Artificial Cognitive Systems

Module 3: Cognitive Architectures

Lecture 1: The role of a cognitive architecture; desirable characteristics; core cognitive abilities; how to design a cognitive architecture

David Vernon
Carnegie Mellon University Africa

www.vernon.eu

What are the characteristics of a cognitive agent?

The chief characteristic of a cognitive agent is the ability to act effectively in a world that is uncertain, under-specified, 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.

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

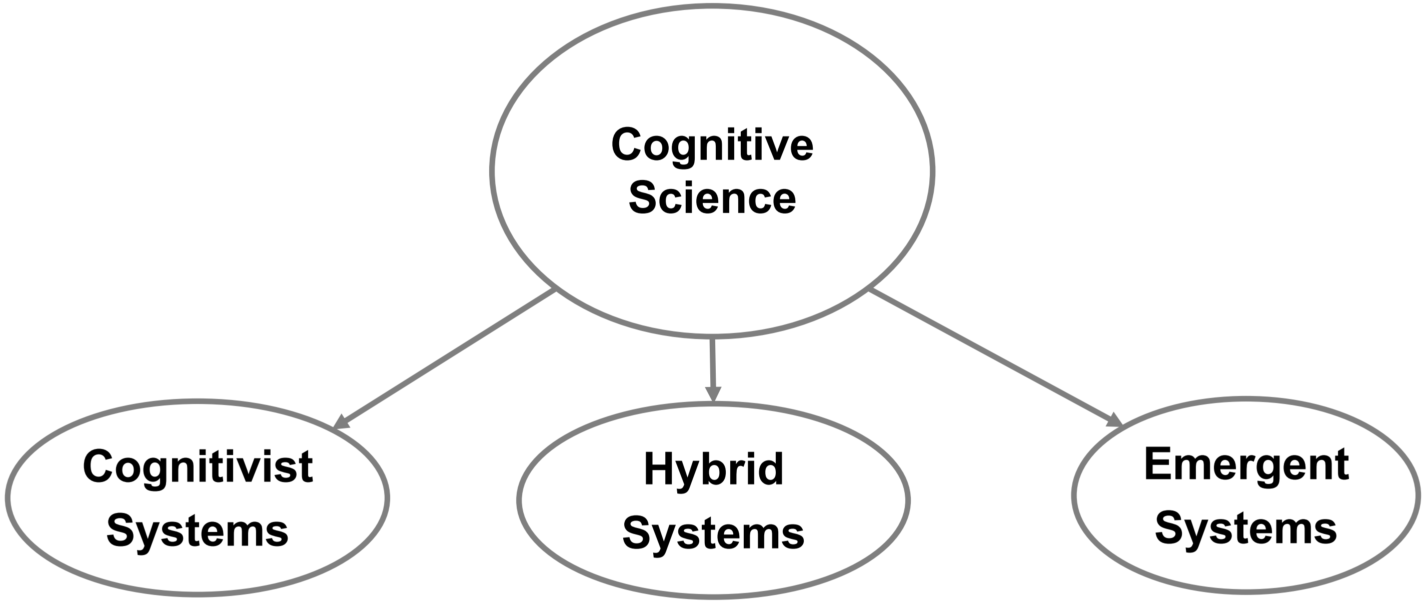
How does a cognitive architecture work?

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

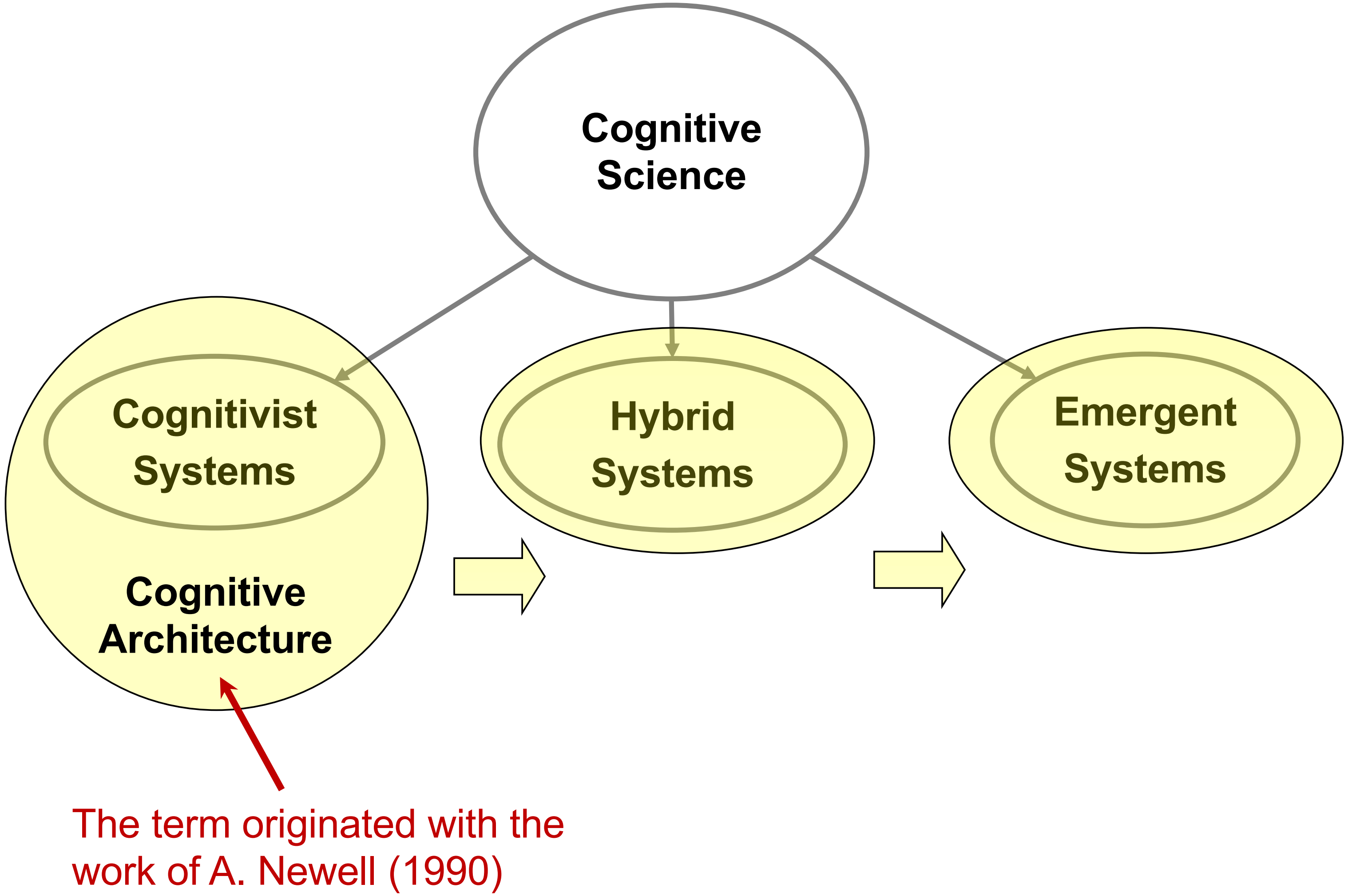
Perception
Attention
Action selection
Memory
Learning
Reasoning
Meta-reasoning
Prospection

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

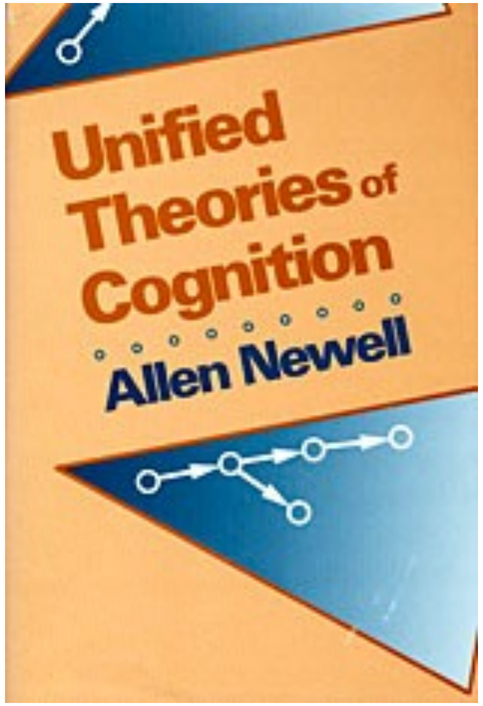
Attempts to create Unified Theories of Cognition (UTC)

UTCs cover a broad range of cognitive issues

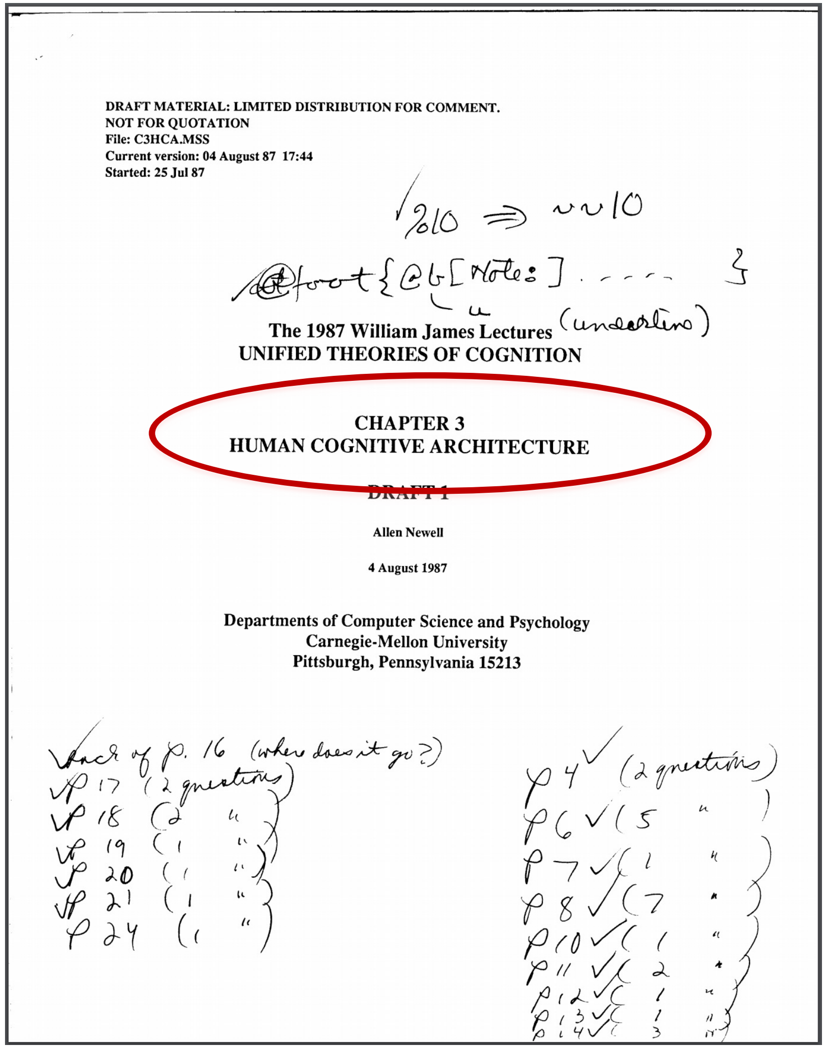
- Attention
- Memory
- Problem solving
- Decision making
- Learning

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 & 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:

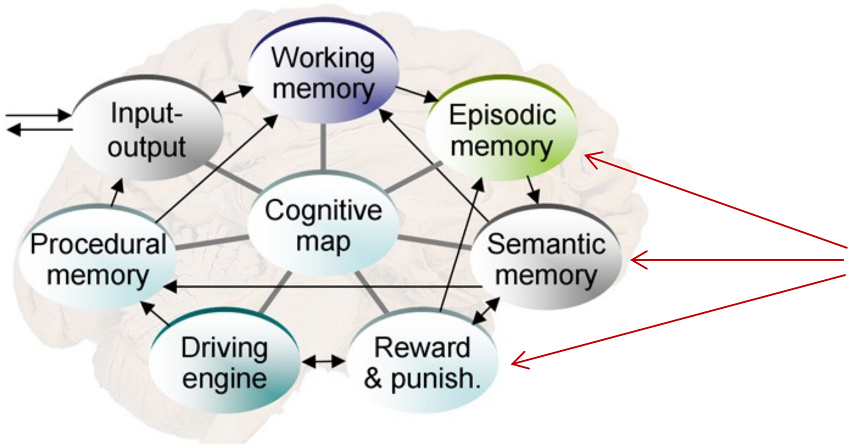
- Determined by the designer (explicitly or implicitly)
- Adapted and augmented by machine learning techniques

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]



(GMU-BICA Architecture: Samsonovich 2005)

Commitment to formalisms for representation and processes

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

Cognitivist Cognitive Architecture

Commitment to formalisms for

- Short-term & long-term memories that store the agent's beliefs, goals, and knowledge
- Representation & organization of structures embedded in memory
- Functional processes that operate on these structures
  - Performance / utilization
  - Learning
- Programming language to construct systems embodying the architectures assumptions

(Langley 05, Langley 06, Langley et al. 09)

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 different 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: growth and development
  - Innate skills
  - Core knowledge (cf. Spelke)

- 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

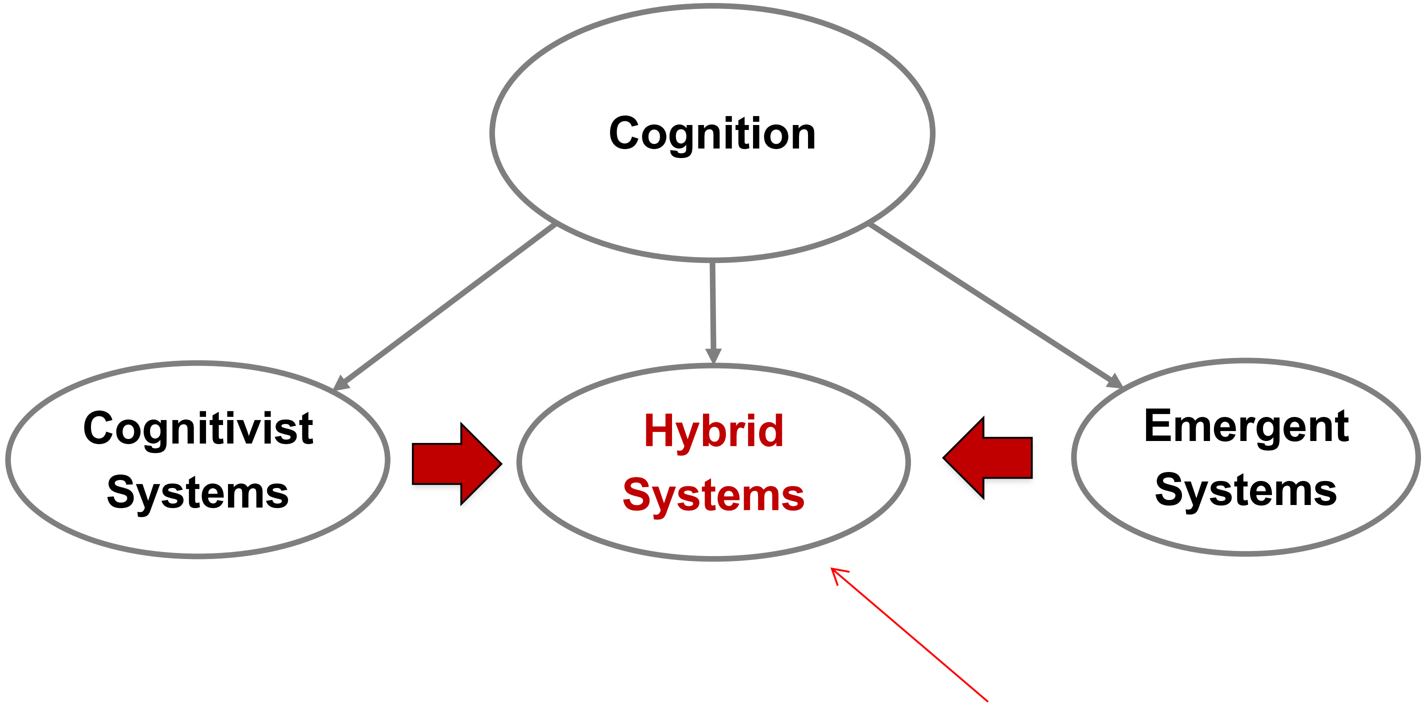
Emergent Cognitive Architecture

The emergent approach rejects:

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

Ziemke, T., The body of knowledge: On the role of the living body in grounding embodied cognition. BioSystems (2016), http://dx.doi.org/10.1016/j.biosystems.2016.08.005

Hybrid Cognitive Architecture



Symbolic & sub-symbolic representation
and computational processing
(i.e. h,not H)

Desirable Characteristics

- Realism
- Behavioral Characteristics
- Cognitive Characteristics

PHILOSOPHICAL PSYCHOLOGY, VOL. 17, NO. 3, SEPTEMBER 2004 — Routledge, Taylor & Francis Group

Desiderata for cognitive architectures

RON SUN

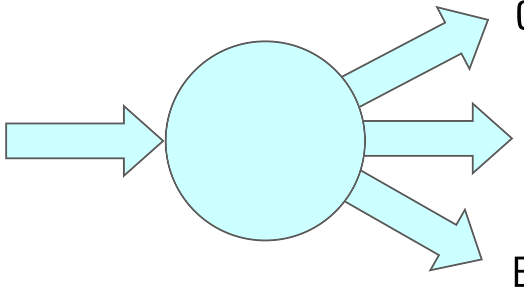
- Functional Capabilities (Langley et al. 2009, Sun 2007)
- Development (Krichmar & Edelman 2006, 2007; Vernon et al. 2016)
- Dynamics

Desirable Characteristics

Realism (Sun 2004):

1. Ecological realism
2. Bio-evolutionary realism
3. Cognitive realism
4. Inclusiveness of prior perspectives

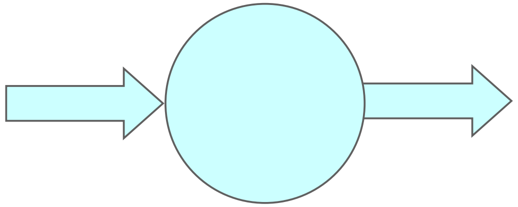
Desirable Characteristics



(Sun 2004)

Desirable Characteristics

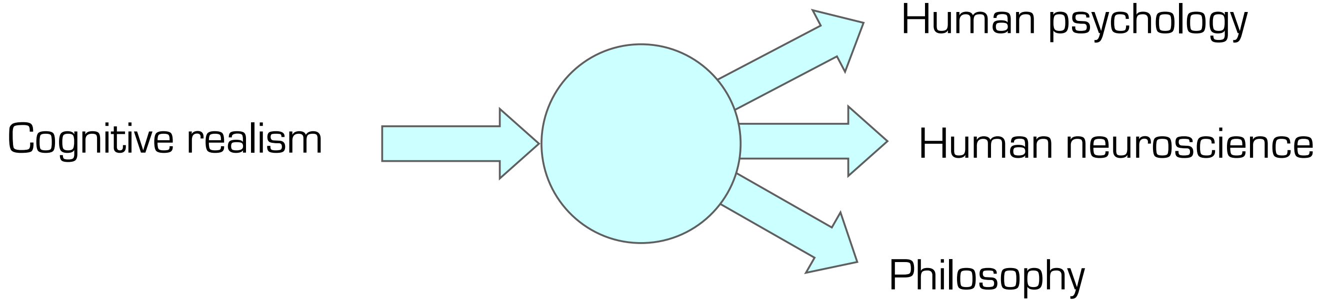
Bio-evolutionary realism



Human intelligence reducible to model of animal intelligence

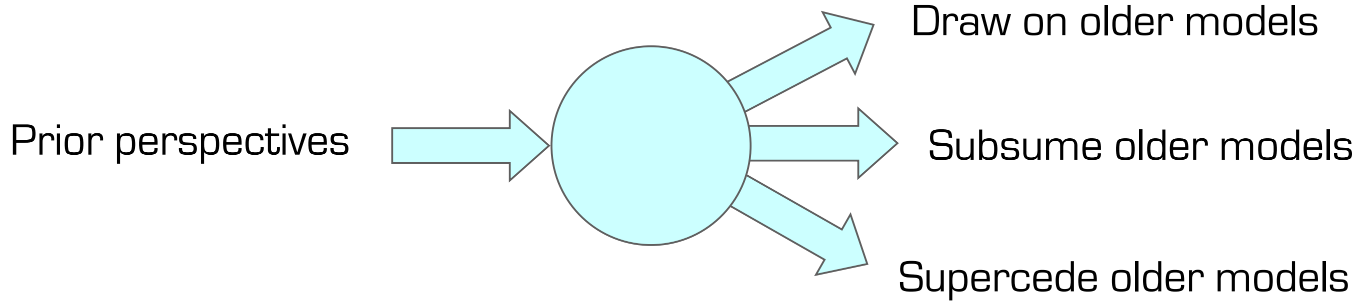
(Sun 2004)

Desirable Characteristics



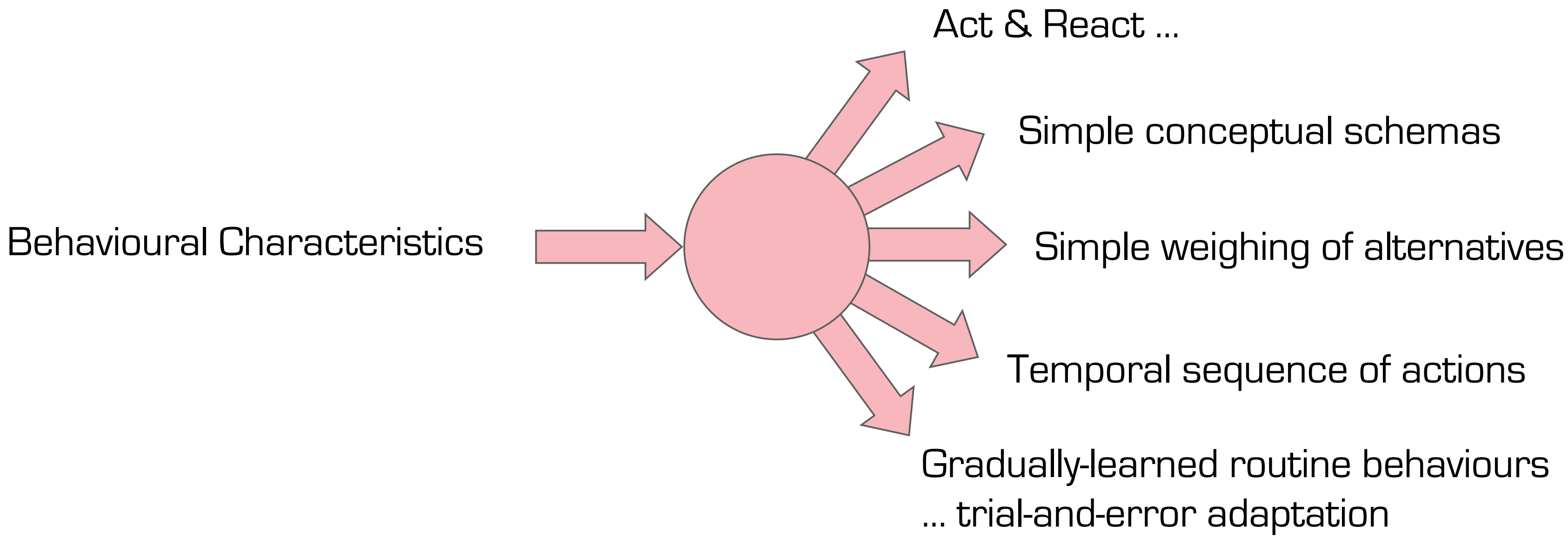
(Sun 2004)

Desirable Characteristics



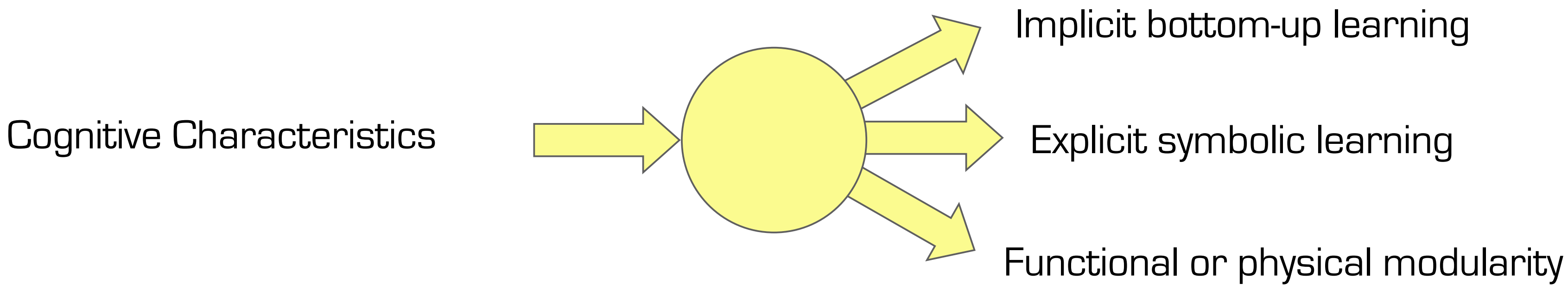
(Sun 2004)

Desirable Characteristics



(Sun 2004)

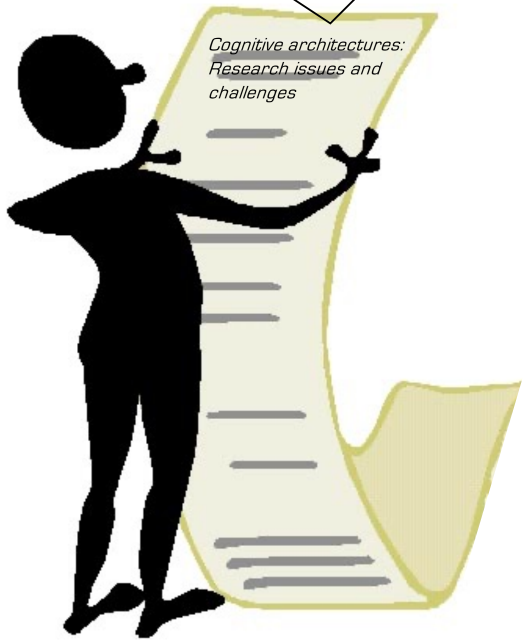
Desirable Characteristics



(Sun 2004)

Desirable Characteristics

Cognitive architectures: Research issues and challenges



(Langley et al. 2009)

1. Recognition & categorization
2. Decision-making & choice
3. Perception & situation assessment
4. Prediction & monitoring
5. Problem solving & planning
6. Reasoning & belief maintenance
7. Execution & action
8. Interaction & communication
9. Remembering, reflection, & learning

Desirable Characteristics

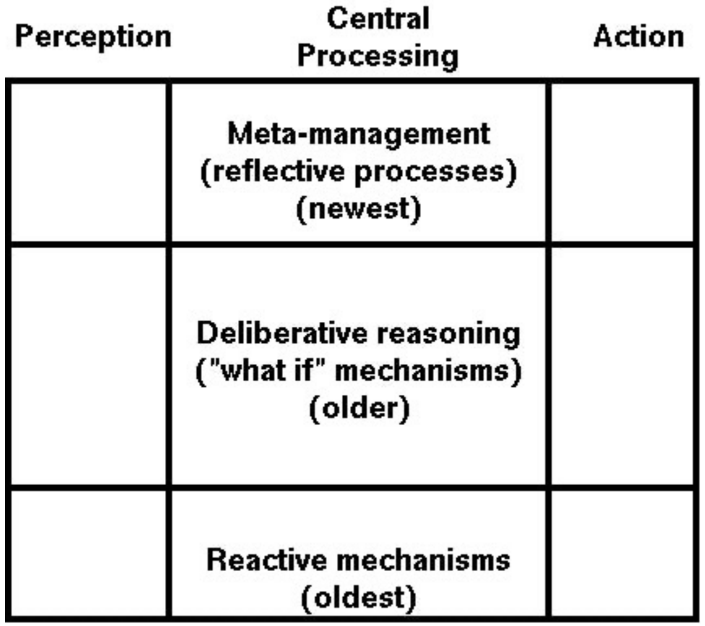
The importance of cognitive architectures …

(Sun 2007)

1. Perception
2. Categorization
3. Multiple representations
4. Multiple types of memory
5. Decision making
6. Reasoning
7. Planning
8. Problem solving
9. Meta-cognition
10. Communication
11. Action control and execution
12. Several types of learning

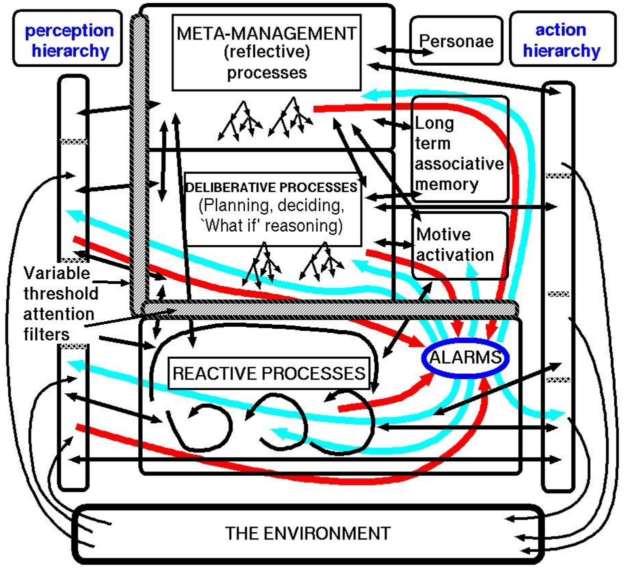
The importance of the interconnectivity between these processes

Desirable Characteristics



Cogaff Cognitive Architecture Schema

(Sloman 2000)



H-Cogaff Cognitive Architecture

(Sloman 2001)

Desirable Characteristics

Development

Cognitive Architectures of Developmental Systems (Krichmar & Edelman 2005, 2006)

1. Address connectivity and interaction between circuits/regions in the brain
2. Effect perceptual categorization, without a priori knowledge
   (a model generator, rather than a model fitter, cf (Weng 04))
3. Embodied & capable of exploration
4. Minimal set of innate behaviours
5. Value system (set of motivations) to govern development

Desirable Characteristics

PHILOSOPHICAL PSYCHOLOGY, VOL. 17, NO. 3, SEPTEMBER 2004

Routledge
Taylor & Francis Group

Desiderata for cognitive architectures

RON SUN

ELSEVIER

Biologically Inspired Cognitive Architectures

Volume 18, October 2016, Pages 116–127

Research article

Desiderata for developmental cognitive architectures

David Vernon[a], Claes von Hofsten[b], Luciano Fadiga[c, d]

Show more

http://dx.doi.org/10.1016/j.bica.2016.10.004

Get rights and content

Desirable Characteristics

Desideratum 1. Value systems and motives

Desideratum 2. Physical embodiment

Desideratum 3. Sensorimotor contingencies

Desideratum 4. Perception

Desideratum 5. Attention

Desideratum 6. Prospective action

Desideratum 7. Declarative and procedural memory

Desideratum 8. Multiple modes of learning

Desideratum 9. Internal simulation

Desideratum 10. Constitutive autonomy

ELSEVIER

Biologically Inspired Cognitive Architectures

Volume 18, October 2016, Pages 116–127

Research article

Desiderata for developmental cognitive architectures

David Vernon[a], Claes von Hofsten[b], Luciano Fadiga[c, d]

Show more

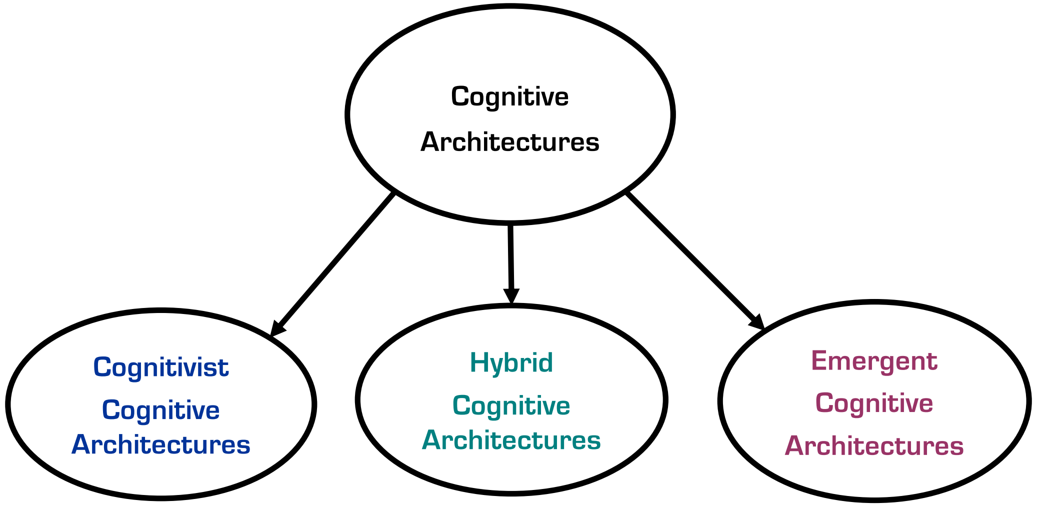
http://dx.doi.org/10.1016/j.bica.2016.10.004

Facets of a Cognitive Architecture

- Component functionality
- Component interconnectivity
- System dynamics

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

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

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.

← Not included in (Kotseruba and Tsotsos 2020)

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

Makes use of many sensory modalities, e.g. vision, audition, and haptic (tactile and kinesthetic)

Perception transforms raw input into the system's internal representation

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

Reduces the information a cognitive system has to process by selecting relevant information & filtering out irrelevant information

Selective mechanisms

Choose one entity from many, e.g. gaze & viewpoint

Restrictive mechanisms

Choose some entities from many:

Priming what to look for or where to look for it

Suppressive mechanisms

Suppress some entities from many

i.e. features, objects, or locations that are not relevant

(Kotseruba and Tsotsos 2018)

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

Determines what the agent should do next

Planning

Determines a sequence of steps to read a certain goal prior to execution of the plan

Dynamic action selection

Selection of one action based on knowledge at the time, typically using winner-take-all, probabilistic, or pre-defined order selection mechanisms

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

Kotseruba and Tsotsos identify six types of memory

Short-term sensory memory	recent percepts
Short-term working memory	information relevant to current task
Long-term episodic memory	key to anticipation; autobiographical
Long-term semantic memory	general knowledge about the world
Long-term procedural memory	motor skills
Long-term global memory	for architectures that don't draw a type-duration distinction:

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

The ability of a system to improve performance over time through the acquisition of knowledge or skill

Declarative learning … explicit knowledge acquisition
Non-declarative learning … perceptual, procedural, associative, non-associative learning

Supervised learning
Unsupervised learning
Reinforcement learning

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

The ability to logically and systematically process knowledge, typically to infer conclusions

Three classical forms of logical inference: deduction, induction, abduction

Reasoning focusses on the practical objective of finding the next (best) action to perform

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

Metacognition refers to the ability of a cognitive system has to monitor its internal cognitive processes, reason about them, and adapt them

Metacognition is needed for social cognition if the agent is to form a theory of mind, also known as perspective-taking, i.e. the ability to infer the cognitive states of other agents with which it is interacting

Core Cognitive Abilities

1. Perception
2. Attention
3. Action selection
4. Memory
5. Learning
6. Reasoning
7. Metacognition
8. Prospection

Prospection – the capacity to anticipate the future – is one of the hallmark attributes of cognition

It lies at the heart of the other core characteristics of a cognitive agent: autonomy, perception, action, learning, and adaptation

[Vernon 2014]

It is central to action since actions are goal-directed and guided by prospective information

[von Hofsten 2009]

Internal simulation plays a key role in prospection

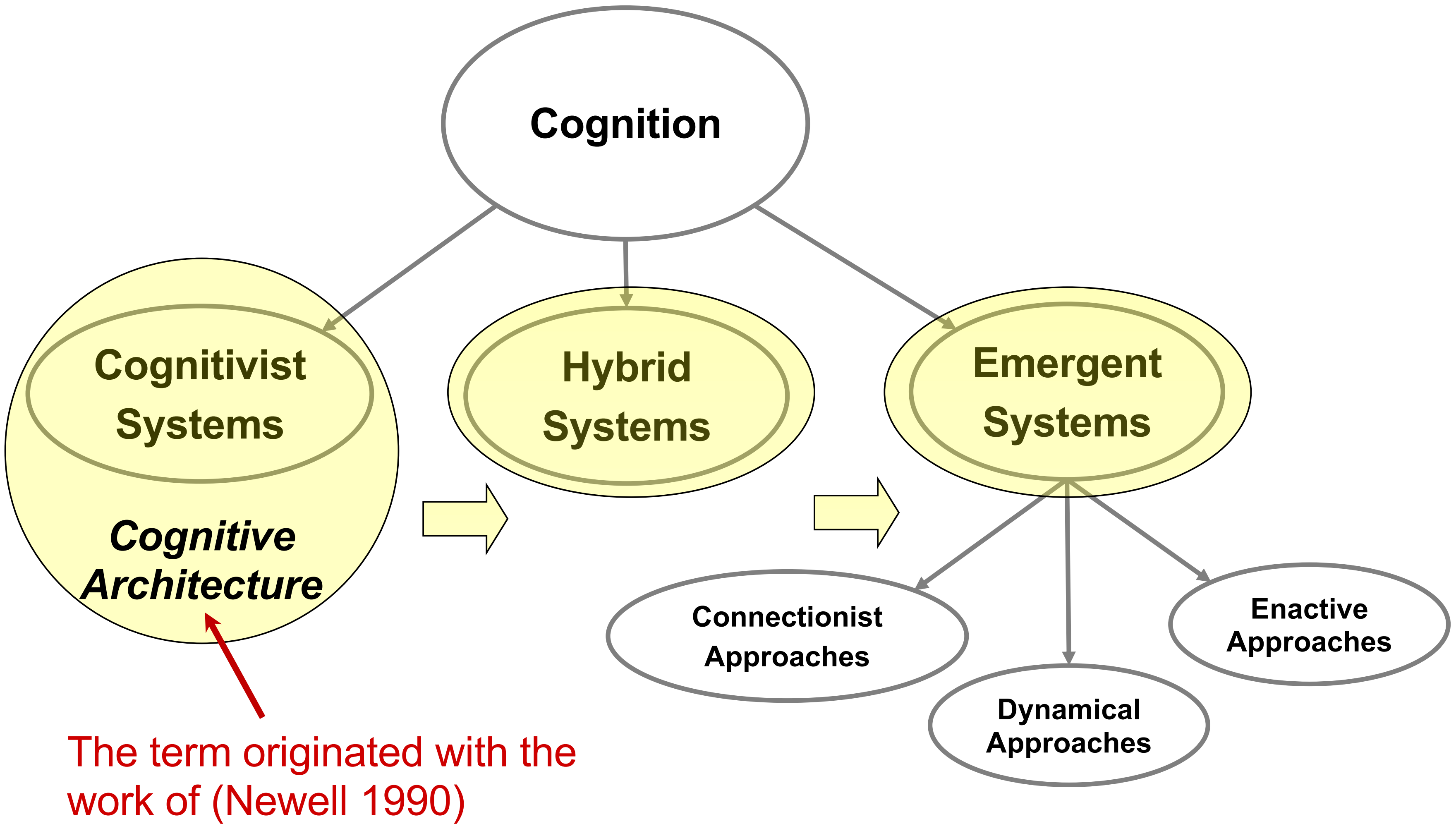
How to Design A Cognitive Architecture

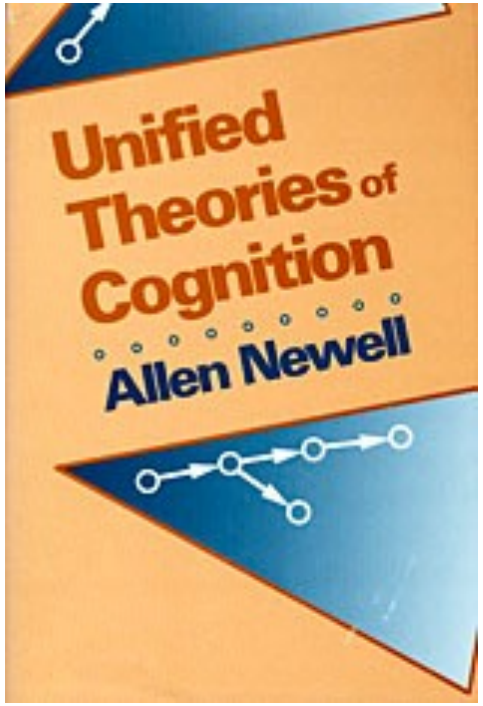
D. Vernon. "Two Ways (Not) To Design a Cognitive Architecture", Proceedings of EUCognition 2016, Cognitive Robot Architectures, European Society for Cognitive Systems, Vienna, 8-9 December, 2016, R. Chrisley. V. C. Müller, Y. Sandamirskaya. M. Vincze (eds.), CEUR-WS Vol-1855, ISSN 1613-0073, pp. 42-43.

There are **two** reasons that people work in cognitive system

1. They want smart systems, e.g. robots

2. **They want to understand cognition**


Cognition
Cognitivist
Systems
Cognitive
Architecture
Hybrid
Systems
Emergent
Systems
Connectionist
Approaches
Dynamical
Approaches
Enactive
Approaches
The term originated with the
work of (Newell 1990)



“all aspects, all perspectives”

What the cognitive scientist wants to design

What the customer might really need

How the customer explained it

Go for what's desirable:
desiderata

Clarify the explanation:
use cases

What the cognitive scientist wants to design

What the customer might really need

Design by Desiderata

Design by Use Case

LEA
Robot Care Systems

MOBILITY PROBLEMS

General weakness of muscles and stiffness for which the elderly uses stroller support

COGNITIVE PROBLEMS

Mild perceptive and cognitive impairment (e.g. impaired hearing, seeing, scent, as well as forgetfulness due to aging)

TREMOR

Tremor in arm muscles due to which elderly have trouble lifting or/and carrying objects

Mr Tucker (82) is suffering from hip injury

Mrs Baker (85) is suffering from dementia

Mrs Janssen (83) is suffering from tremor

LEA from Robot Care Systems

USE CASES TWO SCENARIOS

Use cases for single user at home

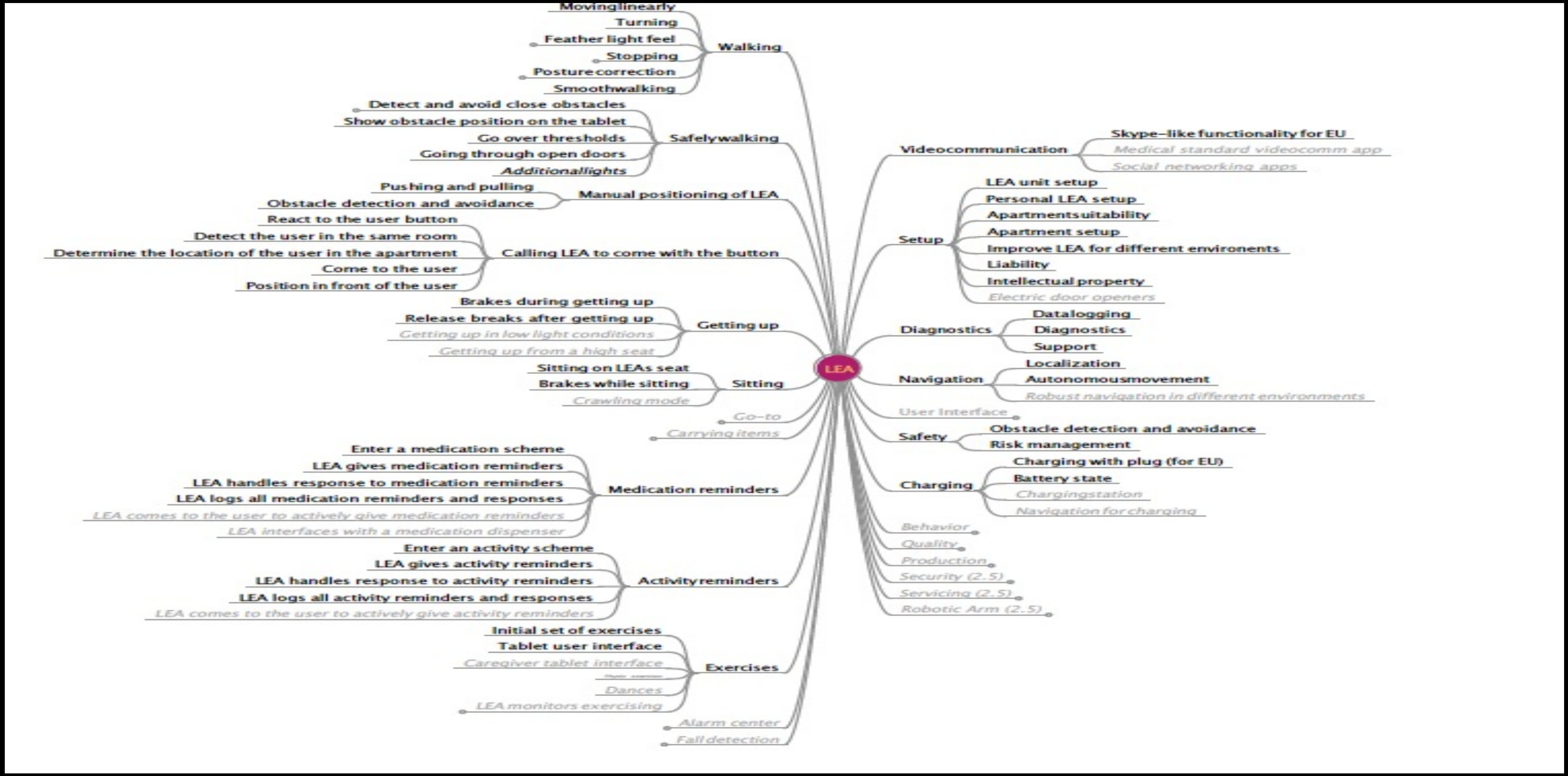
Use case ADL 1: Personal hygiene and grooming
Use case ADL 2: Eating and drinking
Use case ADL 3: Continence
Use case ADL 4: Personal assistant (reminders)
Use case ADL 5: Functional transfers
Use case ADL 6: Exercises
Use case ADL 7: Communication via internet
Use case ADL 8: Remote care (additional functionality)

Use cases for elderly living in care centres

Use case ADL 1: Personal hygiene and grooming
Use case ADL 2: Eating and drinking
Use case ADL 3: Continence
Use case ADL 4: Personal assistant (reminders)
Use case ADL 5: Functional transfers
Use case ADL 6: Exercises
Use case ADL 7: Communication with family, caregivers via internet

NO NAVIGATION FUNCTIONALITIES!!

LEA from Robot Care Systems



LEA from Robot Care Systems

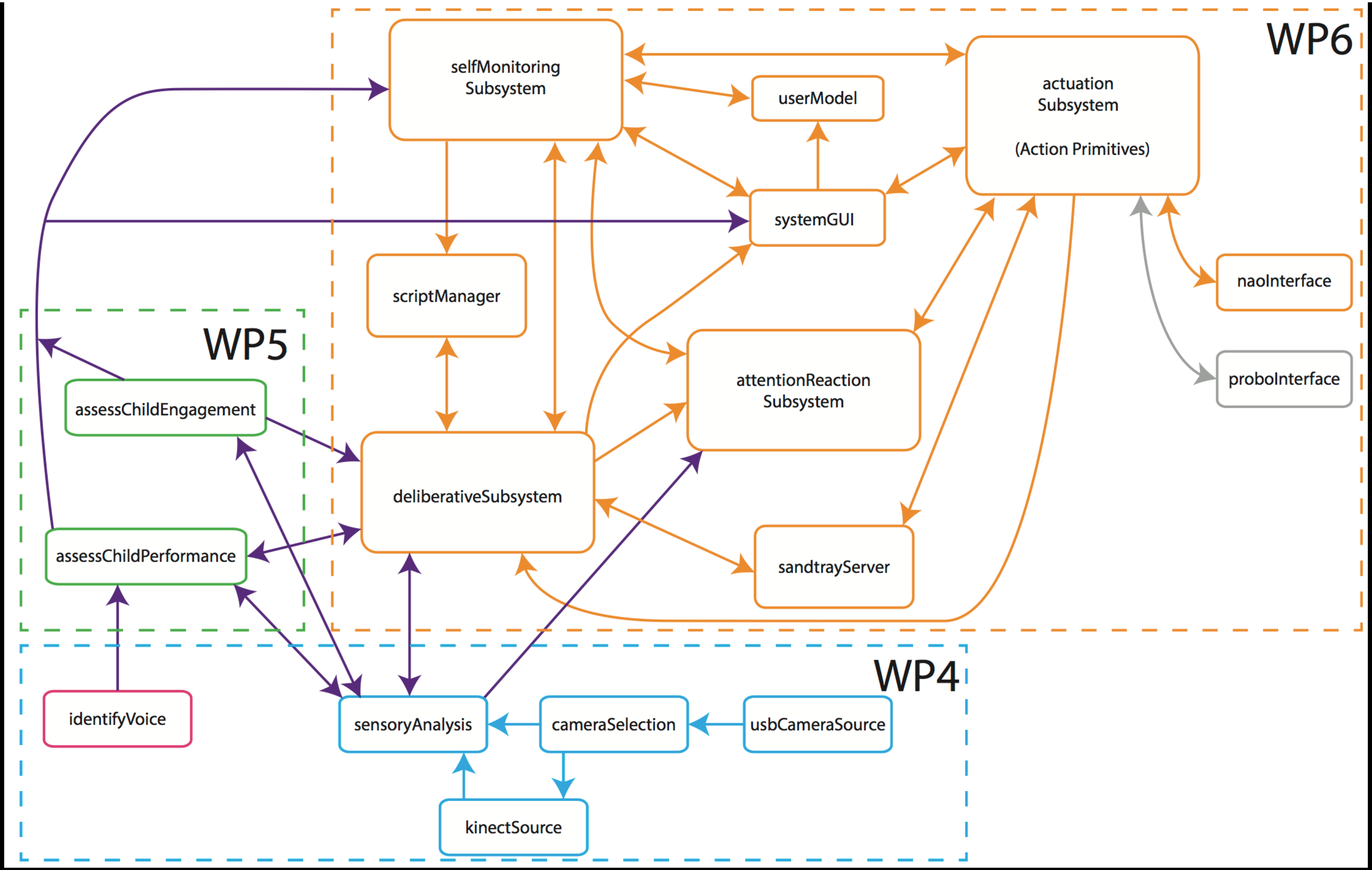
Assisting in Psychotherapy with ASD Children

(**Simple** Perspective Taking in Interaction Tasks)

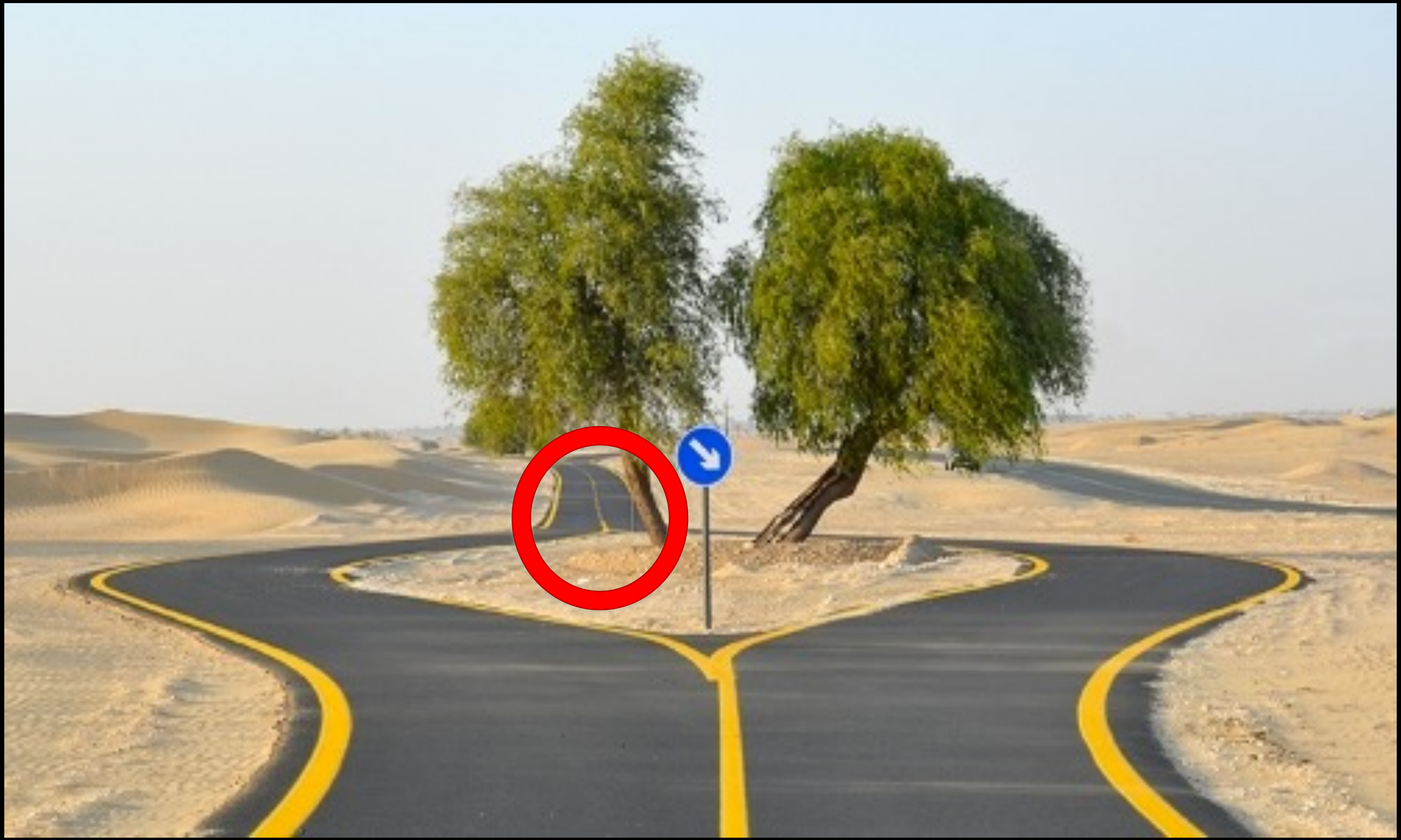
D1.1 Intervention Definition

Joint attention

The actions defined in the intervention tasks	The component movements and sensory cues	The sensory-motor processes provide input for the definition of robot and child behaviour specification	The comments which add explanations of what is happening in the task at that point.
Look at a picture	The robot looks at the picture/object to the left or right	Object detection Object localization Move head to centre gaze on the picture	The robot stares at the picture/object for a specified time
Look back at child	Look for a face UNTIL the child's face is detected Determine the location of the child Move head to look at the child Adjust body posture to face the child	Face detection Face recognition Face localization Move head to centre gaze on the child Move torso to face child and adjust gaze	The child has to identify where the robot looked
Check to see if the child is looking at the picture	Determine the gaze direction of the child	Compute child's head gaze Compute child's eye gaze (ideally)	The child has to look at one of the objects

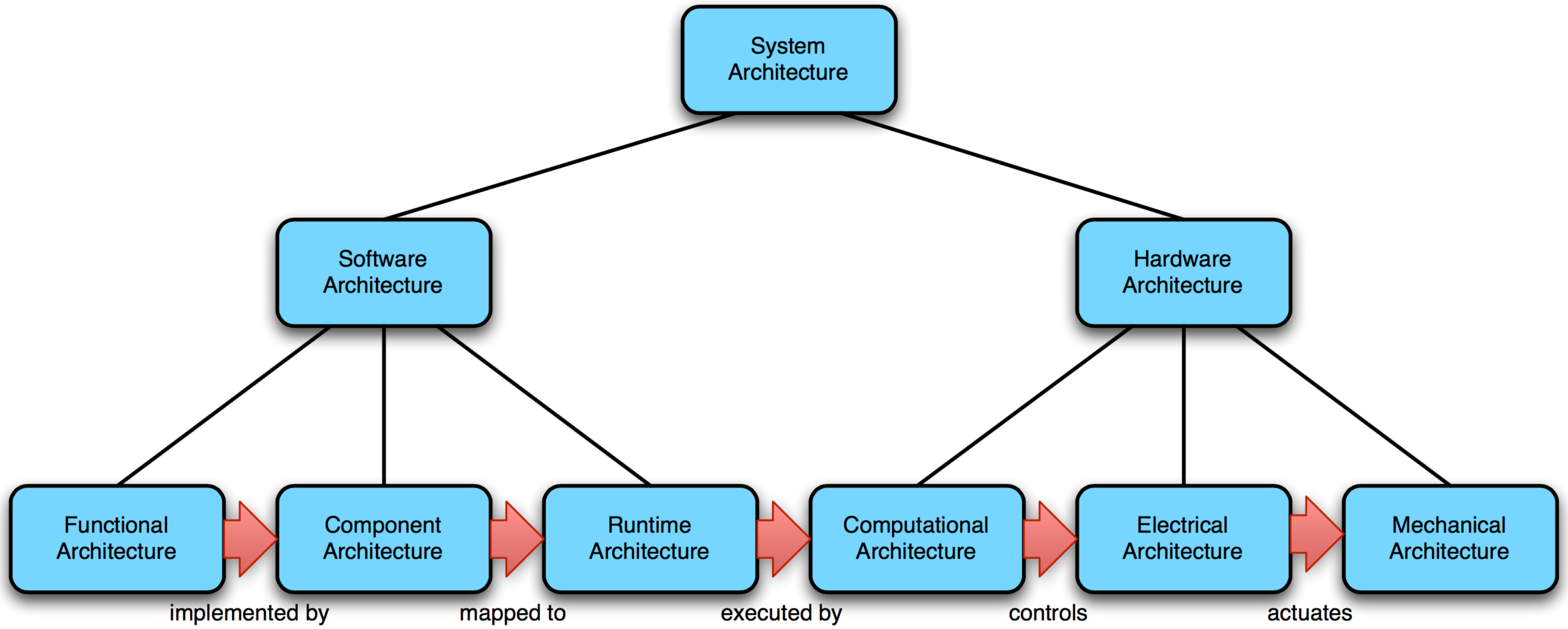

WP6
selfMonitoring Subsystem
userModel
actuation Subsystem
(Action Primitives)
systemGUI
scriptManager
naoInterface
proboInterface
attentionReaction Subsystem
WP5
assessChildEngagement
deliberativeSubsystem
assessChildPerformance
sandtrayServer
WP4
identifyVoice
sensoryAnalysis
cameraSelection
usbCameraSource
kinectSource
European Commission

Design by Desiderata or Use Case



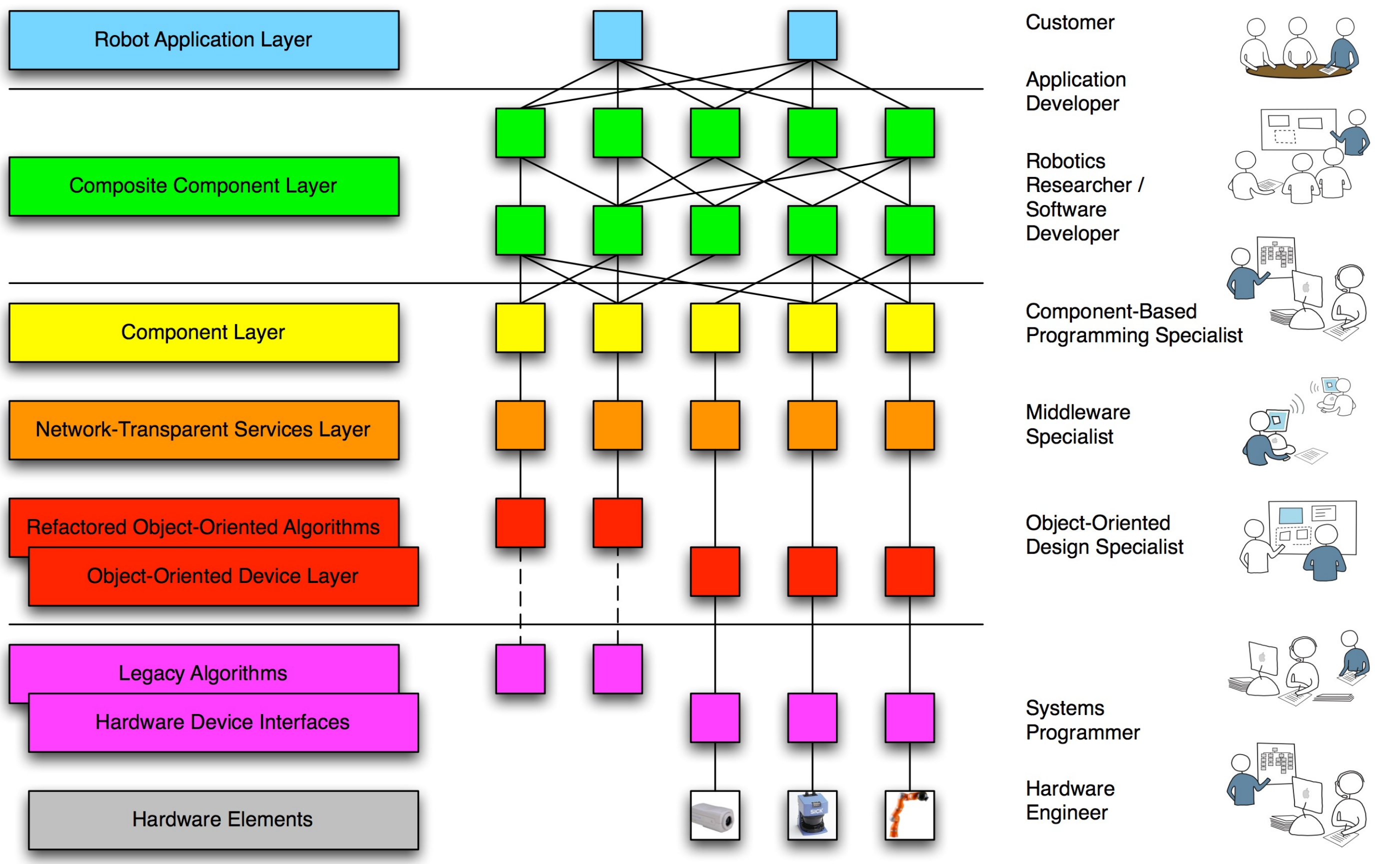
Design by Desiderata or Use Case

How to Design a Cognitive Architecture



G. Kraetzschmar, Software Engineering Factors for Cognitive Robotics, RockEU2 Robotics Coordination Action for Europe Two, Deliverable 3.5, 2017.
https://www.eu-robotics.net/eurobotics/about/projects/rockeu2.html

How to Design a Cognitive Architecture



G. Kraetzschmar, Software Engineering Factors for Cognitive Robotics, RockEU2 Robotics Coordination Action for Europe Two, Deliverable 3.5, 2017.
https://www.eu-robotics.net/eurobotics/about/projects/rockeu2.html

Reading

D. Vernon, *Artificial Cognitive Systems – A Primer*, MIT Press, 2014; Chapter 3, Sections 3.1, 3.2, 3.3, pp. 63-75.

A. Lieto, M. Bhatt, A. Oltramari, and D. Vernon, "The Role of Cognitive Architectures in General Artificial Intelligence", editorial for a special issue on "Cognitive Architectures for Artificial Minds", Cognitive Systems Research, Vol. 48, pp. 1-3.

D. Vernon, "Two ways (not) to design a cognitive architecture", Proceedings of the European Society for Cognitive Systems Meeting, EUCognition 2016, Vienna, 8-9 December, R. Chrisley. V. C. Müller, Y. Sandamirskaya. M. Vincze (eds.), CEUR-WS Vol-1855, ISSN 1613-0073, pp. 42-43, 2016.

D. Vernon, "Cognitive Architectures", in Cognitive Robotics, A. Cangelosi and M. Asada (Eds.), MIT Press, 2022, Sections 10.1 – 10.5.

Further Reading

R. Sun, “Desiderata for cognitive architectures,” Philosophical Psychology, vol. 17, no. 3, pp. 341–373, 2004

D. Vernon, C. von Hofsten, and L. Fadiga, "Desiderata for Developmental Cognitive Architectures", Biologically Inspired Cognitive Architectures, Vol. 18, pp. 116-127, 2016.

D. Vernon, "The Architect's Dilemmas", in Cognitive Architectures, M. Ferreira, J. Sequeira, and R. Ventura (eds.), Intelligent Systems, Control and Automation: Science and Engineering, Vol. 94, Springer, 2019.