C1, C2, and C3, but it has enough of an effect on the structure of agents (both natural and artificial) to warrant its own item here.

Characteristics of Environments, Tasks, and Agents

C5: Tasks can be complex, diverse, and novel.

- A general, intelligent agent must be able to work on a diverse set of novel, complex tasks.
- Tasks can interact so that in some cases achieving one task aids in achieving another but in other cases achieving one task makes achieving another more difficult.
- Tasks can also vary in the time required to achieve them: some must be performed at a time scale close to that of relevant changes in the environment, but others require many orders of magnitude longer to complete.

Characteristics of Environments, Tasks, and Agents

C6: Agent-environment-task interactions are complex and limited.

- There may be many regularities in the environment, but they are relevant only if they can be detected and only if they influence the agent's ability to perform its tasks.
- Thus, an agent must have sufficient sensory capabilities that it can detect (possibly only through extensive learning) task-relevant regularities in the environment.
- An agent also must have mechanisms for acting in the environment in order to pursue a task.
- Although sensing and action modalities can be extensive, they are limited.
- The environment is partially observable, both from inherent physical limits in the sensors and from the size of the environment.
- Sensors are noisy, can be occluded by objects, and have limited range, making the agent's perception of its environment incomplete and uncertain.
- The agent's actions must obey the physical limitations of the environment. For example, actions usually take time to execute and are limited in extent.

Characteristics of Environments, Tasks, and Agents

C7: An agent's resources are limited.

- An agent has physical limits on its resources relative to the dynamics of the environment.
- Here, we focus on energy and computation resources. Energy resources are finite and are consumed by movement and computation, and thus must be replenished.
- Computation resources are limited so that an agent cannot perform arbitrary computation in the time it has available to respond to the dynamics of the environment.
- Thus, an agent has bounded rationality and cannot achieve perfect rationality (or universal intelligence) in sufficiently complex environments and tasks when it has a large body of knowledge.

Characteristics of Environments, Tasks, and Agents

C8: An agent's existence is long-term and continual.

- The agent is always present in its environment, and it must actively pursue tasks (such as self-protection) related to its survival.
- The agent may act to position itself so that the dynamics of the environment have little effect on it for extended times (e.g., it may hide in a protected area), but it has no guarantee that those efforts will be successful. Further, the agent has a long-term existence relative to its primitive interactions with its environment.
- Its activity extends indefinitely across multiple tasks, and possibly across multiple instances of the same task.

Architectural Requirements

R0: Fixed for all tasks.

- An individual agent adapts to its environment not through changes in its architecture but through changes in knowledge.
- The rationale for this requirement is that regularities exist (C3) at time scales that approach or exceed the life of the agent (C8), and these regularities are worth capturing in a fixed architecture, which then provides a stable platform for acquiring and using knowledge.
- There can be some changes in an architecture, such as occurs in humans through development, but those changes occur at time scales of years.
- Fixing the architecture's structure extends to disallowing manually tuned parameters.

Architectural Requirements

R1: Realize a symbol system.

- The consensus in AI and in cognitive science is that in order to achieve human-level behavior a system must support universal computation.
- Symbol systems provide both sufficient and necessary means for achieving universal computation; that is, a symbol system is capable of producing a response for every computable function.
- Symbol systems provide the ability to manipulate a description of some object in the world “in the head” without having to manipulate the object in the real world.
- Symbol structures also provide arbitrary composability to match the combinatoric complexity and regularity of the environment (C1, C3).
- Thus, structures encountered independently can be combined later to create novel structures never experienced together (C5).
- Symbol systems also allow us to accept instructions from another agent and then use those instructions later to influence behavior; thus, symbol systems provide additional flexibility and more generality, so that not everything must be programmed into a symbol system beforehand. In addition, symbols are required for communication that doesn’t cause the meaning to be directly experienced by the agent (C4).

Architectural Requirements

R2: Represent and effectively use modality-specific knowledge.

- Although pure symbol systems support universal computation, they rely on modality independent methods for representing and reasoning to achieve universality and complete composability. However, complete composability isn't always necessary.
- For many problems, modality-specific representations can support more efficient processing through regularities (C3) in sensory processing (C6). For example, some representations and associated processes for visual input have qualitatively different computational properties for image operations. Examples include rotation, inversion, and detecting and reasoning about spatial relations.
- For tasks (C5) where the agent has limited computational resources (C7), modality specific representations are necessary for achieving maximal efficiency; this is especially true of tasks that require real-time performance (C2).
- Modality-specific representations can also provide precision and accuracy for reasoning and motor control that is difficult to achieve with purely symbolic representations, especially when using task-independent transformations from sensory data to symbolic representations.

Architectural Requirements

R3: Represent, effectively use, and efficiently access large bodies of diverse knowledge.

- The agent must be able to represent and use large bodies of knowledge, because of the wealth of available knowledge that arises from the complexity of the environment (C1) and its associated regularities (C3), the variety of tasks the agent must pursue (C5), the agent's complex interaction with the environment (C6), and the agent's continual existence (C8).
- Not only are large bodies of knowledge available; the knowledge is diverse, including memories of experiences, facts of the world, skills, and knowledge about other agents (C4).
- As knowledge grows with experience, access to the knowledge must continue to be fast enough to be useful relative to the dynamics of the environment (C2).

Architectural Requirements

R4: Represent and effectively use knowledge with different levels of generality.

- The agent must represent and use general knowledge that takes advantage of environmental regularities (C3).
- The agent must also be sensitive to details of its current situation and be sensitive to its relationship to its tasks.
- These details are ubiquitous in complex (C1), dynamic (C2) environments in which an agent can have many tasks (C5).

Architectural Requirements

R5: Represent and effectively use diverse amounts of knowledge.

- An agent must be able to take advantage of whatever knowledge is available.
- For novel tasks and environments, its knowledge is limited.
- Even for familiar tasks and environments, its knowledge may be incomplete, inconsistent, or incorrect.
- If there is extensive knowledge available for a task, the agent must be able to represent it and to use it effectively.
- This is due in part to the combination of the fact that there are regularities in the environment worth knowing (C3), the complexity of an agent's limited sensing of its environment (C6), the complexity of its environment and tasks (C5), and limits on its computational resources (C7). Planning systems often fail this requirement.
- They often have a required and fixed set of input knowledge (the task operators and a declarative description of the goal).
- Without this knowledge, they cannot even attempt the problem. Further, if additional knowledge (such as knowledge about the likelihood of an operator leading to the goal) is available, the planner is often unable to use it to improve behavior.

Architectural Requirements

R6: Represent and effectively use beliefs independent of perception.

- The agent must be able to represent and reason about situations and beliefs that differ from its current perception.
- Perceptual information is insufficient because perception is limited (C6), the environment is dynamic (C2), and there are regularities in the environment worth remembering (C3) for task completion (C5).
- Thus, the agent must be able to maintain a history of situations as well as the ability to represent and reason about hypothetical situations, a necessary component of planning.
- Symbol systems provide the means for representing novel combinations of previously sensed information.
- Meeting this requirement makes it possible for an agent to make a decision based not just on its current situation but also on its memory of previous situations and its predictions of future situations.
- This capability is also necessary for perspective taking, in which an agent can reason about what other agents will do in different situations.

Architectural Requirements

R7: Represent and effectively use rich, hierarchical control knowledge.

- The agent must have a rich representation for control, because the actions it can perform are complex (C6).
- Because of the dynamics of the environment (C2) and the multiplicity of the tasks playing out at multiple time scales (C5), some actions may have to occur in rapid sequence whereas others may have to occur in parallel.
- To keep up with a rapidly changing environment (C2) with limited computational resources (C7), the agent must take advantage of the structure of regularities of the environment (C3), maximizing the generality of the knowledge it encodes because of the complexity and variability of the environment and the agent's tasks (C1, C5).
- In many cases, this means organizing knowledge about actions hierarchically. With such an organization, the agent can decompose some of its actions into sequences of simpler actions, using the context of higher-level actions to constrain the choices and thereby reducing the knowledge required to generate action.

Architectural Requirements

R8: Incorporate innate utilities.

- In view of the agent's long-term existence (C8), the accompanying regularities at long time scales (C3), and the agent's limited resources (C7), there are states of the world that either favor or impede the agent's ability to survive and pursue its long-term tasks (C5).
- The exact value of states for the agent depends on the details of the structure of the agent, the environment, and the agent's long-term tasks.
- Some of these may be contingent on interactions with other agents (C4).

Architectural Requirements

R9: Initiate, represent, and manage goals at multiple time scales.

- An agent must be able to initiate tasks on the basis of interaction with other agents (C4) or on the basis of its own situational analysis (R8).
- Many tasks cannot be achieved immediately; others are ongoing and must be continually pursued.
- Thus, an agent must maintain an internal representation of the task it is attempting to achieve so as to direct its behavior in service of the task (Wooldridge 2000).
- We call such an internal representation of a task a goal.
- Different tasks, and thus different goals, can have different temporal extents (C3), some lasting over significant portions of an agent's lifetime (C8).
- Goals can be complex, diverse, and novel (C5), and interactions among goals can require that the agent actively manage them by, for example, determining when it is appropriate to pursue a goal, interrupting pursuit of a goal, or even abandoning unachievable goals.

Architectural Requirements

R10: Access, represent, and effectively use meta-cognitive knowledge.

- It is sometimes necessary for an agent to represent and use knowledge about itself and knowledge about its own knowledge (meta-knowledge).
- An agent invariably faces novel tasks (C5) in which its task knowledge and/or its computational resources (C7) are not sufficient to determine the appropriate behavior, owing to environmental complexity (C1), but in which there are regularities of which the agent can take advantage (C3).
- In these situations, an intelligent agent can detect its lack of task knowledge and then use meta-knowledge to acquire new task knowledge. An agent can use other types of meta-cognitive knowledge to set its own goals and to direct future behavior in preparation for tasks, events, and situations that it expects to arise.
- This is done in response to the characteristics listed above and in response to the fact that the agent exists beyond a single task or problem (C8).
- The exact range of necessary meta-cognitive knowledge isn't clear — some appears to be necessary, but complete meta-cognitive knowledge isn't required (at least, not in humans).
- We humans don't always know exactly what we know, and often we discover what we know only when put in a situation where that knowledge is useful.

Architectural Requirements

R11: Support a spectrum of bounded and unbounded deliberation.

- At one extreme, in tasks with time constraints close to those of the dynamics of the environment (C2), the agent must respond using bounded computation (C5).
- Because of inherent limits to its computational resources (C7), it cannot reason or plan from first principles for all tasks.
- At the most primitive level, the absolute time to respond is bounded by the environmental dynamics for some subclass of responses.
- Reactive behavior is possible if the agent's knowledge of the environment and other agents is complete, correct, and encoded for bounded access below the level of dynamics of the environment.
- However, because of the complexity of the environment (C1), the diversity of tasks (C5), and the limitations on environmental interaction (C6), that generally isn't possible.
- Moreover, at the other extreme, when there are sufficient computational resources available relative to the dynamics of the environment and task, the agent should have the ability to compose novel responses using its knowledge to take advantage of regularities in the tasks and environment (C3).
- This composition is the basis for planning.
- It takes time, but it allows the agent to integrate its diverse and potentially large bodies of knowledge for novel situations (R1 – R10).
- In between these two extremes, the agent must use its knowledge of the situation to balance between reaction and deliberation.

Architectural Requirements

R12: Support diverse, comprehensive learning.

- An agent with long-term existence (C8) requires different learning mechanisms when exposed to diverse environments (C1) and tasks (C5) involving complex interactions (C6).
- Learning takes advantage of regularities (C3), some of which can be extracted from a single situation in which all of the information is available at the same time, whereas in other cases the information may be spread across time.
- Although general learning mechanisms exist, they are invariably biased toward specific regularities and types of knowledge that are available to the agent in different ways and often at different time scales.
- Moreover, a general cognitive architecture should be able to learn all the types of task-specific knowledge it represents and uses — a property we call the learning completeness principle.
- A significant component of our research is exploring what types of regularities are available to an agent and what types of knowledge and associated learning mechanisms are required to extract and later retrieve those regularities and achieve complete learning.

Architectural Requirements

R13: Support incremental, online learning.

- An agent with long-term existence (C8) in a complex active environment (C1, C2) with regularities (C3) must learn and modify its knowledge base so as to take advantage of environmental regularities (C3) when they are available.
- Once the experience has happened, it is gone. Only the information that the agent itself stores while it is behaving is available to guide its future behavior.
- This is not to suggest that an agent cannot recall situations and perform additional analysis at some future time (R6); however, some primitive learning mechanism must store the experience for that more deliberative future learning.
- Moreover, the mechanisms for storing and retrieving those experiences must operate in real time even as more and more experiences are captured. Incremental learning incorporates experiences when they are experienced.

Connections between environment, task, and agent characteristics and architectural requirements.

	C1 (complex environment)	C2 (dynamic environment)	C3 (task regularities)	C4 (social environment)	C5 (complex tasks)	C6 (limited interaction)	C7 (limited resources)	C8 (long-term existence)
R0 (fixed structure)			X					X
R1 (symbol system)	X		X	X	X			
R2 (modality-specific knowledge)		X	X		X	X	X	
R3 (large bodies of knowledge)	X	X	X	X	X	X		X
R4 (levels of generality)	X	X	X		X			
R5 (amount of knowledge)			X		X	X	X	
R6 (non-perceptual representations)		X	X		X	X		
R7 (rich action representations)	X	X	X		X	X	X	
R8 (innate utilities)			X	X	X		X	X
R9 (goals across multiple time scales)			X	X	X			X
R10 (meta-cognitive knowledge)	X		X		X		X	X
R11 (spectrum of deliberation)	X	X	X		X	X	X	
R12 (comprehensive learning)	X		X		X	X		X
R13 (incremental learning)	X	X	X				X	X

Open Issues

- Episodic Memory & Reflective processes
- Natural Language
- Emotions
- Enhanced learning
- And many more…

Resources and References

- http://www.cambridgecognition.com/blog/entry/what-is-cognition
- https://en.wikipedia.org/wiki/Cognitive_model
- Sun, R. (2007). The importance of cognitive architectures: An analysis based on CLARION. Journal of Experimental & Theoretical Artificial Intelligence, 19(2), 159-193.
- Laird, J. E. (2012). The Soar cognitive architecture. MIT press.