CERA-CRANIUM: A Test Bed for Machine Consciousness Research

Raúl Arrabales, Agapito Ledezma and Araceli Sanchis

Computer Science Department
Carlos III University of Madrid
http://Conscious-Robots.com/Raul

International Workshop on Machine Consciousness 2009
PolyU, Hong Kong, 14th June 2009

Contents

- Introduction and objectives.
- Related work.
- Objectives.
- CERA-CRANIUM.
- Experimentation settings.
- Example of application.
- Conclusions.

Introduction (I)

- Theories of consciousness
  - Philosophical or psychological background.
  - Metaphorical descriptions.

- Need to bridge the gap between theories and implementation.

Introduction (II)

- Cognitive theories of consciousness
  - Global Workspace Theory (Baars, 1997).
  - Multiple Draft Model (Dennett, 1991).

- Inspiration for the design of a partial computational model of consciousness.

Introduction (III)

- Common ground
  - Non-unitary mechanisms producing the unity of self.
  - In other words, conscious contents emerge as a result of competition/collaboration (Minsky, Dennett, Hofstadter, Baars, Shanon).

Introduction (IV)

- Objectives
  - To understand how cognitive skills associated with consciousness can be integrated effectively.
  - To test different machine consciousness models.

Introduction (V)

- Functionalism:
  - Is this a reductionist approach?
  - Does this mean that phenomenal aspects are ignored?

Related work (I)

- Shanahan's (2005, 2006) cognitive architecture.
- IDA and LIDA (Ramamurthy et al., 2006).
- Computational Agent Framework for Consciousness (Moura & Bonzon, 2004).
- CERA-CRANIUM (Arrabales et al., 2007, 2008).

Related work (II)

- Common denominator:
  - Shared workspace and specialized processors.
- Different approaches in terms of:
  - Architecture.
  - Problem domain.
  - Perceptual flow.
  - Decision taking.
  - Modulation.

Objectives

- Experimentation platform.
- Test bed for high level cognitive approaches.
- Generic but configurable cognitive architecture.

CERA-CRANIUM (I)

- CERA-CRANIUM Provides:
  - Mechanisms for specialized processors
    - Creation,
    - Association,
    - Combination,
    - Competition.
  - Mechanisms to regulate the former processes.

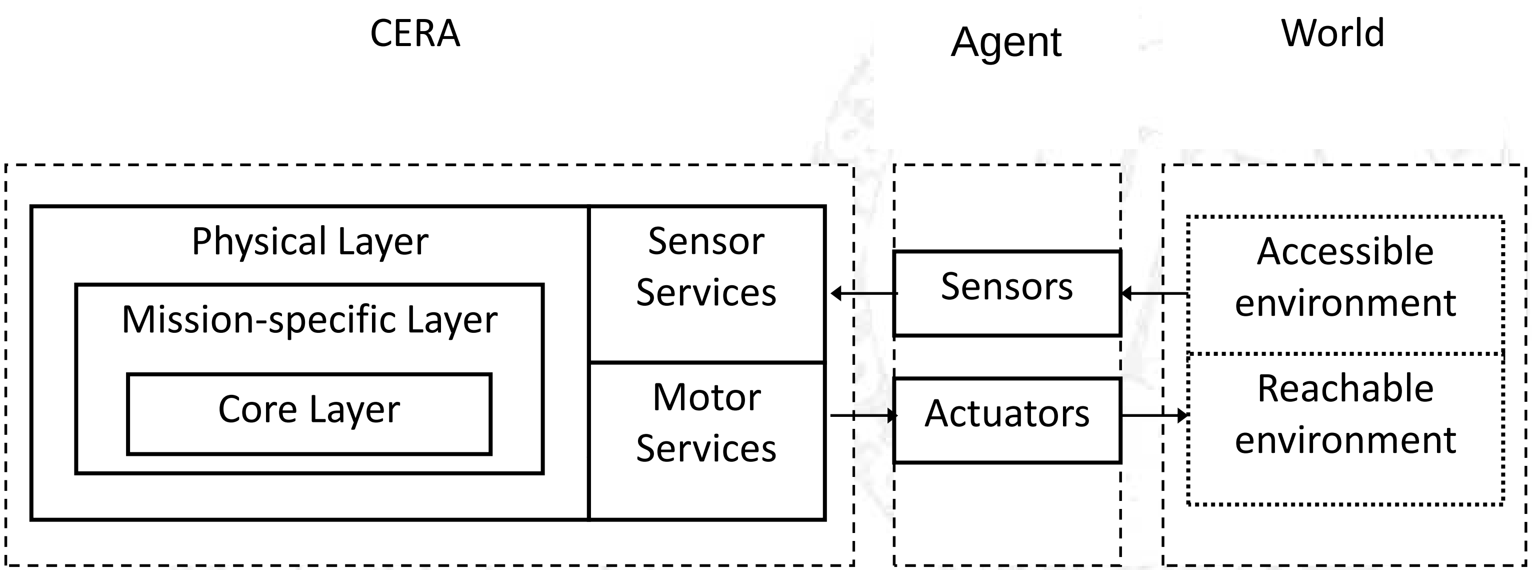
CERA-CRANIUM (II)

- CERA: layered control architecture.
- CRANIUM: runtime tool for the creation and management of high amounts of parallel processes in shared workspaces.

CERA-CRANIUM (III)

- CERA: layered design
  - Sensorimotor services layer.
  - Physical layer.
  - Mission-specific layer.
  - Core layer.

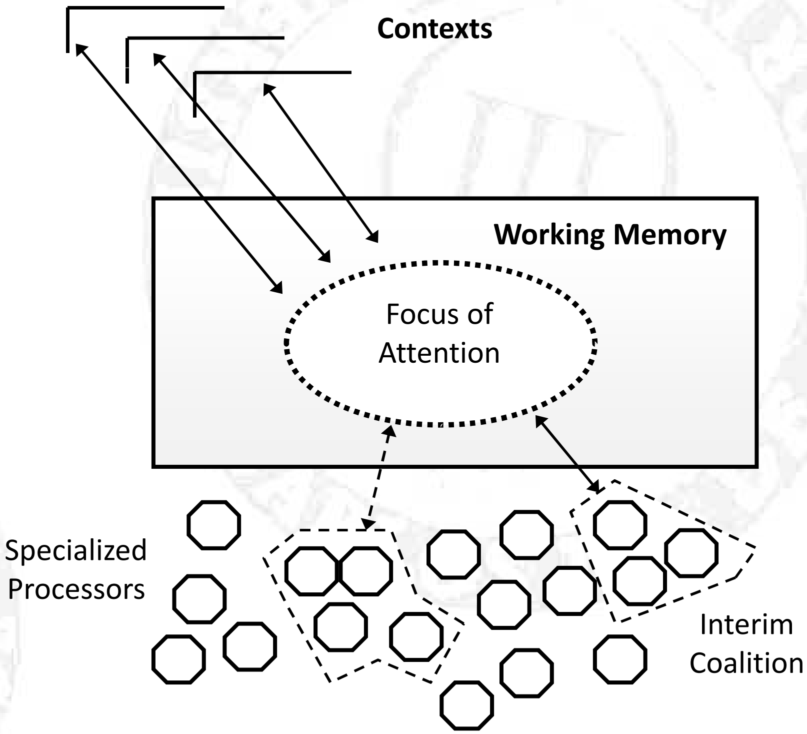
CERA-CRANIUM (IV)



Different levels of description

CERA-CRANIUM (V)

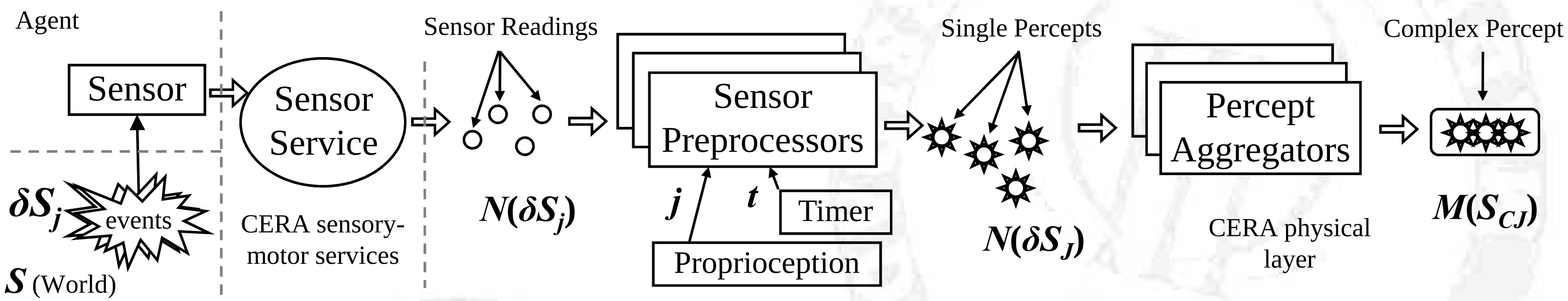
- CRANIUM
  - Blackboard
  - Pandemonium



CERA-CRANIUM (VI)

- Perceptual flow:
  - Sensory data.
  - Single percepts.
  - Complex percepts.
  - Mission percepts.

CERA-CRANIUM (VIbis)

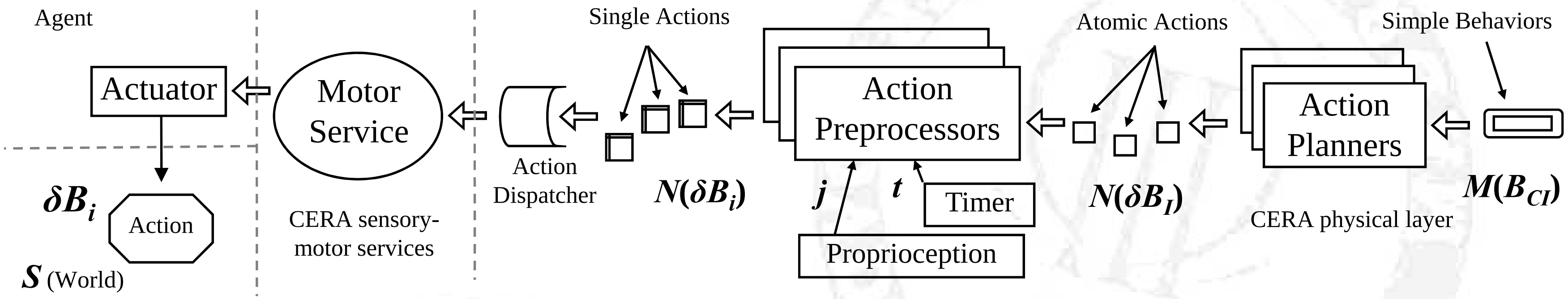


Knowledge representation

CERA-CRANIUM (VII)

- Behavior generation:
  - Mission behaviors.
  - Simple behaviors.
  - Single actions.
  - Motor controller commands.

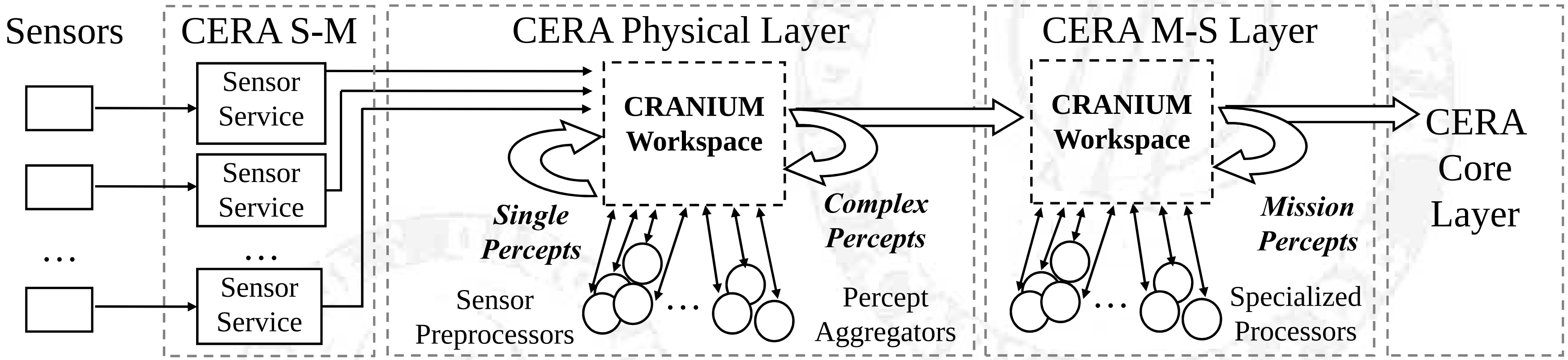
CERA-CRANIUM (VIIbis)



Action generation

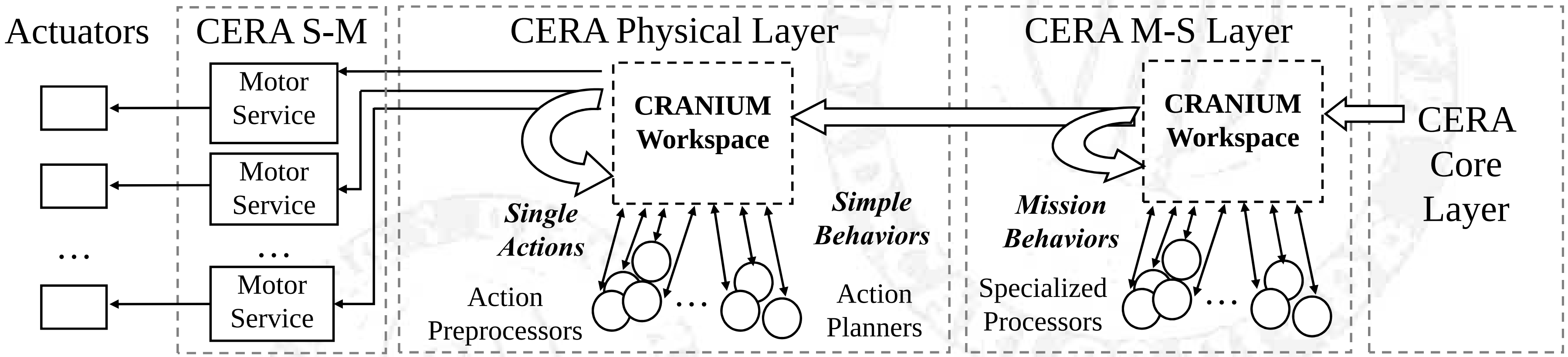
CERA-CRANIUM (VIII)

- Bottom-up flow



CERA-CRANIUM (IX)

- Top-down flow



CERA-CRANIUM (X)

- CRANIUM processor types
  - Sensor preprocessors
    - Raw sensory data → single percepts

CERA-CRANIUM (XI)

- CRANIUM processor types
  - Action preprocessors
    - Atomic actions → Single actions.

CERA-CRANIUM (XII)

- CRANIUM processor types
  - Percept aggregators
    - Single percepts → complex percepts.
    - Complex percepts → complex percepts.

CERA-CRANIUM (XIII)

- CRANIUM processor types

  - Reactive processors
    - Single percept → simple behavior.
    - Complex percept → simple behavior.

CERA-CRANIUM (XIV)

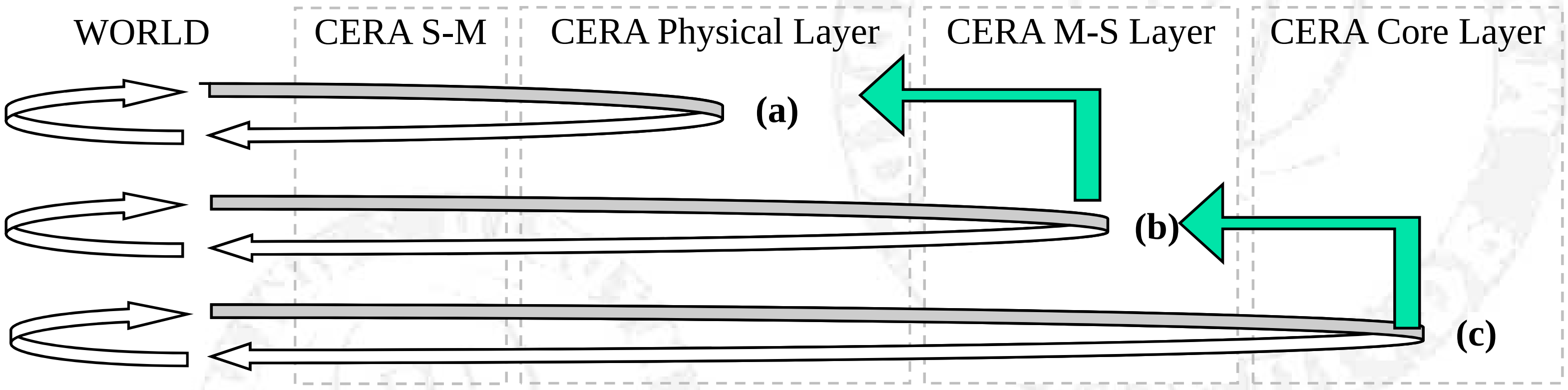
- CRANIUM processor types
  - Action planners
    - Simple behavior → atomic actions.

CERA-CRANIUM (XV)

- CRANIUM processor types
  - Sensory predictors
    - Single percepts → mismatch complex percept.
    - Complex percepts → mismatch complex percept.

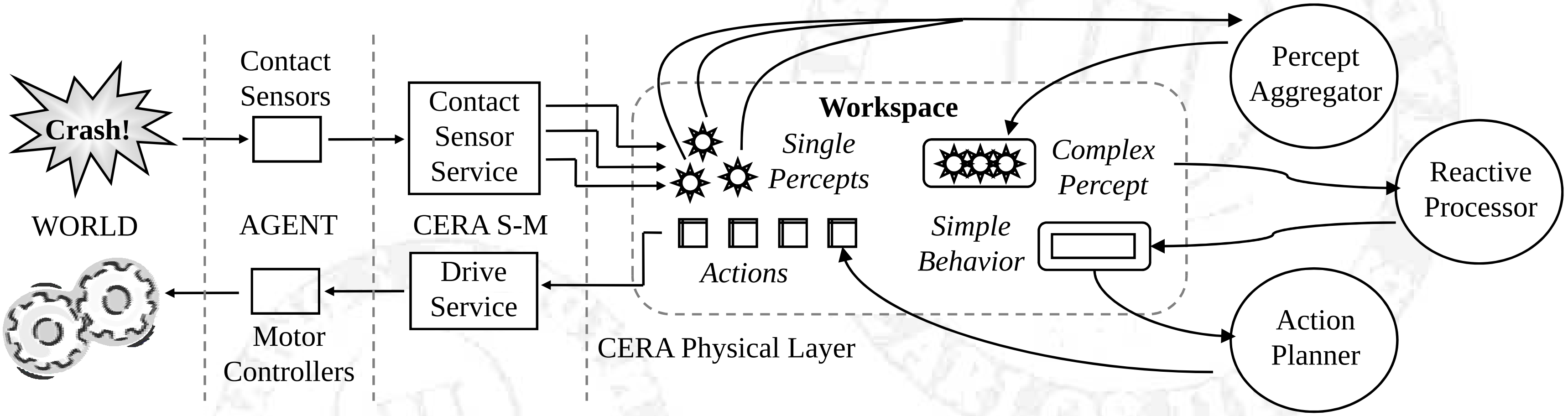
CERA-CRANIUM (XVI)

- Multi-level concurrent feedback loops



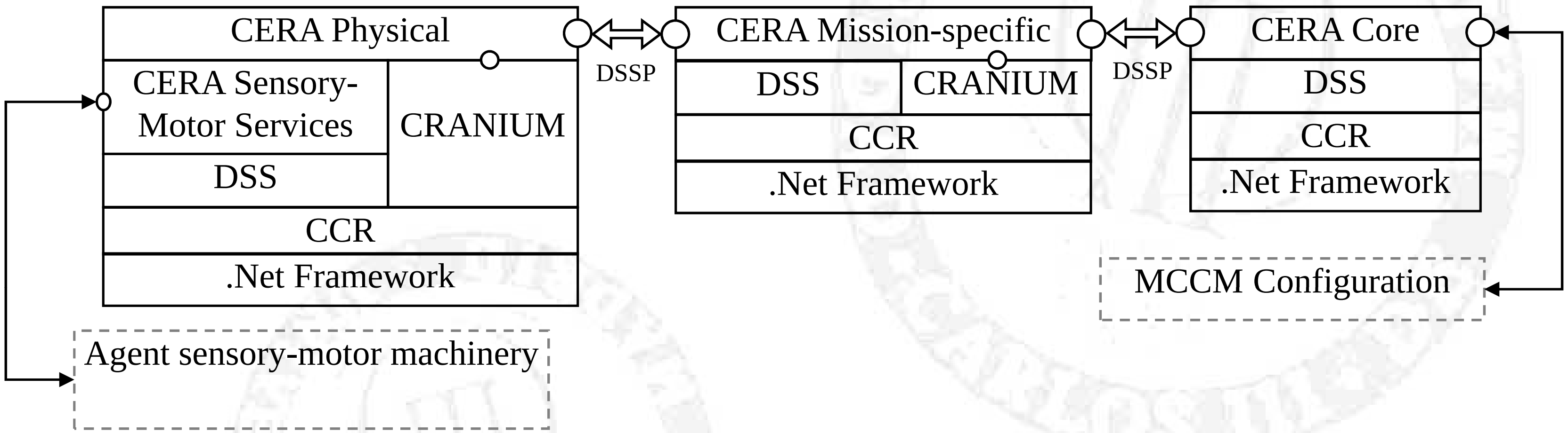
CERA-CRANIUM (XVII)

- Physical level feedback loops



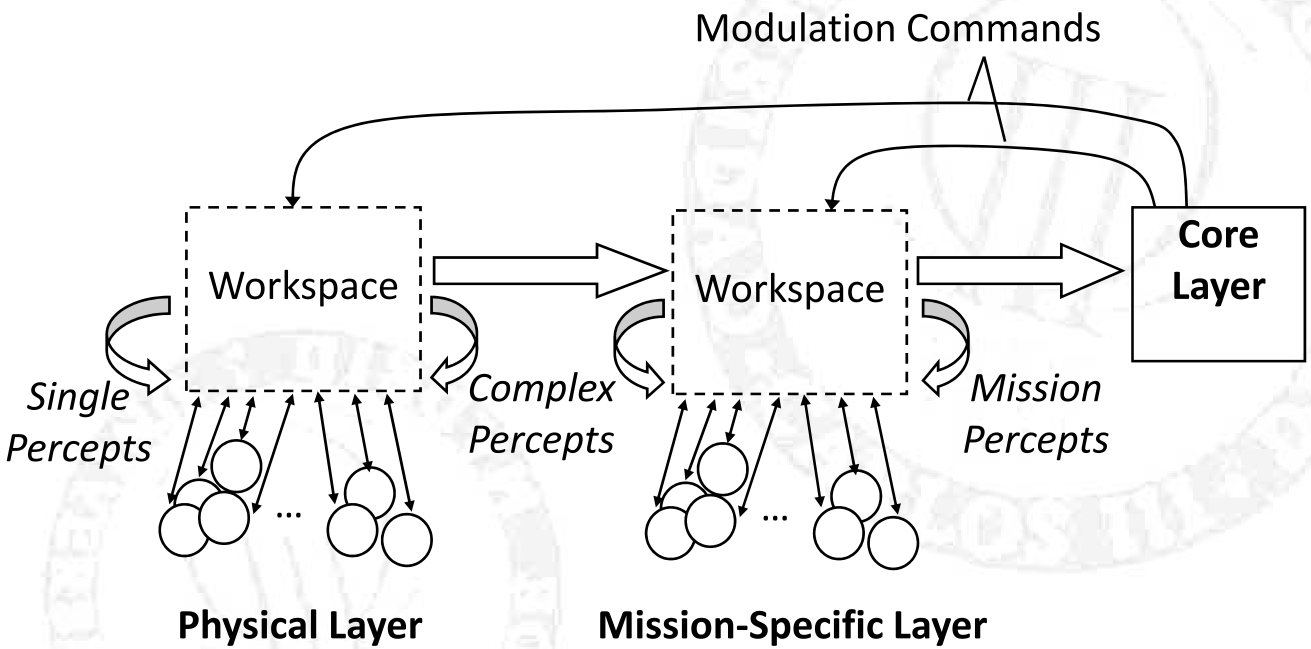
CERA-CRANIUM (XVIII)

- Software architecture



CERA-CRANIUM (XVIIIbis)

- Software architecture



CERA-CRANIUM (XIX)

- The proposed cognitive architecture is focused on:
  - Selecting next content of conscious perception.
  - Selecting next action to be executed.

Experimentation settings

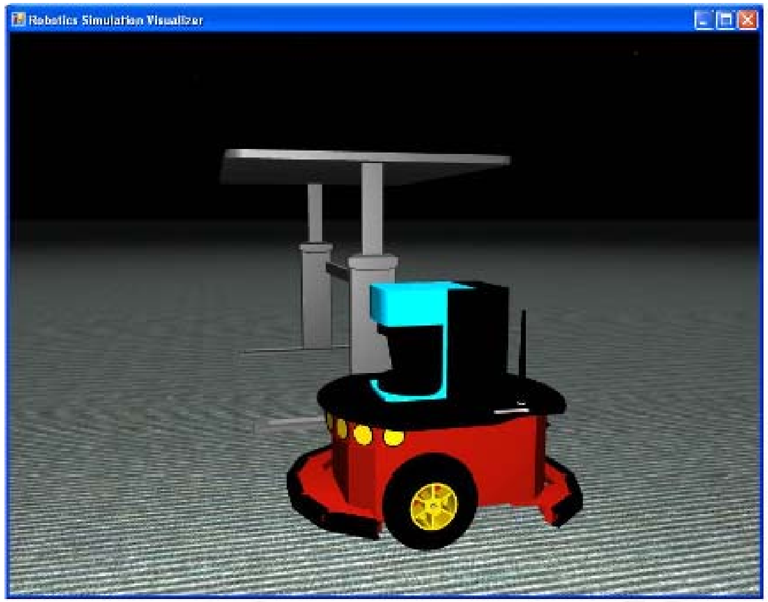
- Variables:
  - Cognitive model of consciousness.
  - Agent (physical or simulated).
  - Problem domain and mission.
  - Environment (physical or simulated).

Applications

- Current applications:

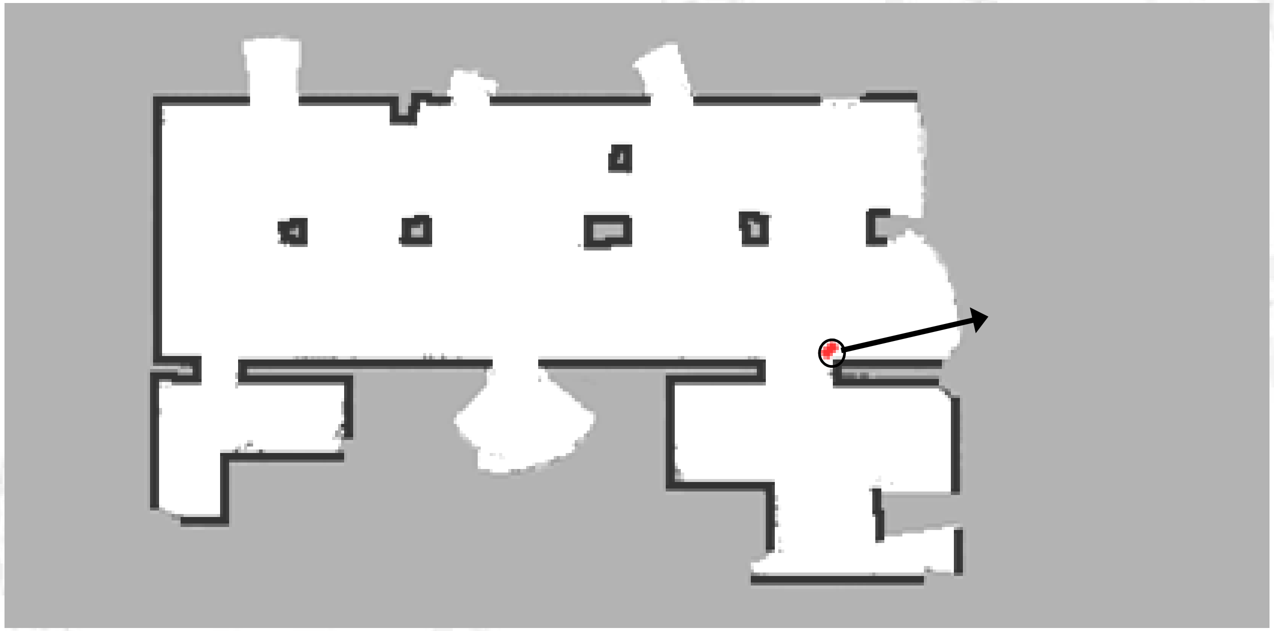
Application example (I)

- Autonomous exploration and mapping



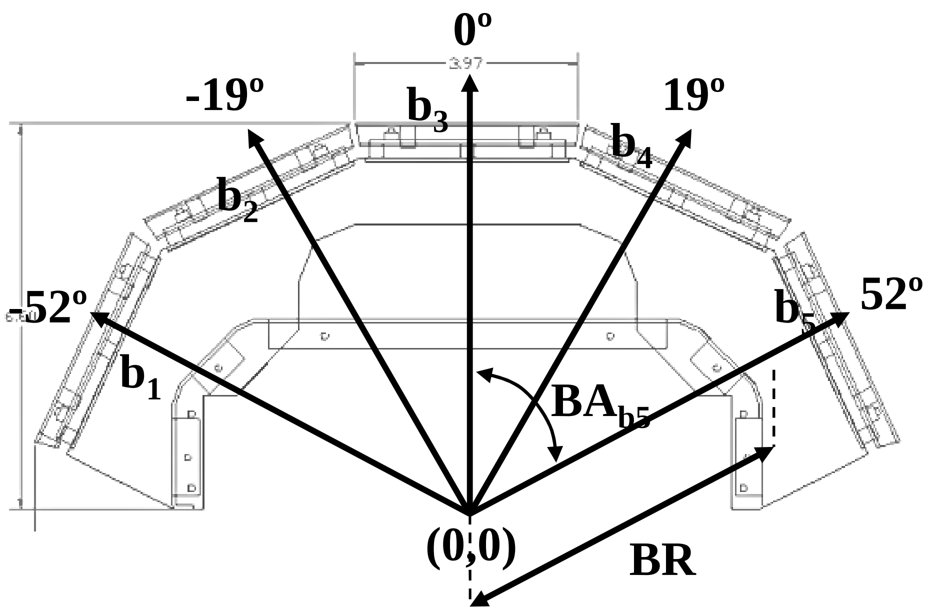
Application example (II)

- Autonomous exploration and mapping.



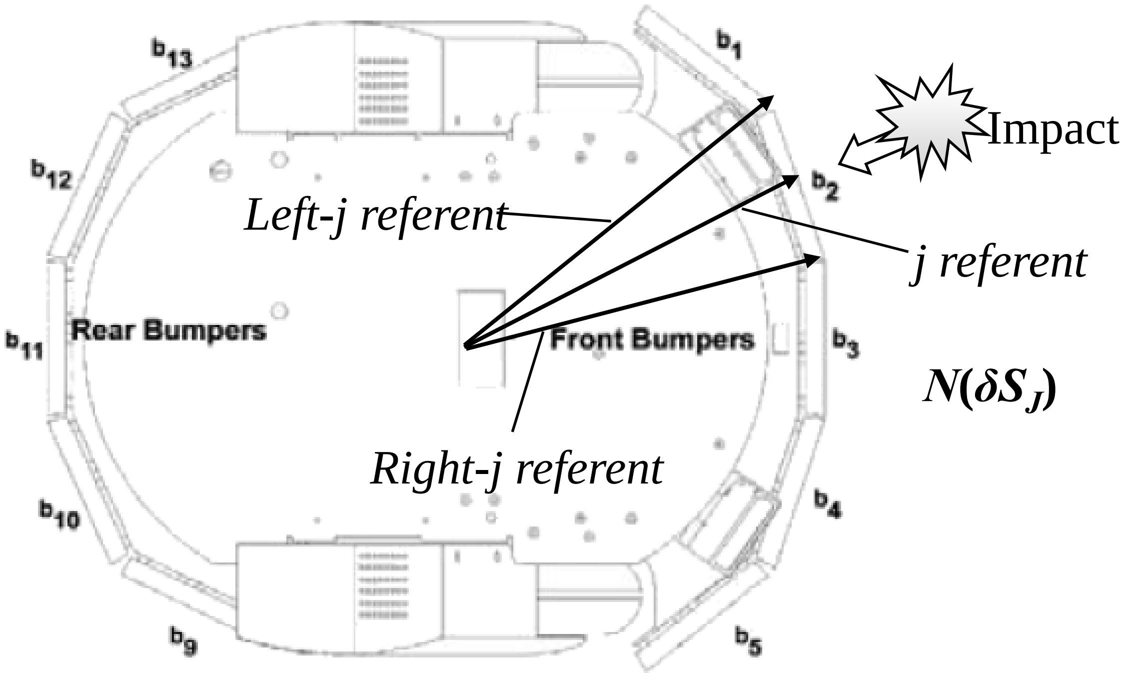
Application example (III)

- Percepts (bumpers).



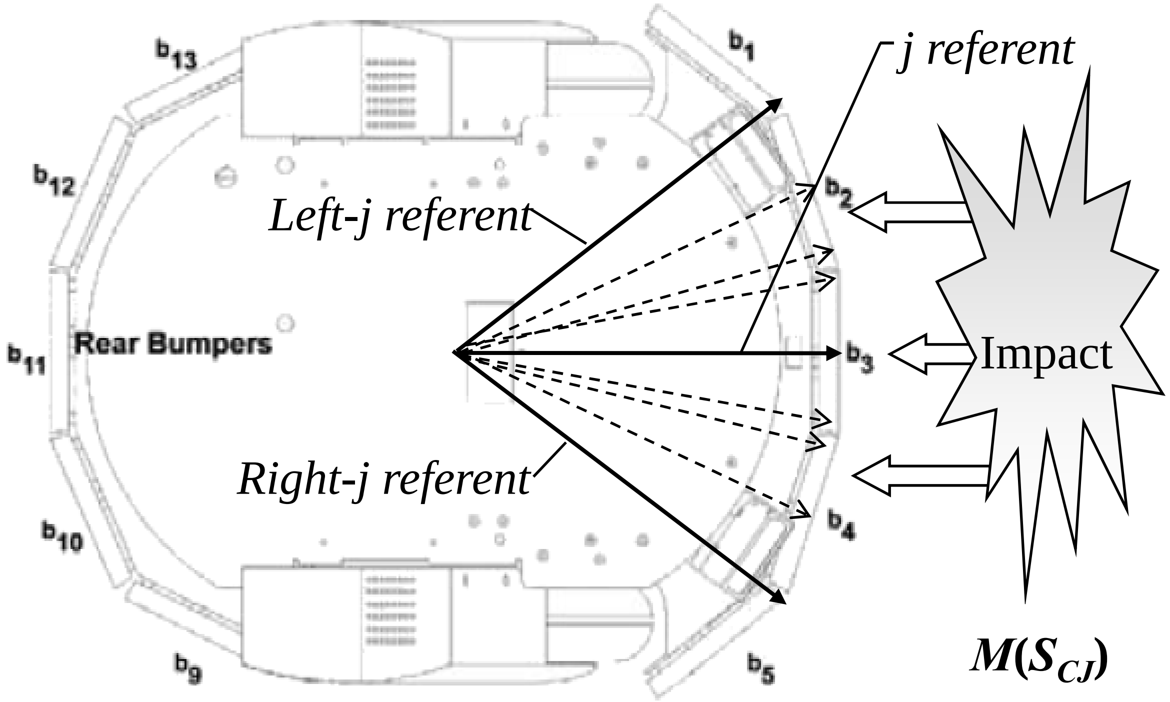
Application example (IV)

- Single percept (bumper).



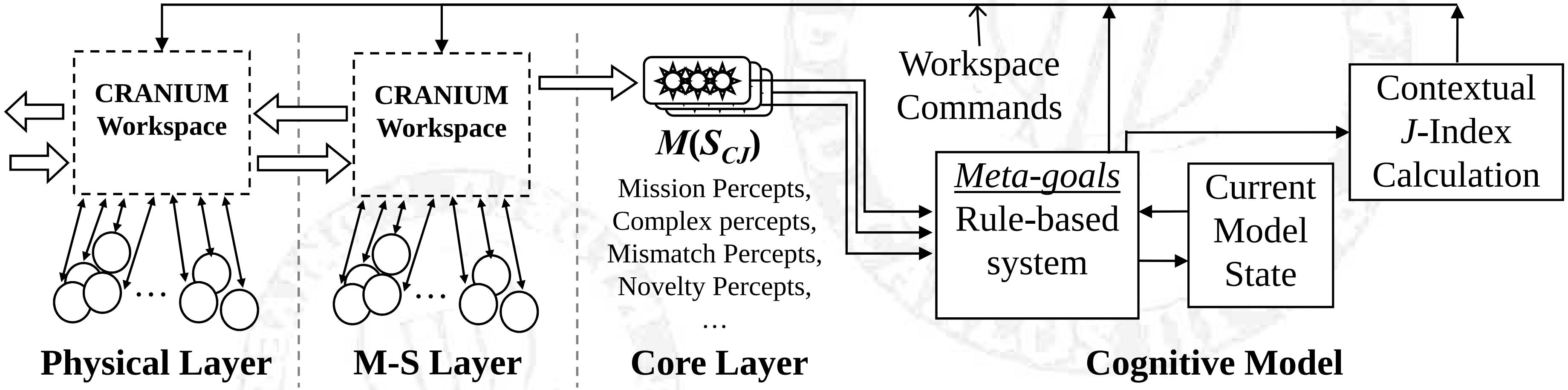
Application example (V)

- Complex percept (bumper).



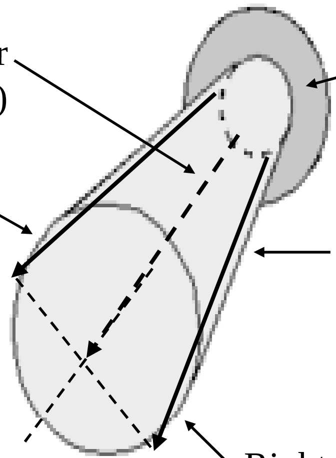
Application example (VI)

- CERA Core Layer Design.



Conclusions (I)

- From sensory to qualia?



Conclusions (II)

- Different implementations of CERA layers.
- Explore workspace modulation techniques.

Thank you