Lecture 26

Cognitive Architecture-I
Cognitive Architecture

A cognitive architecture specifies the *infrastructure* for an intelligent system that remains constant across different domains and knowledge bases. This infrastructure includes a commitment to

- formalisms for representing knowledge,
- memories for storing the domain content,
- performance processes that utilize this knowledge, and
- learning mechanisms that acquire it.

Research aims

- To support the same broad capabilities as humans exhibit, and usually attempts to remain at least qualitatively consistent with established psychological findings
An artificial computational process which mimics certain acts of cognition (Langley 2006)

- Focuses on cognition as a whole not some aspect of it like learning or problem solving
- The concept of modularity is central. Mostly fixed, but some allows growth and modification over time.
- Usually with learning capability
- Concurrency is normally required.
  Three kind of processes all running in parallel:
  - Reactive process: equivalent to things that we are doing without thinking,
  - Deliberate process: core to reasoning and planning
  - Reflective process: change the mental state of the robot

→ Cognitive system:
  adaptive, anticipatory and purposive goal-directed behavior
Cognitive Architecture

- **Additional features for socially interactive systems** (Scassellati 2001):
  - acts with the same reaction time to the modeled cognitive system
  - can operate in the same complex, noisy, and cluttered environment people inhabit such that it can recognize the social cues and express itself with such cues
  - Robust behavior and reliability

- **Approaches:**
  - Cognitivist approach:
    - based on symbolic information processing representational systems
    - Cognition as “symbolic, rational, encapsulated, structured, and algorithmic,“
  - Emergent systems approach
    - based on principles of self-organization
    - Cognition as emergent, self-organizing, and dynamical
  - Hybrid approach incorporating aspects of these two (Vernon 2007, Clark 2001)
COGNITIVE ARCHITECTURES
reviewed in this paper

<table>
<thead>
<tr>
<th>Cognitivist</th>
<th>Emergent</th>
<th>Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soar</td>
<td>AAR</td>
<td>HUMANOID</td>
</tr>
<tr>
<td>EPIC</td>
<td>Global Workspace</td>
<td>Cerebus</td>
</tr>
<tr>
<td>ACT-R</td>
<td>I-C SDAL</td>
<td>Cog: Theory of Mind</td>
</tr>
<tr>
<td>ICARUS</td>
<td>SASE</td>
<td>Kismet</td>
</tr>
<tr>
<td>ADAPT</td>
<td>DARWIN</td>
<td></td>
</tr>
</tbody>
</table>
Cognitivism

- The focus is on the aspects of cognitions, constant over time and relatively independent of the task.

- Cognitive model:
  - The combination of a given cognitive architecture and a particular knowledge set
  - The cognitive architecture:
    - defines the manner in which a cognitive agent manages the resources at its disposal, where resources are the computational system in which the physical symbol system is realized.
    - specifies the formalisms for knowledge representations and the memory used to store them, the processes that act upon that knowledge, and the learning mechanisms that acquire it

- ISAC (Intelligent Soft Arm Control, a parallel distributed cognitive control system for a humanoid robot) (Kawamura 2004), Soar architecture (Rosenbloom 1993), ACT-R (Adaptive Control of Thoughts-Rational) architecture (Anderson 1996), and Society of Mind (Minsky 1986)
Cognitivism

- ACT-R System (like a programming language)
  - To reproduce the theories about human cognition with assumption that human knowledge can be divided into two irreducible kinds of representations: declarative and procedural

- Two types of modules
  - Perceptual-motor modules
    - which take care of the interface with the real world.
    - The most well-developed perceptual-motor modules are the visual and the manual modules.
  - Memory modules: two kinds of memory modules in ACT-R:
    - Declarative memory: consisting of facts such as 2+2=4
    - Procedural memory: made of productions
      - Productions represent knowledge about how we do things, e.g., how to perform addition.
Applications:
- Learning and memory
- High level cognition
- Problem solving
- Decision making
- Natural language, including syntactic parsing, semantic processing and language generation
- Perception and Attention

Cognitivism
ACT-R introduced an architecture which can imitate human behavior, but modeling in this framework is a hard task.
- It needs both knowledge of modeling cognitive process and programming.
- The strength of this method is in its general approach to modeling of brain and its relevance with reality.

Future works
- Enhancing the cognitive models through expanded results and cognitively plausible behaviors and reasoning mechanisms
- Adding learning capabilities to the models to acquire new knowledge and skills through interaction with humans and while performing tasks
Emergent Approaches

- A very different view of cognition from cognitivism:
  Cognition is the process by which an autonomous system becomes viable and effective in its environment *through a process of self-organization* through which the system is continually *re-constituting* itself in real-time to maintain its operational identity through moderation of mutual system-environment interaction and co-determination.

  - Co-determination: the cognitive agent is specified by its environment and at the same time the cognitive process determines what is real or meaningful for the agent (Granlund 1999, Vernon 2007).
  - Emergent approaches are intrinsically embodied and the physical experiences of the robot play a central role in its cognition.
  - Architecture: not functionally modular, but parallel, real-time, and distributed

- Assertions:
  - The primary model for cognitive learning is anticipative skill construction rather than knowledge acquisition
  - Processes that both guide action and improve the capacity to guide action *while doing so*, are taken to be the root capacity for all intelligent systems.
Emergent Approaches

- Major subcategories: *connectionist, dynamical models and enactive system models*
  
  - Connectionist systems with simple abstractions of the brain, rely on parallel processing of non-symbolic distributed activation patterns using statistical properties, rather than logical rules, to process information and achieve effective behaviors (Medler 1998).
    - Local aspects of the problem, rather than giving insight into how a complete system might be organized
  
  - Dynamical system theory: used to complement classical approaches in AI (Reiter 2001) and deployed to model natural and artificial cognitive systems (Gelder & Port 1995).
    - An open dissipative nonlinear non-equilibrium system with relatively low order is used to model the system behavior.
    - The ability to characterize the behavior of a high-dimensional system with a low-dimensional model, is one of the features that distinguishes dynamical systems from connectionist systems.
Dynamical Models

- The core is the application of the mathematical tools of dynamics to the study of cognition.
  - A cognitive dynamical system:
    - The components of the system must be related and interact with one another.
    - Any change in one component or aspect of the system must be dependent on the states of the other components: “they must be interactive and self contained.”
    - “Cognition is nonsymbolic, nonrepresentational and all mental activity is emergent, situated, historical, and embodied.”
  - Cognitive system:
    - Not a discrete sequential manipulator of static representational structures, rather a structure of mutually and simultaneously influencing change in real time of ongoing change of environment, the body and the nervous system.
Enactive System Models

- Goal: the complete treatment of the nature and emergence of autonomous, cognitive, social systems
  - Based on the concept of autopoiesis – literally self-production – whereby a system emerges as a coherent systemic entity, distinct from its environment, as a consequence of processes of self-organization.
  - Enaction involves different degrees of autopoiesis and three orders of system
    - First-order autopoietic systems correspond to cellular entities that achieve a physical identity through structural coupling with their environment.
    - Second-order systems are metacellular systems that engage in structural coupling with their environment, this time through a nervous system that enables the association of many internal states with the different interactions in which the organism is involved and has processes of self-development.
    - Third-order systems exhibit coupling between second-order (i.e., cognitive) systems, i.e., between distinct cognitive agents.
Comparison

Cognitivist:

- Two central ideas: the *dualism* that speculates the logical separation of mind and body and the *functionalism* that speculates that cognitive mechanisms are independent of the physical platform (Freeman 1999)
- The construction is top-down.
- A symbolic structure is crucial for mimicking human behavior.
- Independent of their physical platform, general-purpose systems of cognition operating as a brain-in-the-box
- The adaptation is achieved by gaining new knowledge.

Emergent approach:

- Reflect or recognize the morphology of the physical body in which it is embedded and of which it is an intrinsic part
- It is bottom-up.
- Dynamical and organizational properties of the systems, e.g. a connectionist approach is sufficient for such a task of mimicking human behavior.
- Emergent systems are basically embodied metaphors of their environment, the physical instantiation plays a direct fundamental part in the cognitive process.
- Adapts by changing its structure to attain new dynamics.
## Comparison

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Cognitivist</th>
<th>Emergent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Construction</td>
<td>Top-Down</td>
<td>Bottom-Up</td>
</tr>
<tr>
<td>Representational Framework</td>
<td>Symbolic</td>
<td>Connectionist</td>
</tr>
<tr>
<td>Embodiment</td>
<td>Not implied</td>
<td>Cognition implies embodiment</td>
</tr>
<tr>
<td>Adaptation</td>
<td>Gaining new knowledge</td>
<td>Develop new structure</td>
</tr>
<tr>
<td>Perception</td>
<td>Abstract symbolic representations</td>
<td>Response to perturbation</td>
</tr>
<tr>
<td>Relevance of Autonomy</td>
<td>Not necessarily implied</td>
<td>Cognition implies autonomy</td>
</tr>
</tbody>
</table>
Hybrid Structures

- Combine aspects of cognitivist and emergent approaches, but against the uses of explicit designer knowledge in an AI system.
  
  - Their goal is to develop an active ‘animate’ perceptual system in which the behaviors themselves become the focus, rather than the perceptual abstractions.
  
  - The adaptability of emergent systems and the advanced starting point of cognitivist systems
    - still use representational invariance but the difference is these representations should only be constructed by the system itself as it interacts with and explores the world rather than through a priori specification

Ex) Combining a NN-based perception-action component and a symbolic component

- The Polyscheme model (Cassimatis 2002) and the Cog project (Brooks 1999, Scassellati 2001)
Polyscheme Model

- To achieve human-level intelligence by integrating multiple methods of representation (Cassimatis 2002)
  - based on Specialists which are modules based on a particular representation
    - Each specialist executes each of the basic operations of the procedural substrate with some policies for guiding the focus of attention.

- The assumptions:
  - Most high-order reasoning and problem solving algorithms can be implemented using the same set of basic computational operations - forward inference, sub-goaling, grounding, representing alternate worlds and identity matching
  - Each basic operation can be implemented using multiple representations, including those arising from perceptual processes and cognition about a basic set of relations (involving times, space, events, identity, causality and belief).
Polyscheme Model

- **Powerful level of integration**
  - Since all high-level reasoning and planning algorithms are implemented by a focus of attention that integrates all lower-level representations and perceptual and motor processes, *every single step of reasoning and planning* is guided by multiple representations and is informed by new perceptual information.
    - Changing environment and noisy sensors can quickly invalidate plans and inferences.

- **Robonaut, NASA’s Humanoid robot**
  - ACT-R to create three different spatial representations for object identification and combined them with Polyscheme and its multiple representation architecture.
  - Encapsulate representations in specialists to be integrated into inferences that agents make about a situation or possible situations (Sofge 2004).
Socially intelligent robots need a theory of mind to discover the hidden properties of their environment, e.g. the intent of the instructor rather than explicit actions.

Humans are not general purpose machines and they do not perform all the tasks equally.

Some neglected part of human intelligence may be substantial to the ability of human to autonomously learn, and assimilate immense numbers of skills:

- Development: forms the framework by which humans successfully acquire increasingly more complex skills and competencies.
- Embodiment and physical coupling: allow humans to use the world itself as a tool for organizing and manipulating knowledge.
- Integration: allows humans to maximize the efficacy and accuracy of complementary sensory and motor systems.
- Social interaction: allows humans to exploit other humans for assistance, teaching, and knowledge.
The Cog Project

- The Robot considering all the neglected essence of human intelligence

  - Development
    - Building systems developmentally facilitates learning both by providing a structured decomposition of skills and by gradually increasing the complexity of the task to match the competency of the system.

  - Physical embodiment
    - The direct physical coupling between action and perception reduces the need for an intermediary representation. For an embodied system, internal representations can be ultimately grounded in sensory-motor interactions with the world (Lakoff, 1987).

  - Integration of multiple sensory and motor systems
    - The computational control is a heterogeneous network of many different processor types operating at different levels in the control hierarchy.

- Social Interaction
  - Building social skills into an artificial intelligence provides not only a natural means of human-machine interaction but also a mechanism for bootstrapping more complex behavior.
  - A motivational system works to maintain robot drives within homeostatic bounds and motivates the robot to learn behaviors that satiate them.
A Theory of Mind

- The ability to recognize what another person can see, the ability to know that another person maintains a false belief, and the ability to recognize that another person likes games that differ from those that he/she enjoys.

- Central to what defines human interactions
  - Normal social interactions depend upon the recognition of other points of view, the understanding of other mental states, and the recognition of complex non-verbal signals of attention and emotional state.

- Leslie’s Model of Theory of Mind: The world is naturally decomposed into three classes of events.
  - One class for mechanical agency: The Theory of Body module (ToBY) explained events by the rules of mechanics.
  - One for actional agency: system 1 of the Theory of Mind module (ToMM-1) explaining events in terms of the intent and goals of agents, i.e., their actions.
  - One for attitudinal agency: system 2 of the Theory of Mind module (ToMM-2) which explains events in terms of the attitudes and beliefs of agents.
Baron-Cohen’s Model of Theory of Mind: The set of precursors to a theory of mind can be decomposed into four distinct modules.

- The first module, called the intentionality detector (ID), interprets self-propelled motion of stimuli in terms of the primitive volitional mental states of goal and desire and produces dyadic representations that describe the basic movements of approach and avoidance, such as “he wants the food” or “she wants to go over there.”

- The second module, called the eye direction detector (EDD), processes visual stimuli that are eye-like to determine the direction of gaze, such as “agent sees me” and “agent looking-at not-me”.

- The third module, the shared attention mechanism (SAM), takes the dyadic representations from ID and EDD and produces triadic representations of the form “John sees (I see the girl)”.

- The fourth module, the theory of mind mechanism (ToMM), provides a way of representing epistemic mental states in other agents and a mechanism for tying together our knowledge of mental states into a coherent whole as a usable theory, such as “John thinks (Elvis is alive).”
An Embodied Theory of Mind

- A hybrid architecture called the embodied theory of mind connects modules similar to Leslie’s ToBY and Baron-Cohen’s EDD, ID, and SAM together with real perceptual processes and with links to physical behaviors and extends the current models to behavior selection, attention, and more complex behavioral forms.

Overview of the hybrid theory of mind model