Design & Implementation of Cognition-enabled Robot Agents

Module 10: Cognitive Architectures
Lecture 1: Foundations

Institute for Artificial Intelligence
Universität Bremen

Winter Term 2020/21

Learning Goals

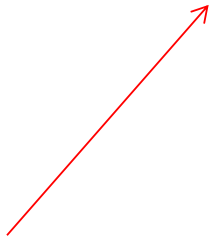
1. Describe the three paradigms of cognitive science
2. Explain the characteristics of a cognitive architecture in each of the three paradigms
3. Describe the key components of a hybrid cognitive architecture
4. Sketch the design of the ISAC and CRAM cognitive architectures and explain how they operate

Lecture Contents

1. What are the characteristics of a cognitive agent?
2. What is a cognitive architecture?
3. How does a cognitive architecture work?
4. The three paradigms of cognitive science
   - Cognitivist (symbolic) cognitive architectures
   - Emergent cognitive architectures
   - Hybrid cognitive architectures
5. Lecture summary
6. Recommended reading & references

1. What are the characteristics of a cognitive agent?

The chief characteristic of a cognitive agent is the ability to act effectively in pursuing goals in a world that is uncertain, under-specified, and dynamic, possibly cooperating with other cognitive agents



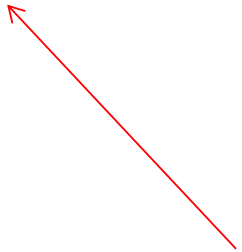
To achieve goals adaptively and robustly in these circumstances requires a complex system that can

- Construct models of the way the world works,
- Use them to guide actions prospectively, and
- Update them dynamically as the system continually learns through its interactions

A cognitive architecture is the way we specify what is required to achieve this.

2. What is a cognitive architecture?

A cognitive architecture is a software framework that integrates all the elements required for a system to exhibit the characteristic attributes of a cognitive agent



The design of a cognitive architecture requires the specification of the formalisms for all the processes and knowledge representations used by that framework

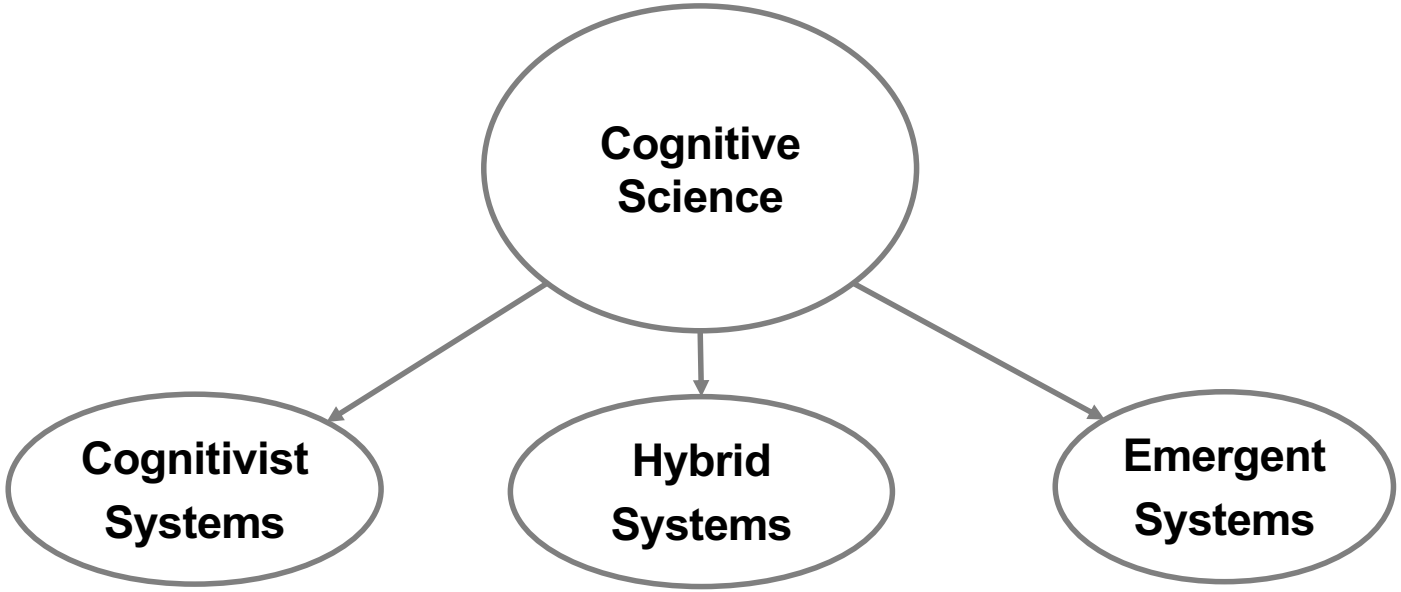
3. How does a cognitive architecture work?

A cognitive architecture integrates the core cognitive abilities so that these abilities can be dynamically coordinated

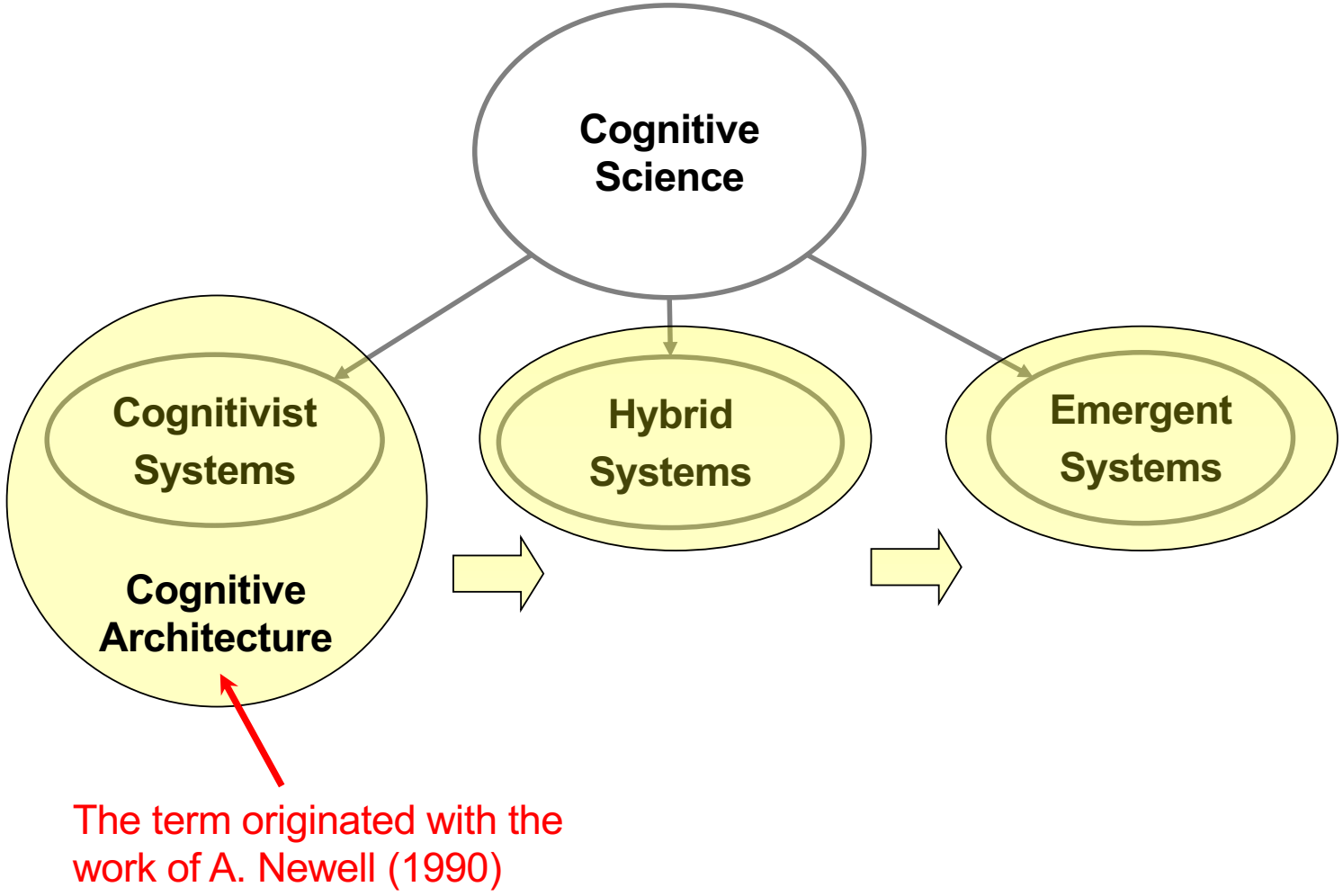
Perception
Attention
Action selection
Memory
Learning
Reasoning
Meta-reasoning
Prospection

Allowing the agent to exhibit flexible context-sensitive behaviour, prospectively selecting and controlling the actions that are required to achieve given goals

A cognitive architecture should also be able to develop autonomously so that its performance improves over time with experience



There are three paradigms of cognitive science


Cognitive Science
Cognitivist Systems
Cognitive Architecture
Hybrid Systems
Emergent Systems
The term originated with the work of A. Newell (1990)

Cognitivist Cognitive Architecture

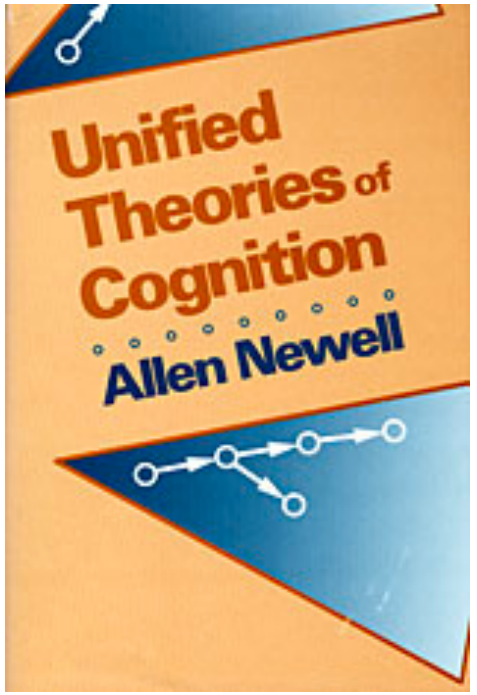
Attempts to create Unified Theories of Cognition (UTC)

UTCs cover a broad range of cognitive issues

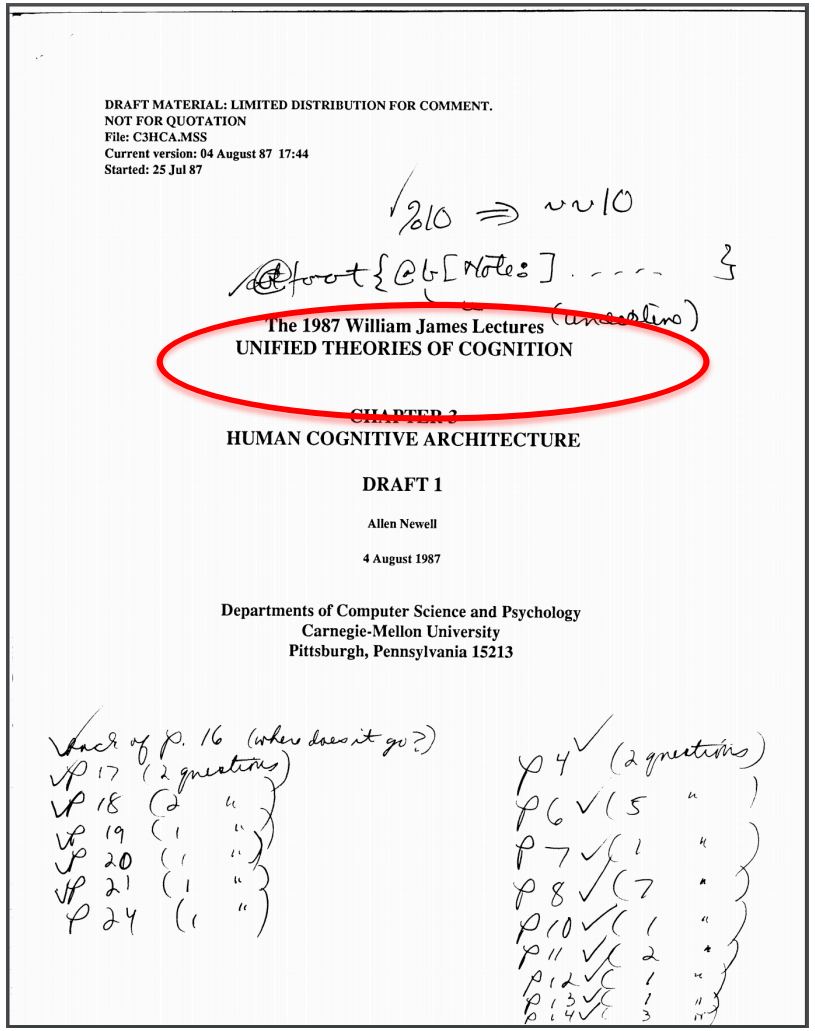
- Attention
- Memory
- Problem solving
- Decision making
- Learning

from several aspects

- Psychology
- Neuroscience
- Computer Science



https://www.hup.harvard.edu/catalog.php?isbn=9780674921016



DRAFT MATERIAL: LIMITED DISTRIBUTION FOR COMMENT.
NOT FOR QUOTATION
File: C3HCA.MSS
Current version: 04 August 87 17:44
Started: 25 Jul 87

The 1987 William James Lectures
UNIFIED THEORIES OF COGNITION

~~CHAPTER 3~~
HUMAN COGNITIVE ARCHITECTURE

DRAFT 1

Allen Newell

4 August 1987

Departments of Computer Science and Psychology
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

http://digitalcollections.library.cmu.edu/awweb/awarchive?type=file&item=352120

Cognitivist Cognitive Architecture

An encapsulation of a scientific hypothesis about those aspects of human cognition that are

- relatively constant over time and
- relatively independent of task

(Ritter and Young 2001)

Cognitivist Cognitive Architecture

- Generic computational model:
  - Not domain-specific
  - Not task-specific
- Knowledge provides the required specificity:

> Cognitive Architecture + Knowledge = Cognitive Model

  - Lehman et al. (1998) put it slightly differently:

BEHAVIOR = ARCHITECTURE x CONTENT

Cognitivist Cognitive Architecture

Knowledge is typically:

- Determined by the designer (explicitly or implicitly)
- Adapted and augmented by machine learning techniques

Cognitivist Cognitive Architecture

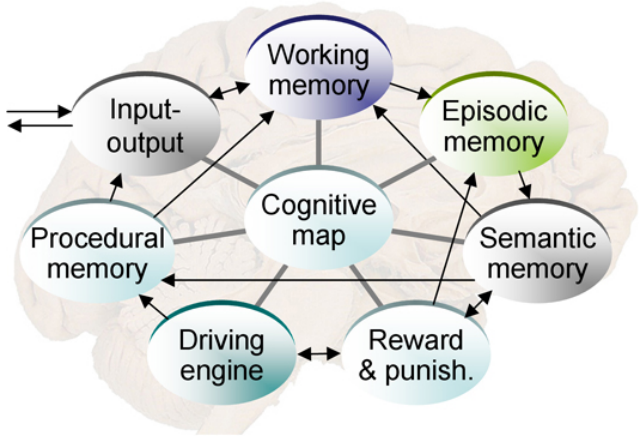
Overall structure and organization of a cognitive system

- Essential modules
- Essential relations between these modules
- Essential algorithmic and representational details in each module

(Sun 2007)

Commitment to formalisms for representation and processes

(Langley 2005, Langley 2006, Langley et al. 2009)



GMU BICA Architecture (Samsonovich 2010)

Emergent Cognitive Architecture

Emergent approaches focus on development

- From a primitive state
- To fully cognitive state, over the system's lifetime

https://childmaltreatmentresearchblog.wordpress.com/about/

Emergent Cognitive Architecture

- Two different views of development
  - Individual
  - Social
- Two complementary theories of cognitive development

Jean Piaget
1896–1980
https://en.wikipedia.org/wiki/Jean_Piaget

Lev Vygotsky
1896–1934
https://en.wikipedia.org/wiki/Lev_Vygotsky

Emergent Cognitive Architecture

The cognitive architecture is the system's phylogenetic configuration

- The basis for ontogenesis, i.e. growth and development
  - Innate skills
  - Core knowledge
- A structure in which to embed mechanisms for
  - Perception
  - Action
  - Adaptation
  - Anticipation
  - Motivation
  - … Development of all these

Emergent Cognitive Architecture

Strong focus on

- Autonomy-preserving, anticipatory, adaptive skill construction
- The morphology of the physical body in which the architecture is embedded

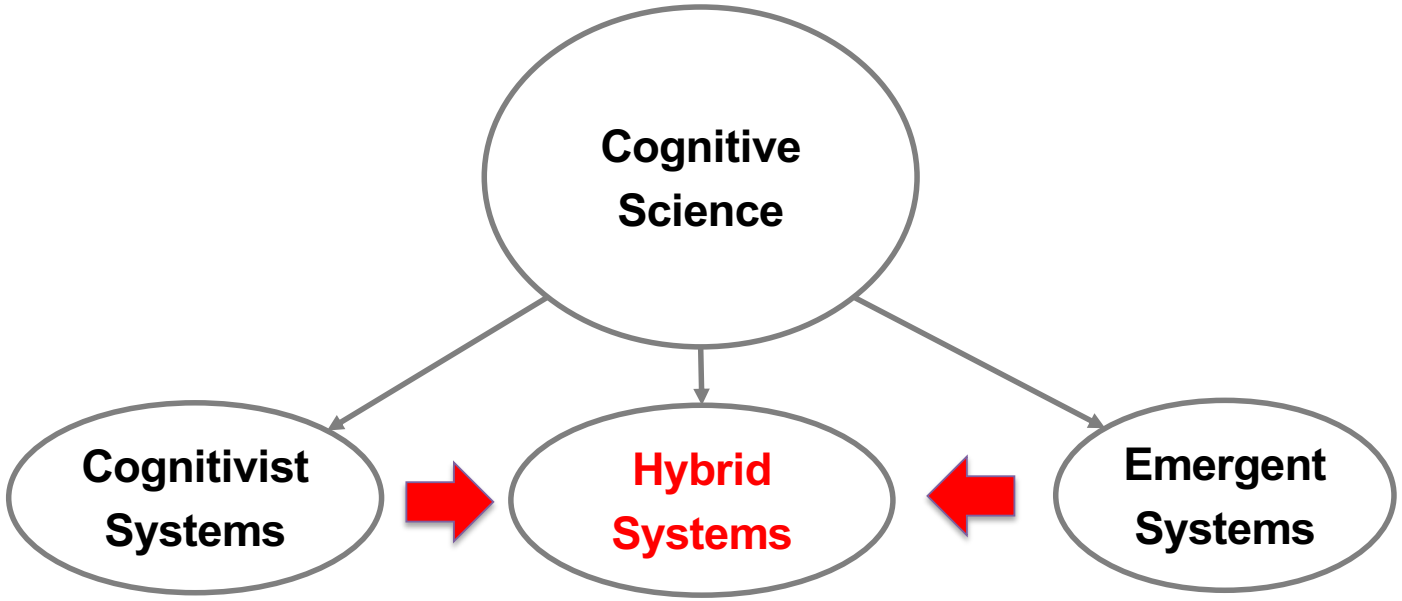
Emergent Cognitive Architecture

The emergent approach rejects:

- Dualism between mind and body
- Functionalism that treats cognitive mechanisms independently of the physical platform
  - Computational functionalism
  - Robotic functionalism

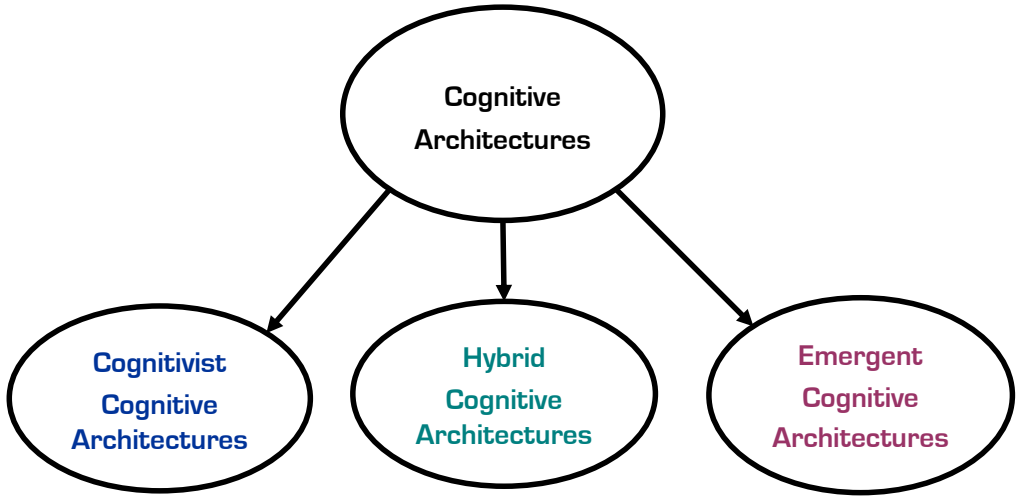
(Ziemke 2016)

Hybrid Cognitive Architecture



Organizational decomposition

- Explicit inter-connectivity
- Representational formalism
- Algorithmic formalism



Framework in which to embed knowledge

- Memories
- Formalisms for learning
- Programming mechanism

Phylogeny - basis for development

- Innate skills & core knowledge
- Memories
- Formalism for autonomy
- Formalism for development

Lecture Summary

1. The chief characteristic of a cognitive agent is the ability to act effectively in pursuing goals in an uncertain, incompletely-specified world
2. A cognitive architecture is a framework that integrates all the elements and core abilities of a cognitive system
3. It dynamically recruits these elements so that the agent exhibits flexible context-sensitive behaviour and prospective selection and control of its actions
4. A cognitive architecture specifies the formalisms for knowledge representation and processing
5. There are three types of cognitive architecture: cognitivist (symbolic), emergent, and hybrid

Recommended Reading

D. Vernon, Artificial Cognitive Systems – A Primer, MIT Press, 2014; Chapter 3: Cognitive Architectures.

D. Vernon. "Cognitive Architectures", in Cognitive Robotics, A. Cangelosi and M. Asada (Eds.), MIT Press, in press.

References

Krichmar, J. L. (2012). Design principles for biologically inspired cognitive architectures. Biologically Inspired Cognitive Architectures 1, 73–81.

Langley, P. (2005). An adaptive architecture for physical agents. In IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Compiegne, France, pp. 18–25. IEEE Computer Society Press.

Langley, P. (2006). Cognitive architectures and general intelligent systems. AI Magazine 27(2), 33–44.

Langley, P., J. E. Laird, and S. Rogers (2009). Cognitive architectures: Research issues and challenges. Cognitive Systems Research 10 (2), 141– 160.

Lehman, J. F., J. E. Laird, and P. S. Rosenbloom (1998). A gentle intro- duction to soar, an architecture for human cognition. In S. Sternberg and D. Scarborough (Eds.), Invitation to Cognitive Science, Volume 4: Methods, Models, and Conceptual Issues. Cambridge, MA: MIT Press.

Newell, A. (1990). Unified Theories of Cognition. Cambridge MA: Harvard University Press.

References

Ritter, F. E. and R. M. Young (2001). Introduction to this special issue on using cognitive models to improve interface design. International Journal of Human-Computer Studies 55, 1–14.

Samsonovich, A. (2010). Toward a unified catalog of implemented cog- nitive architectures. In Proc. the Conference on Biologically Inspired Cognitive Architectures, pp. 195–244.

Sun, R. (2004). Desiderata for cognitive architectures. Philosophical Psychology 17(3), 341–373.

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

Vernon, D., C. von Hofsten, and L. Fadiga (2016). Desiderata for developmental cognitive architectures. Biologically Inspired Cognitive Architectures 18, 116–127.

Ziemke, T. (2016). The body of knowledge: On the role of the living body in grounding embodied cognition. BioSystems 148, 4–11.