MAS S66

New Destinations in Artificial Intelligence

Goals and Directions for Future Research

joscha@mit.edu

Overview

- *Creating Meaning (see: Sloman, Towards a Gibsonian Model of Vision)*
- *Criteria for understanding Minds (see Sloman, Exploring the Space of Possible Minds*
- *Basic functional organization of a human-like mind*

Questions

- How does vision work?
- What are quintessential functions of the mind?
- Which ones are specific for the human mind?
- Which ones could be different?
- What are the basic modules/structures we need to look at?

Perception and Representation

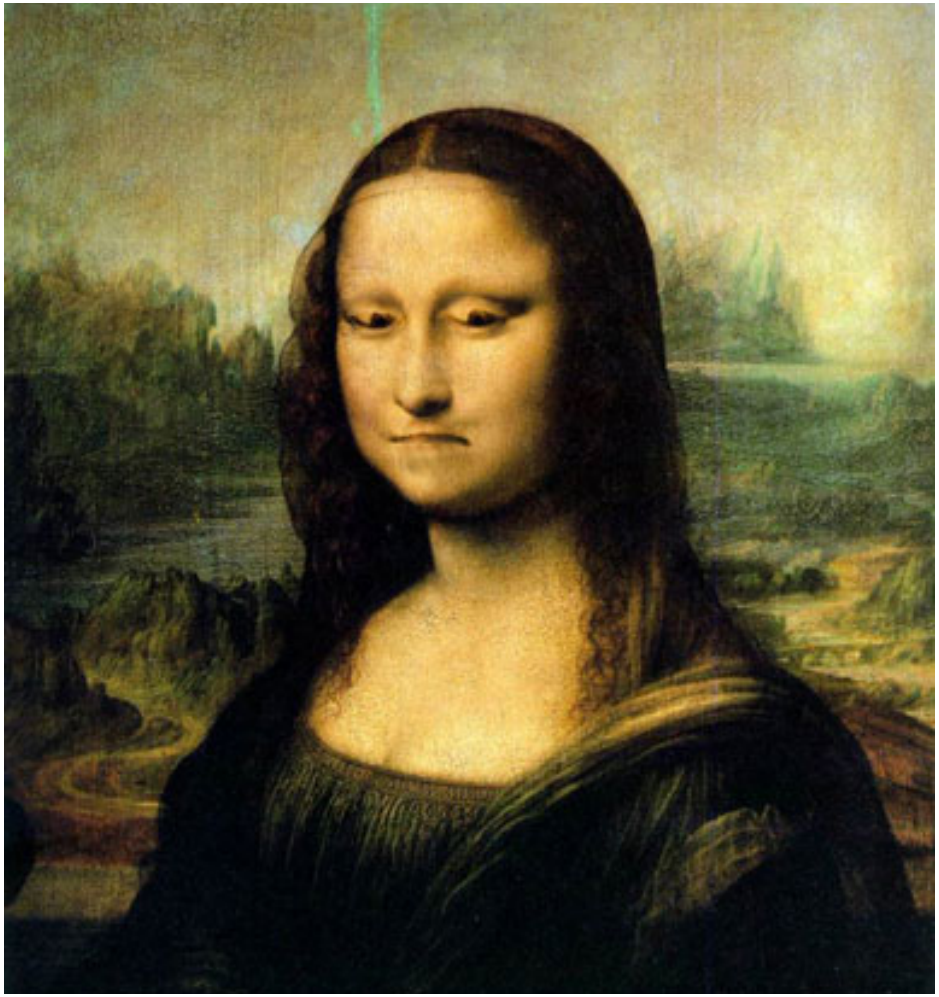
- How are sensory data related to conceptual structures?
- How can conceptual structures be represented?

→ We do not see what we are seeing!

What is perception?

this text is upside down

This txet is upisde donw



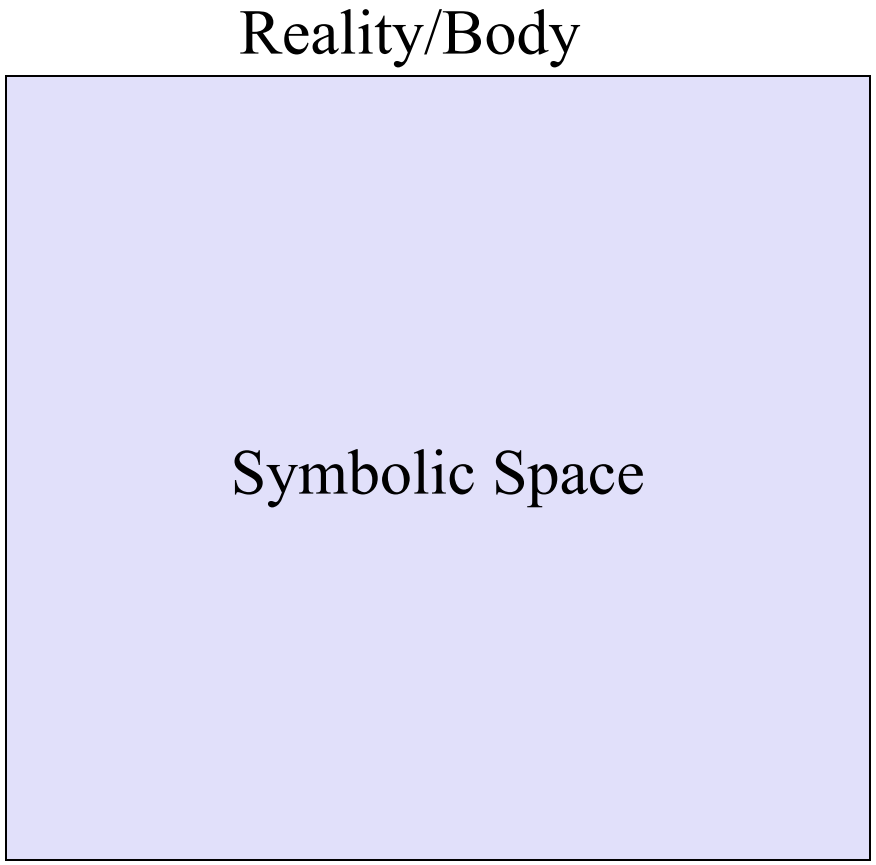
This txet is upisde donw

Thatcher Illusion

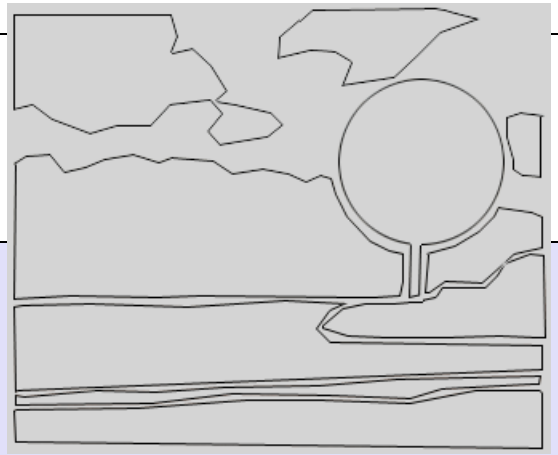
Thatcher Illusion

Thatcher Illusion

The Symbolic Space



The Symbolic Space

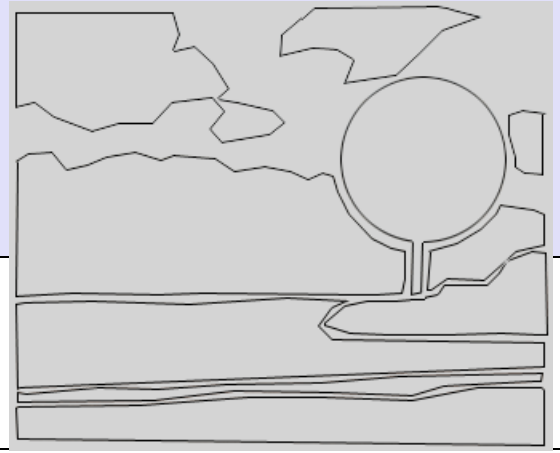


Symbolic Space

Imagination/Mind

The Symbolic Space

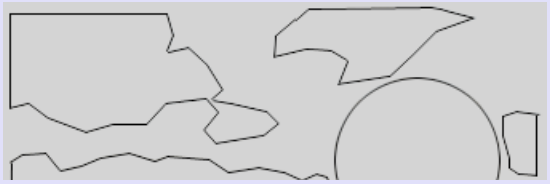
Symbolic Space



The Symbolic Space

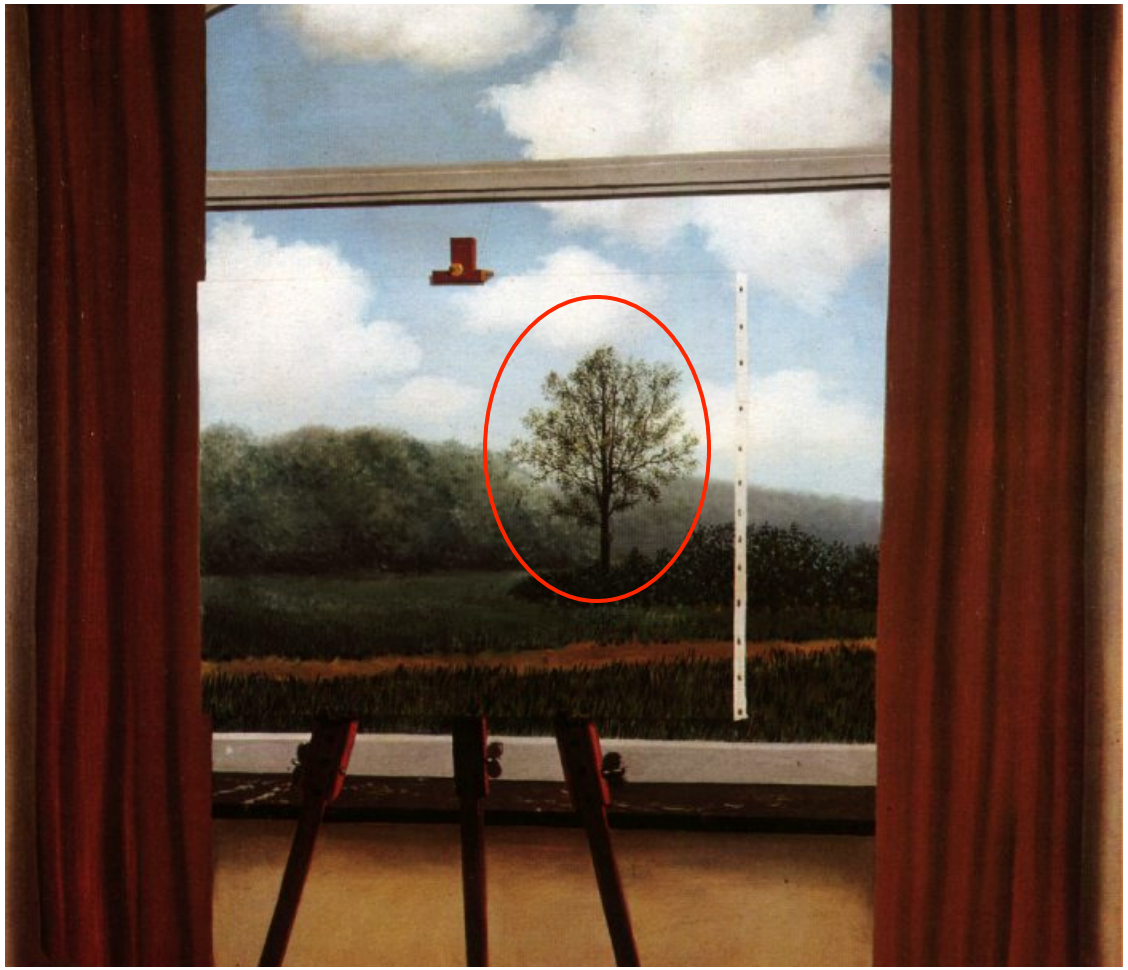
Reality?

Symbolic Space



The Symbolic S

How does perception work?



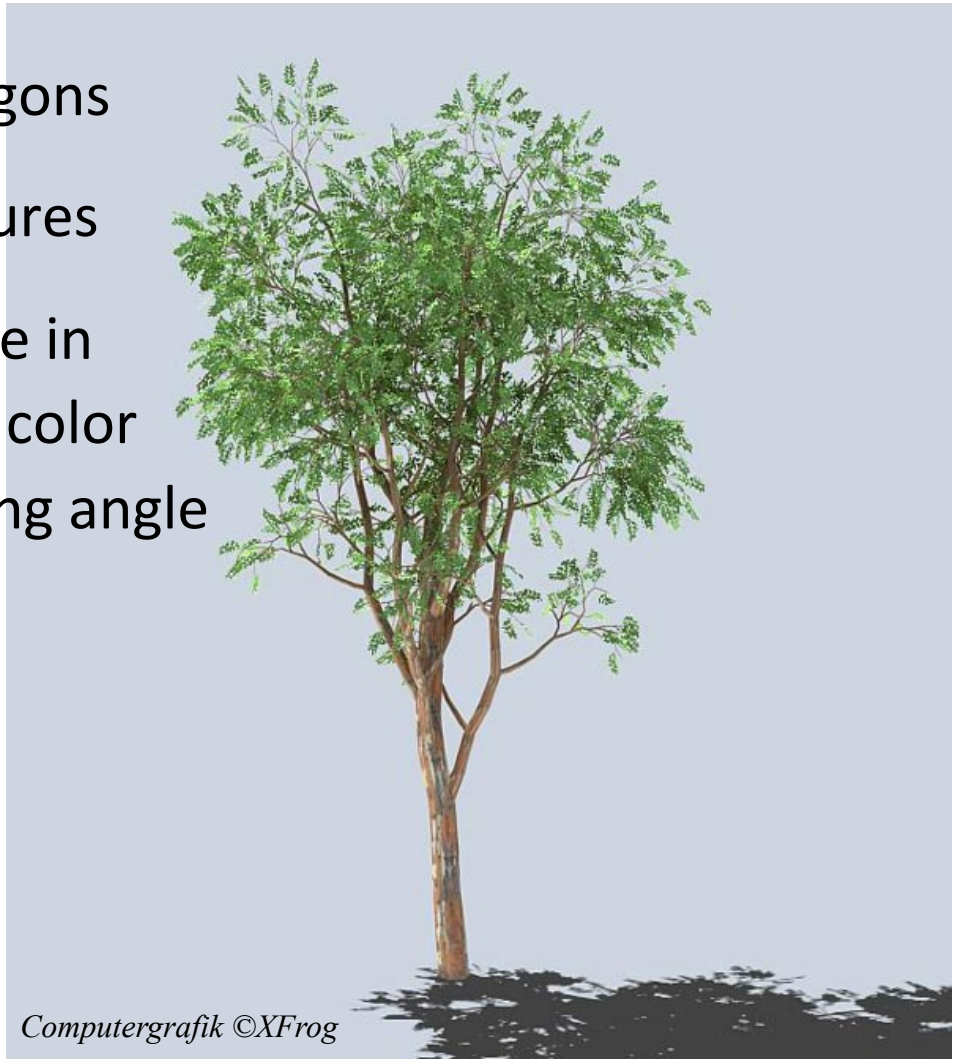
Magritte: *The Human Condition II*

How to recognize a tree?

130.000 Polygons

fractal struktures

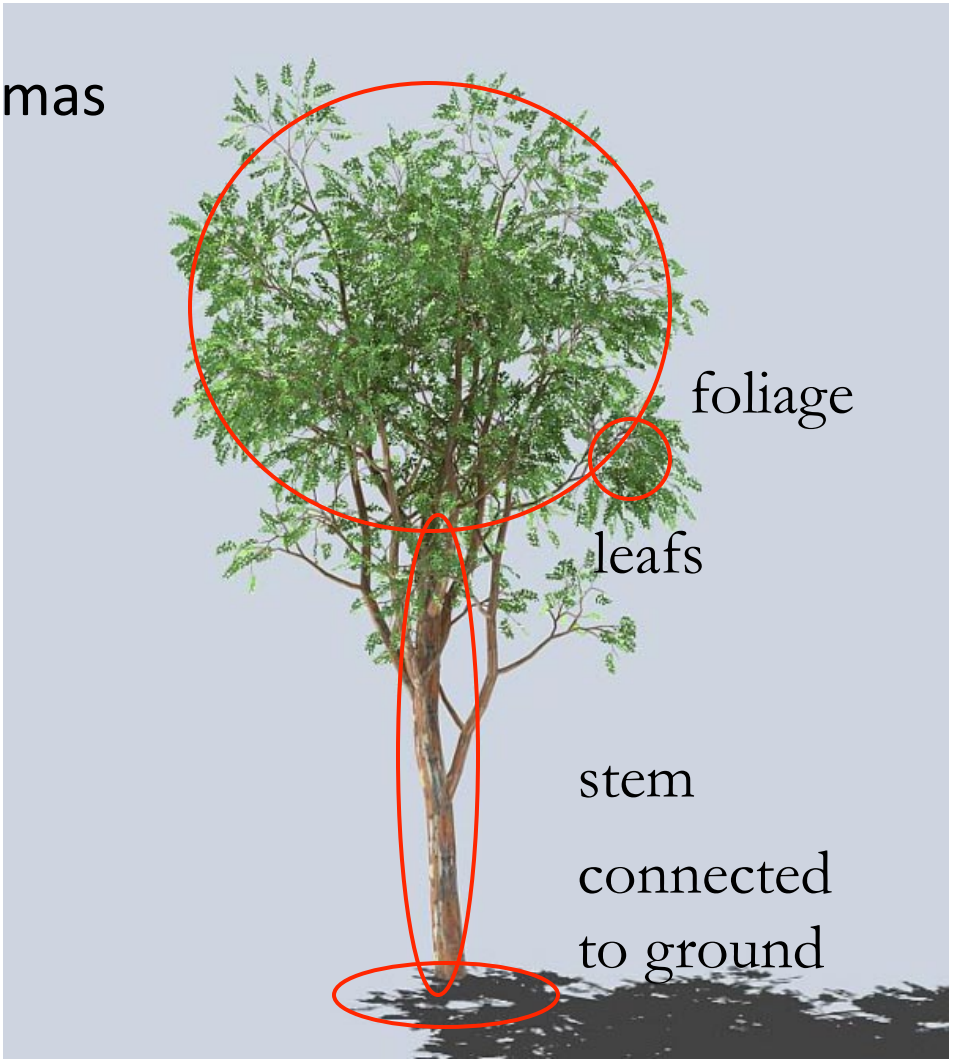
large variance in
- shape and color
- with viewing angle



How to recognize a tree?

How to recognize a tree?

Abstract schemas



Schematic hierarchies

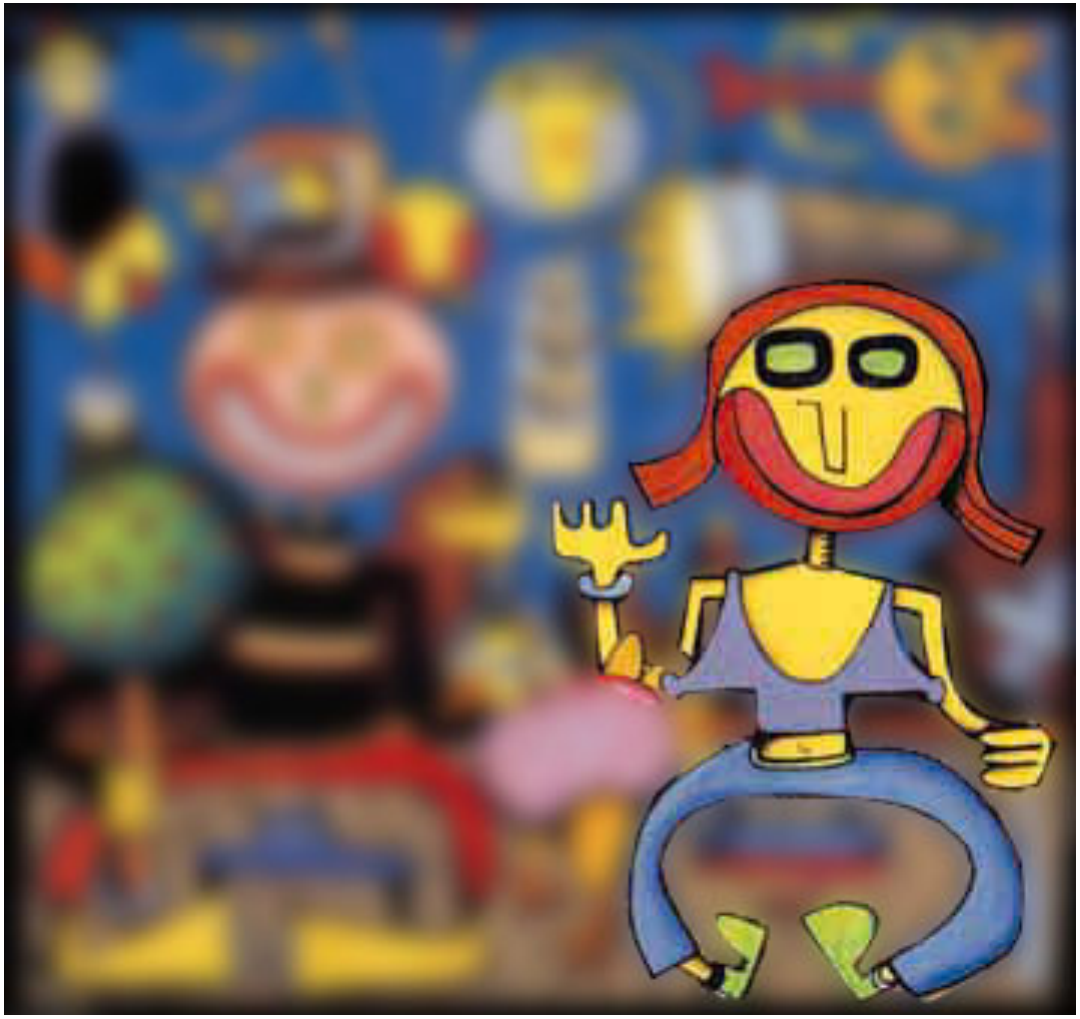
Image recognition tells us something about perceptual mechanisms

How do we perceive reality?

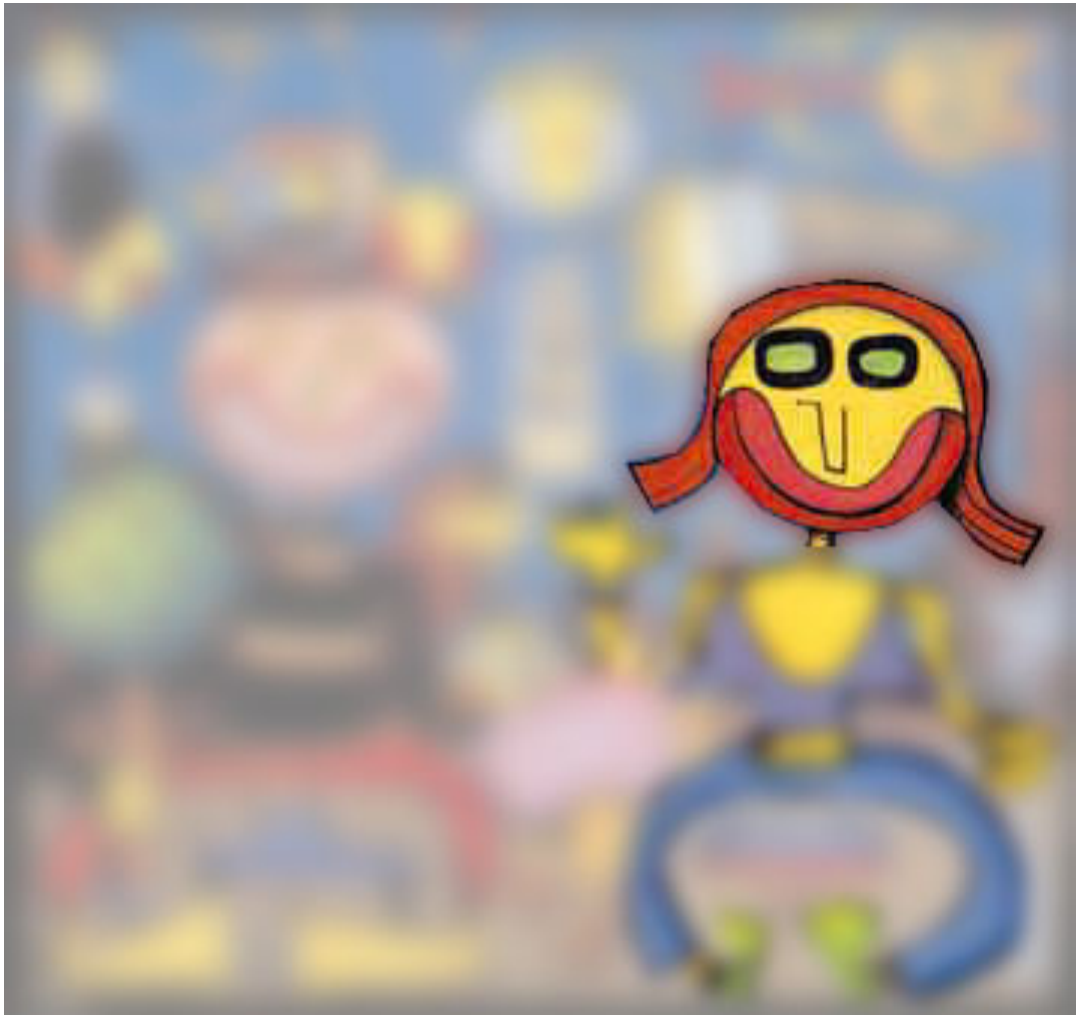
Jim Avignon: “”All my friends”

?

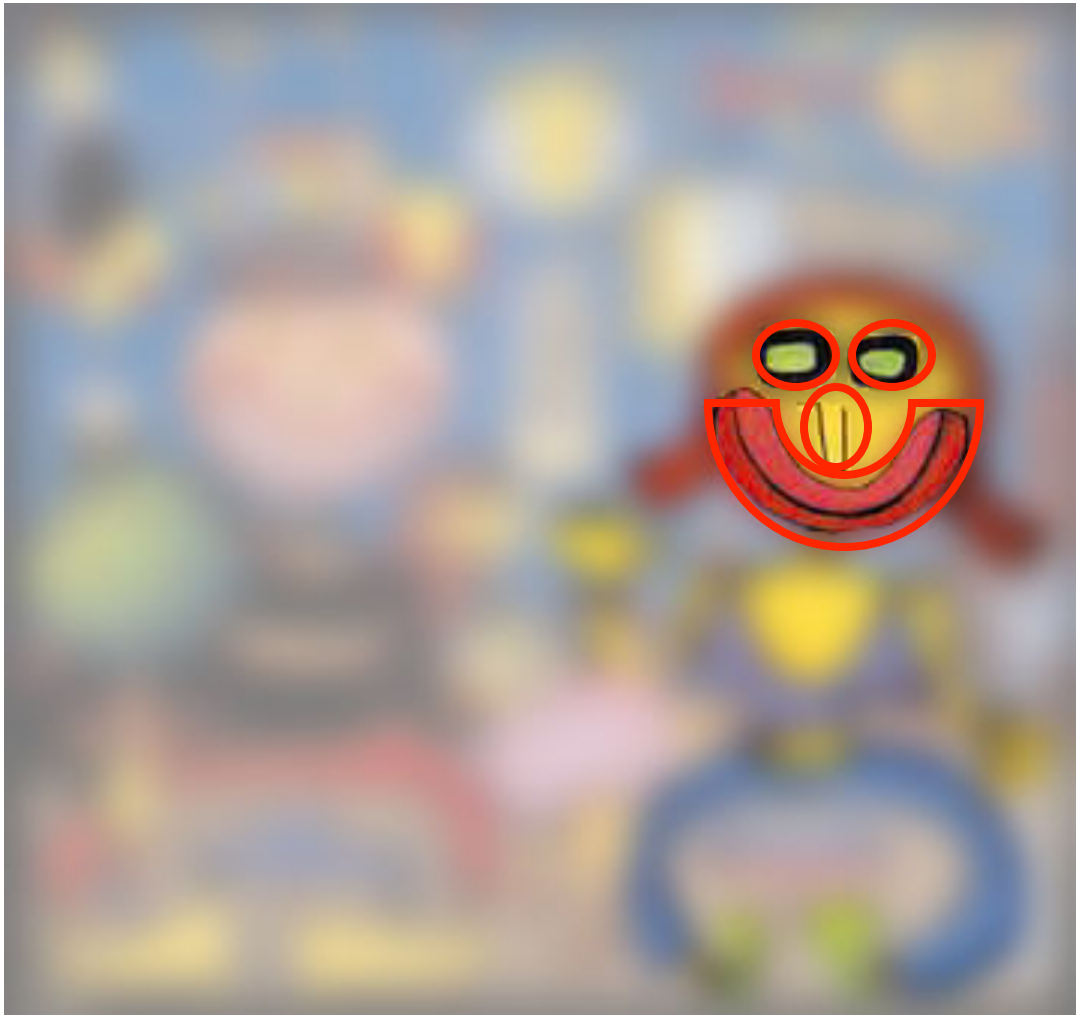
How do we perceive reality?



How do we perceive reality?

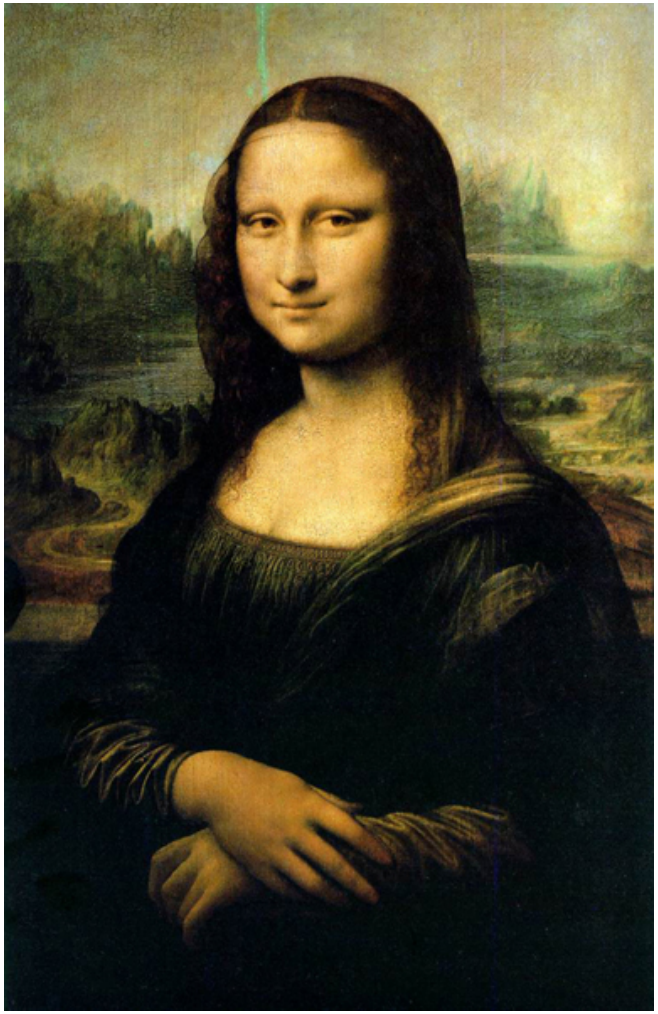


How do we perceive reality?

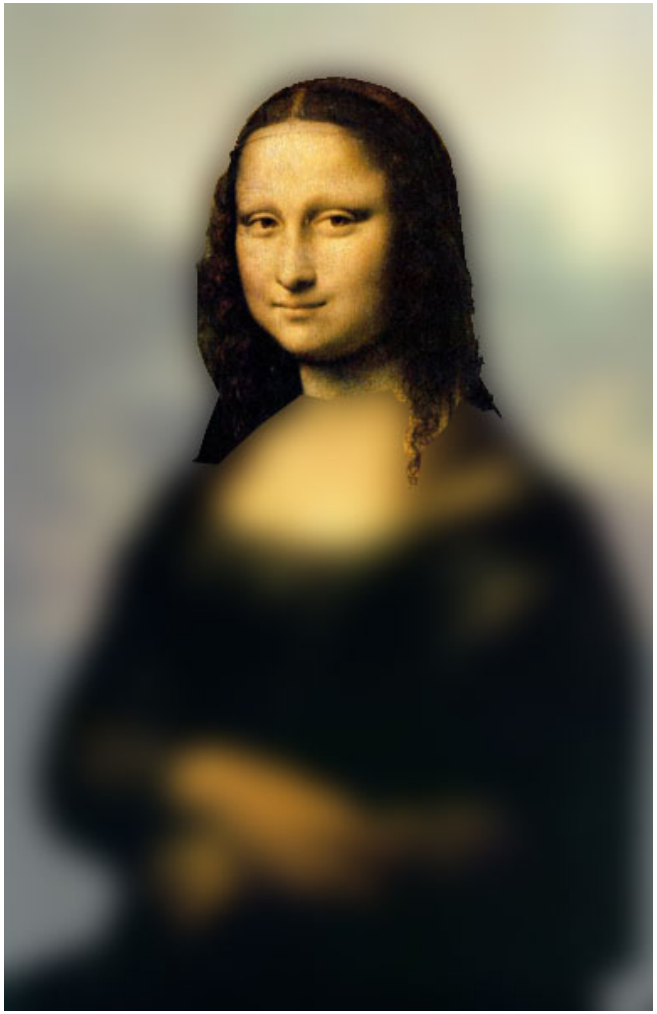


face schema

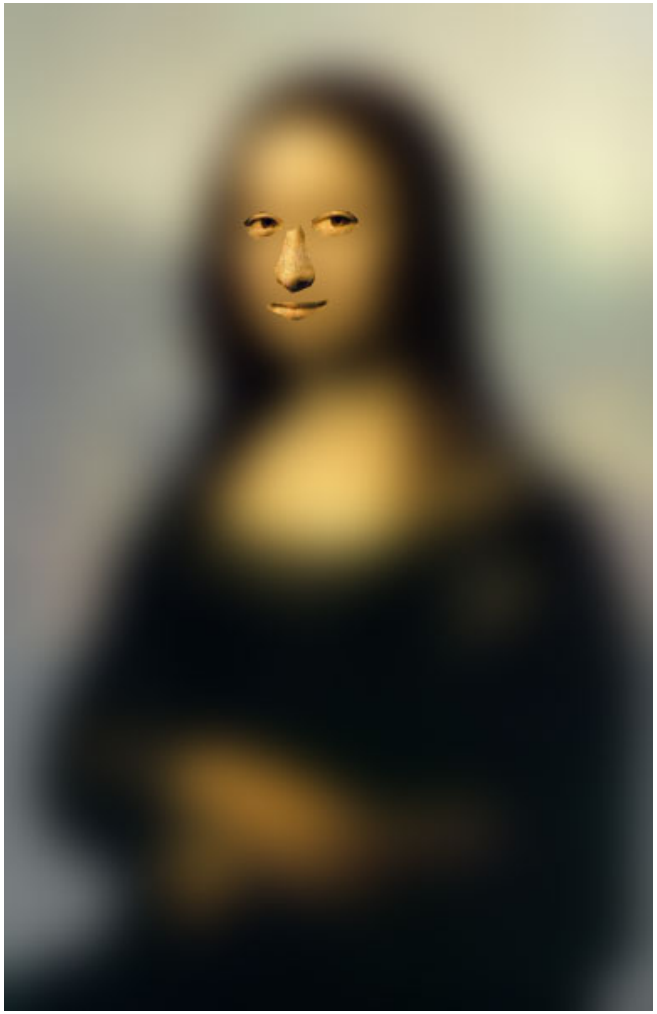
How do we perceive reality?



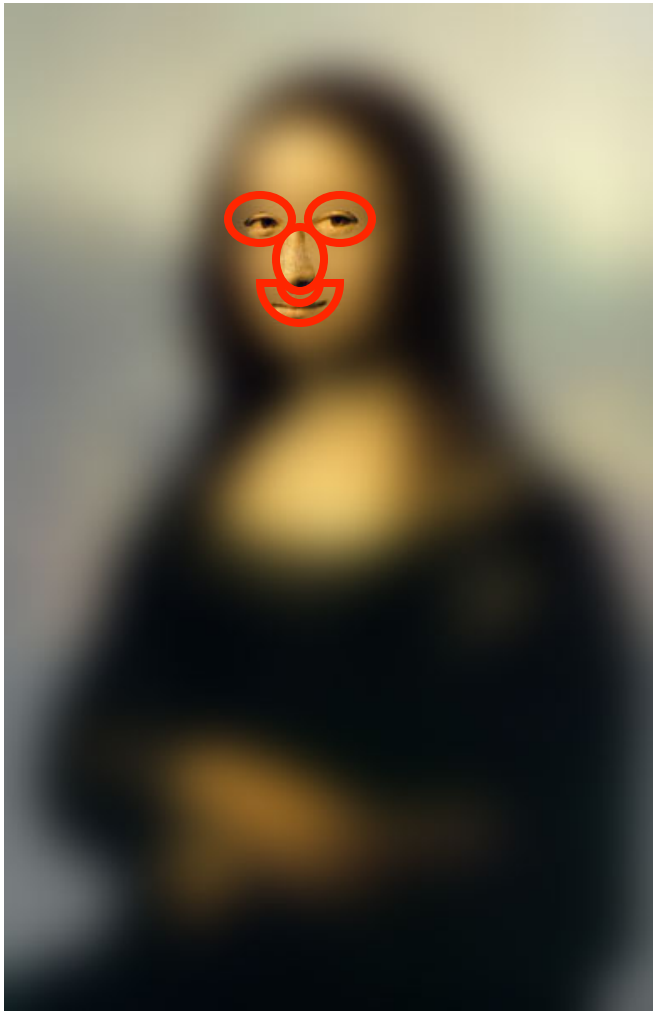
How do we perceive reality?



How do we perceive reality?

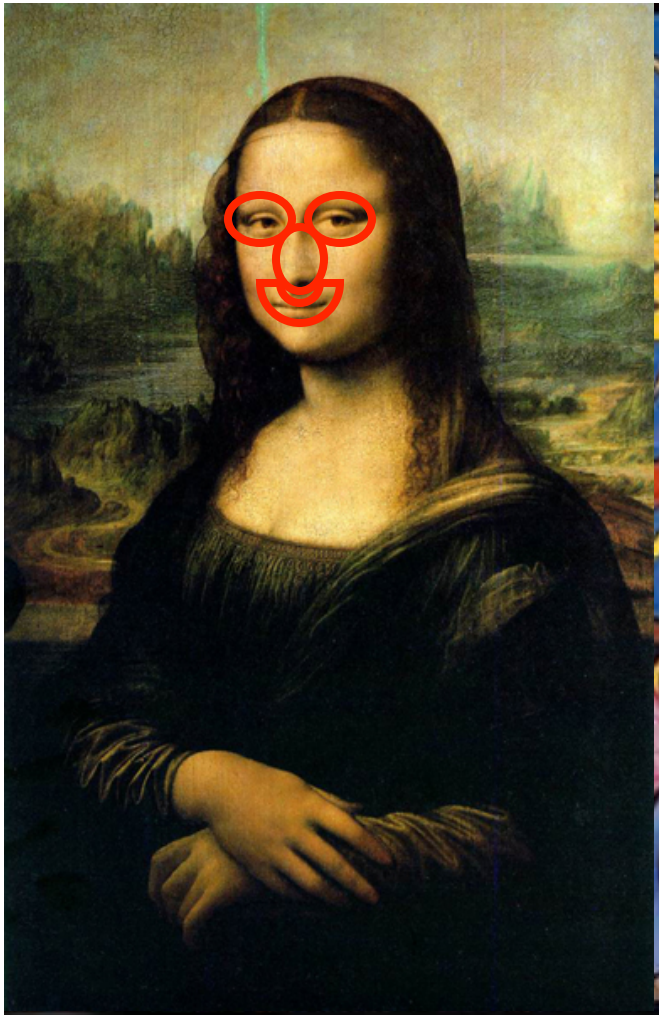


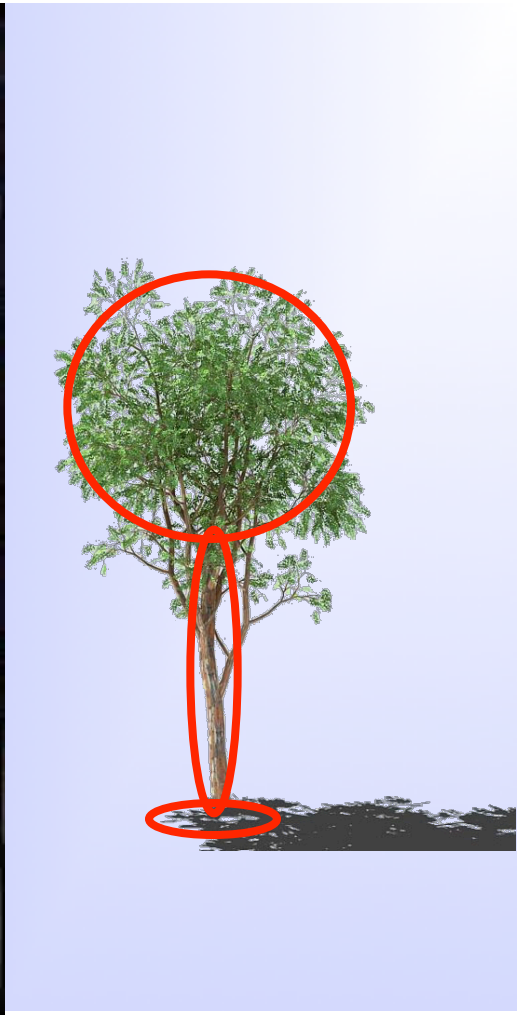
How do we perceive reality?



face schema

How do we perceive reality?





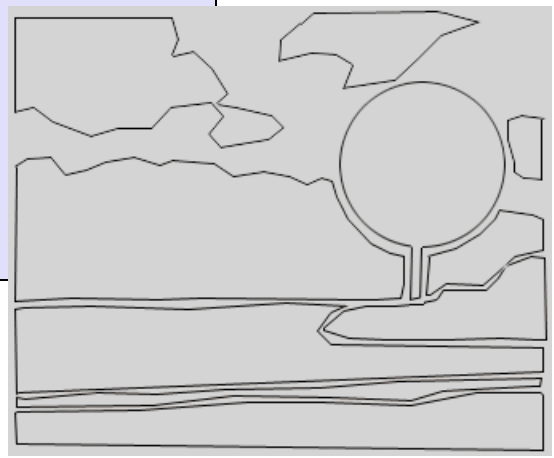
The symbolic space

? ? ? ? ?

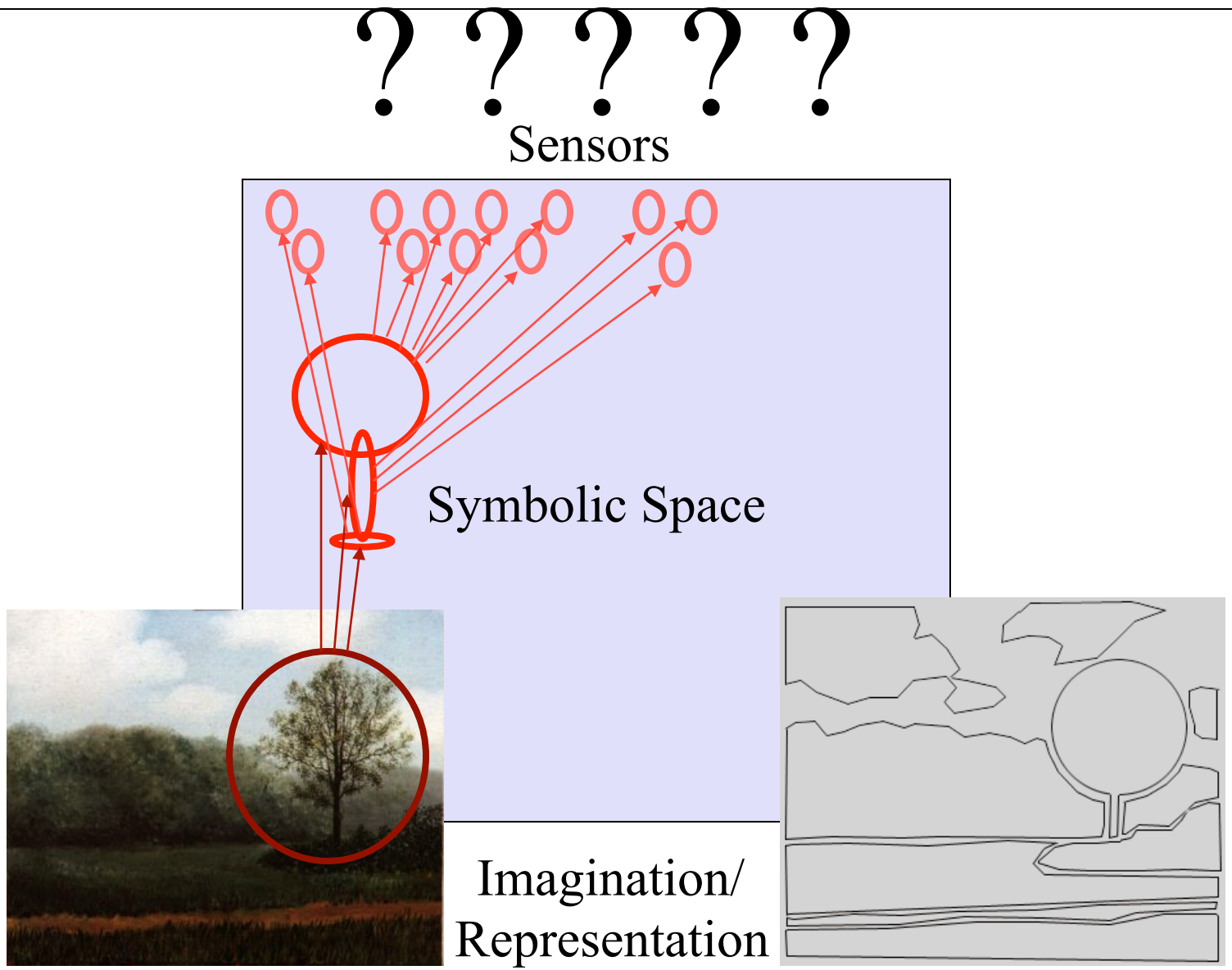
Sensors

Symbolic Space

Imagination/
Representation



The symbolic space



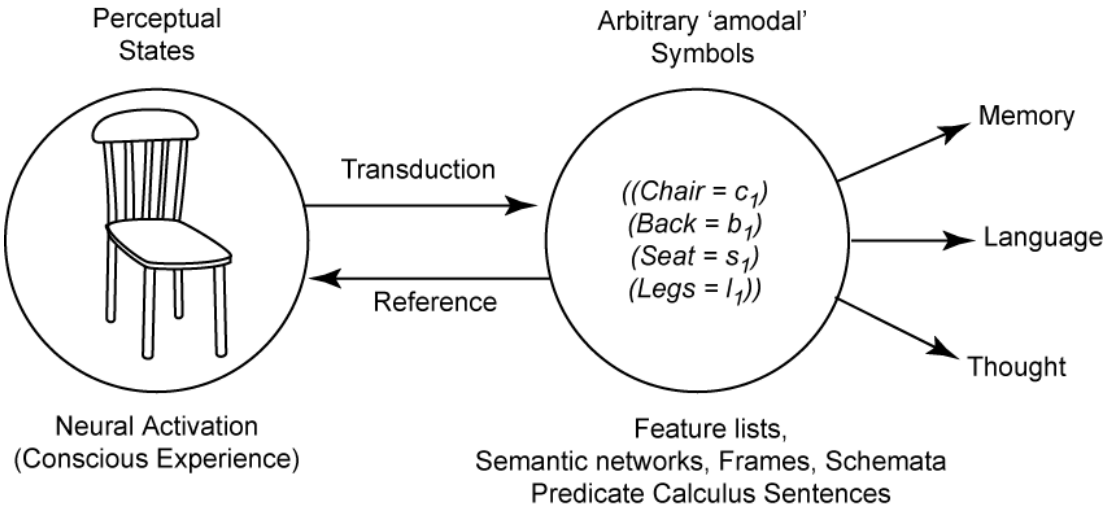
Perception

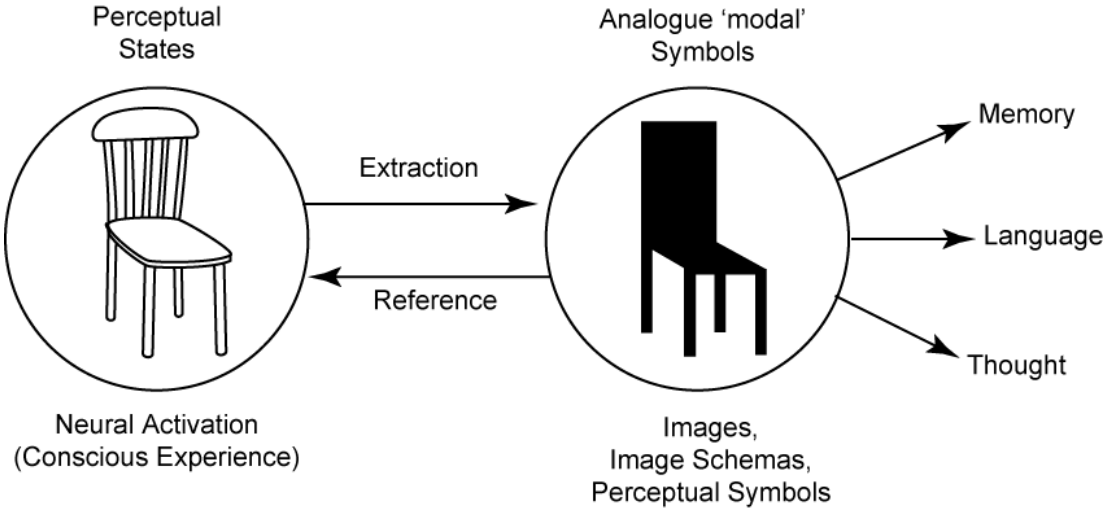
→ Perception is not simply constructed from receptor input (bottom up)

→ Verification of hypothesis via sensors

→ No difference between perceptual and imaginative datastructures

Modal vs. amodal representation (Barsalou 99)





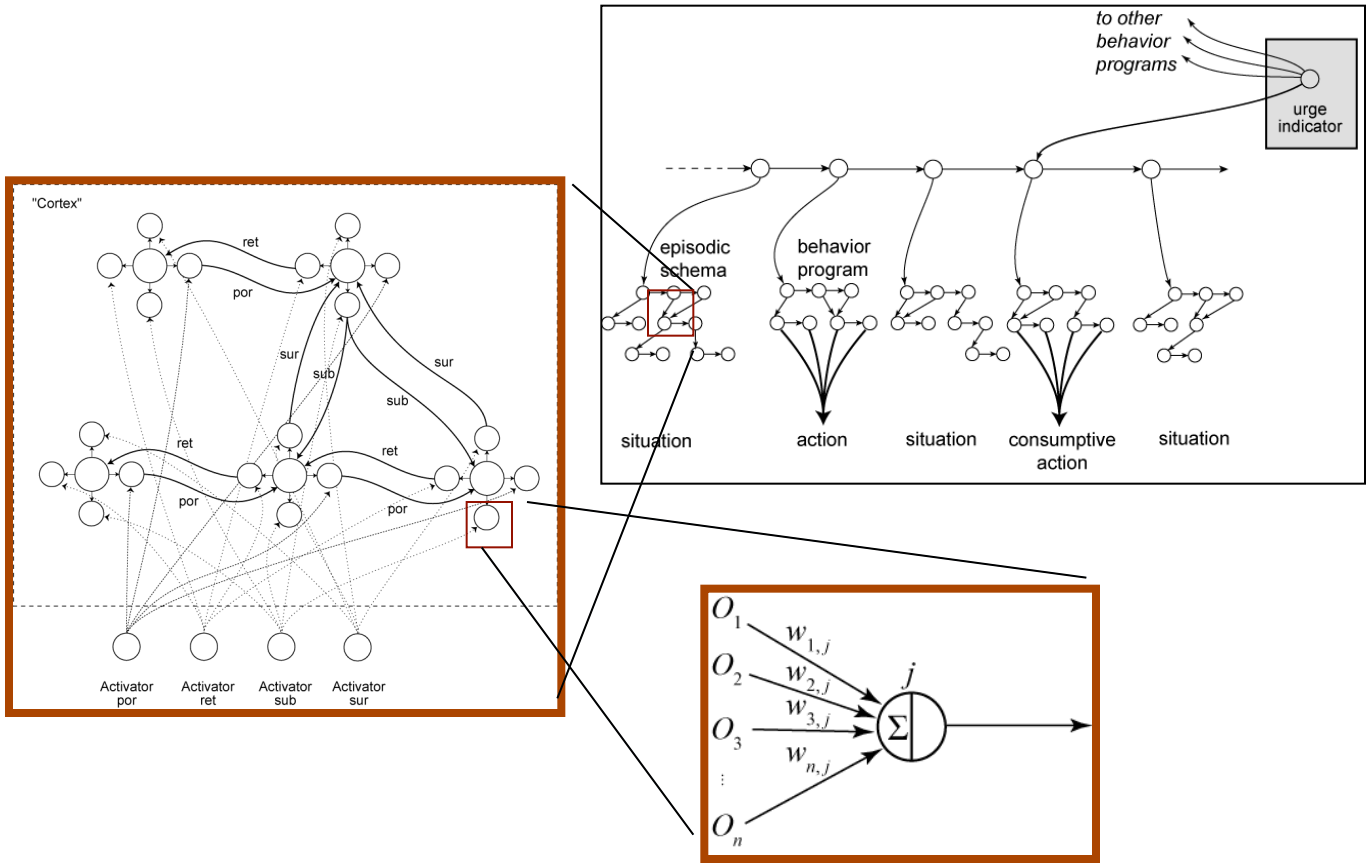
Cognitive Artificial Intelligence

Methods should focus on components and performances necessary for intelligence:

- **Whole, testable architectures**
- **Universal Representations:**
  Grounded neuro-symbolic representations (integrate both symbolic and distributed aspects)
- **(Semi-) Universal Problem Solving:**
  Learning, Planning, Reasoning, Analogies, Action Control, Reflection ...
- **Universal Motivation:**
  Polythematic, adaptive goal identification
- **Emotion and affect**

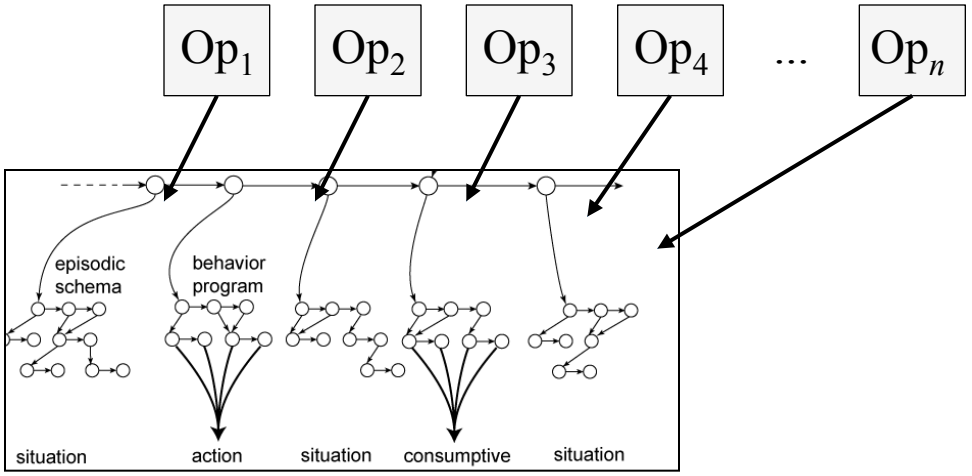
Components for Cognitive AI

- Universal mental representations
  (compositional + distributed → neurosymbolic)



Components for Cognitive AI

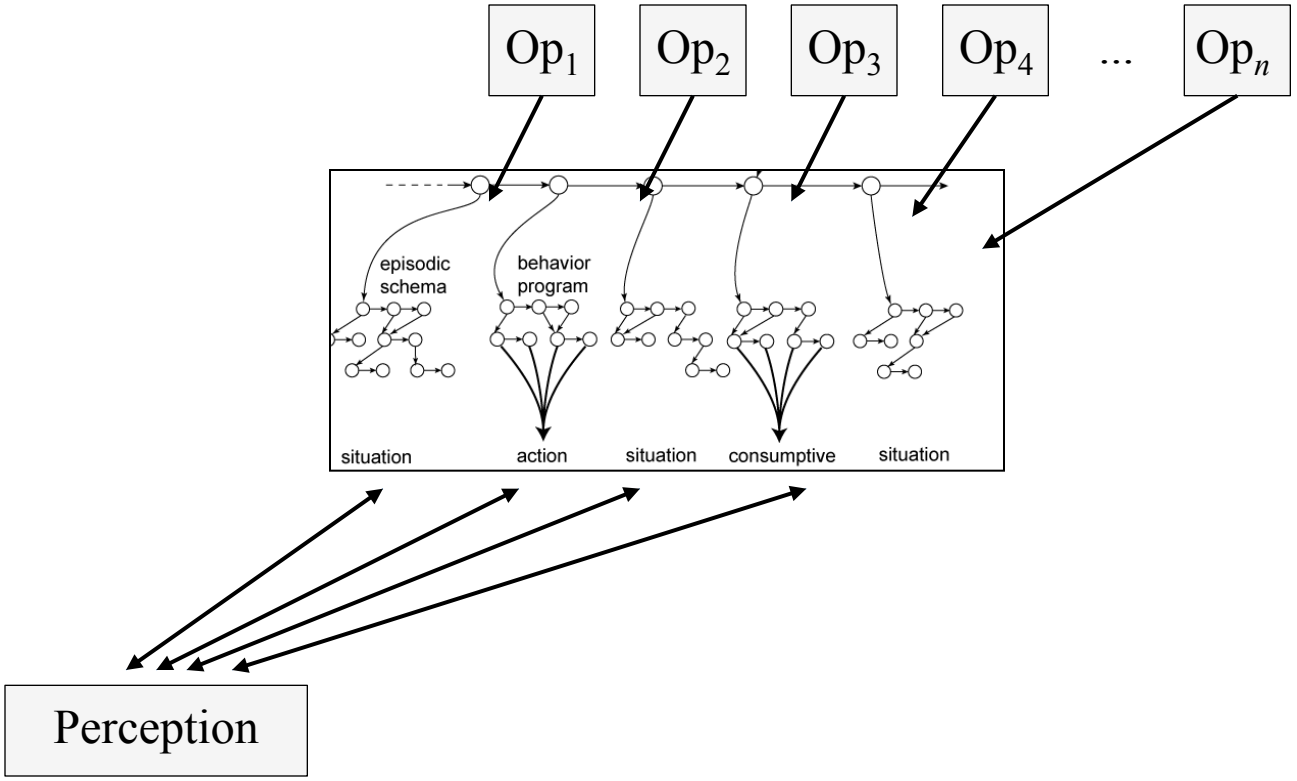
- (Semi-) General problem solving: Operations over these representations



(neural learning, categorization, planning, reflection, consolidation, ...)

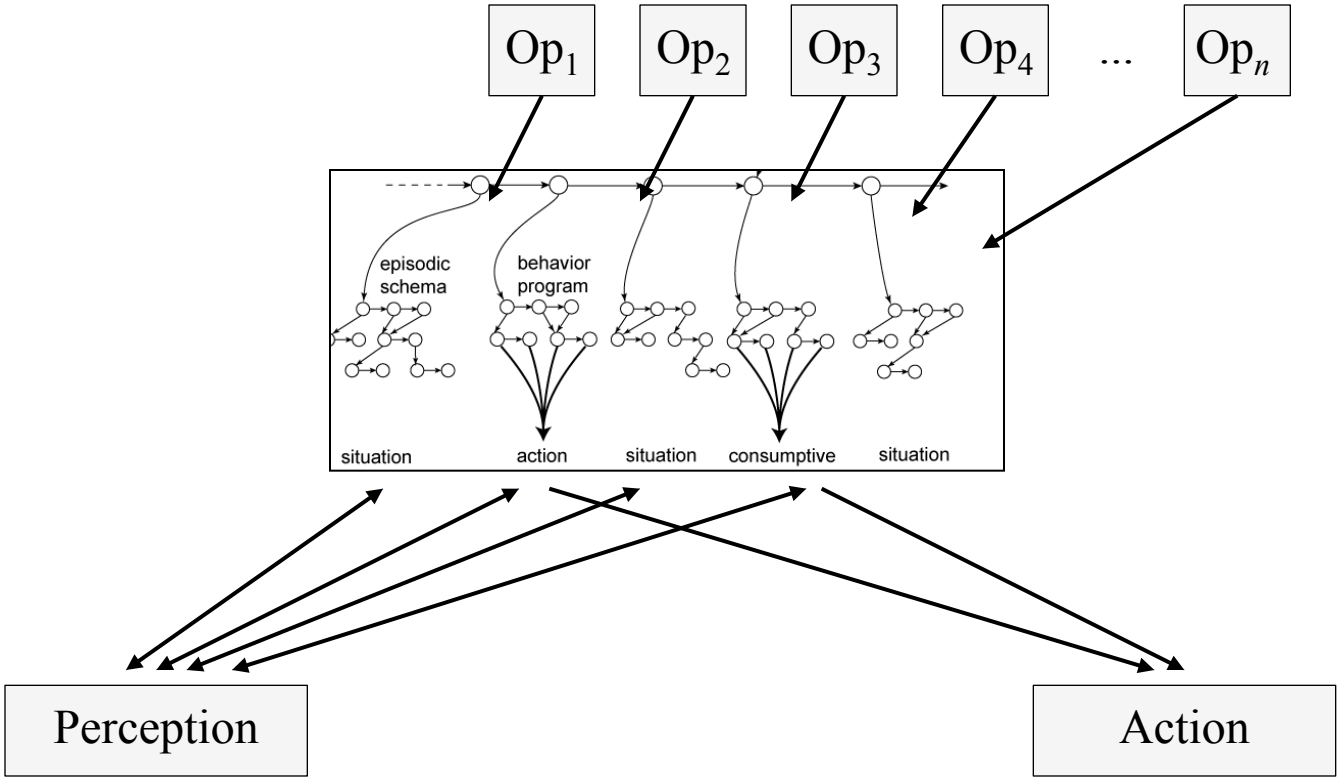
Components for Cognitive AI

- Perceptual grounding



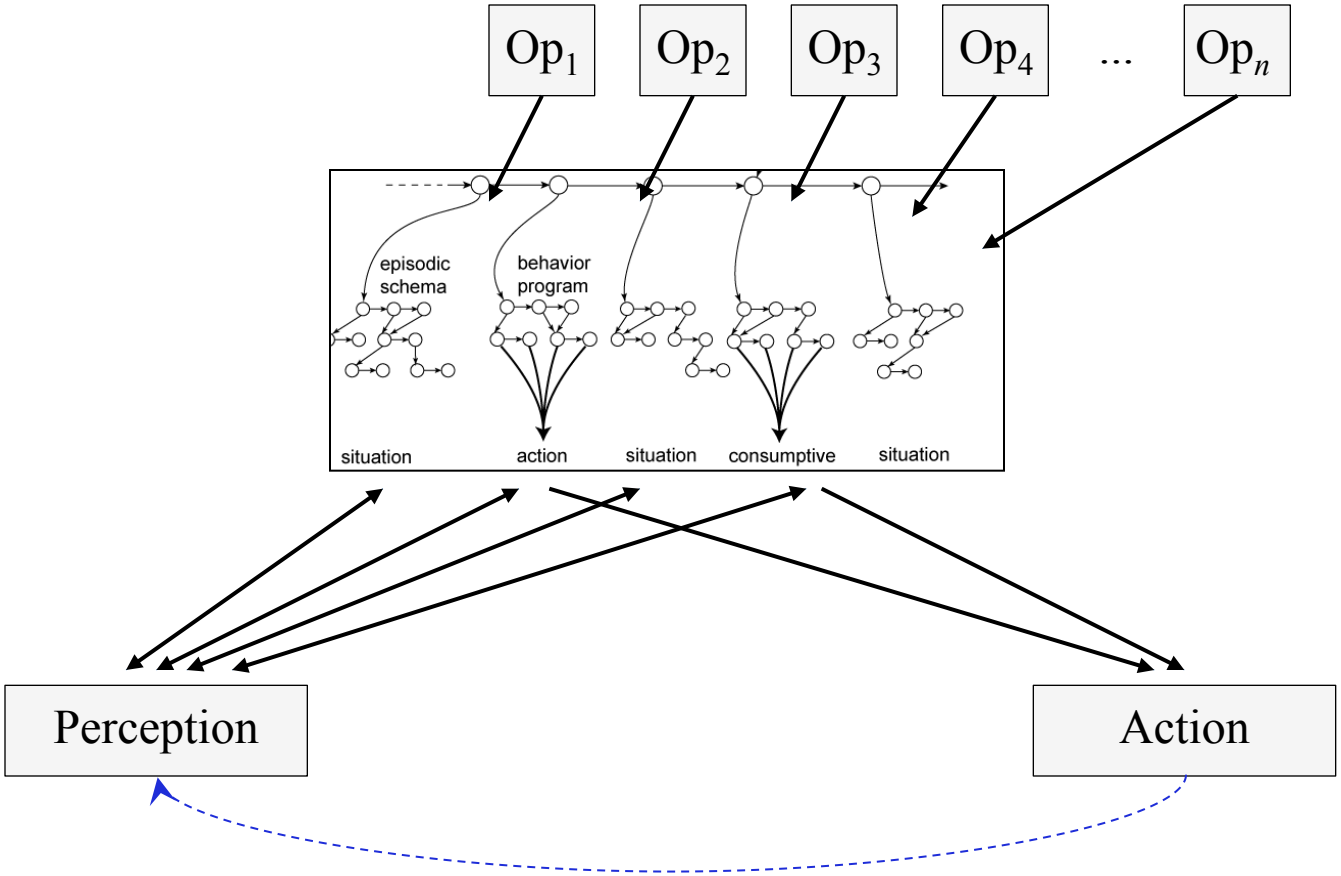
Components for Cognitive AI

- Perceptual grounding and action



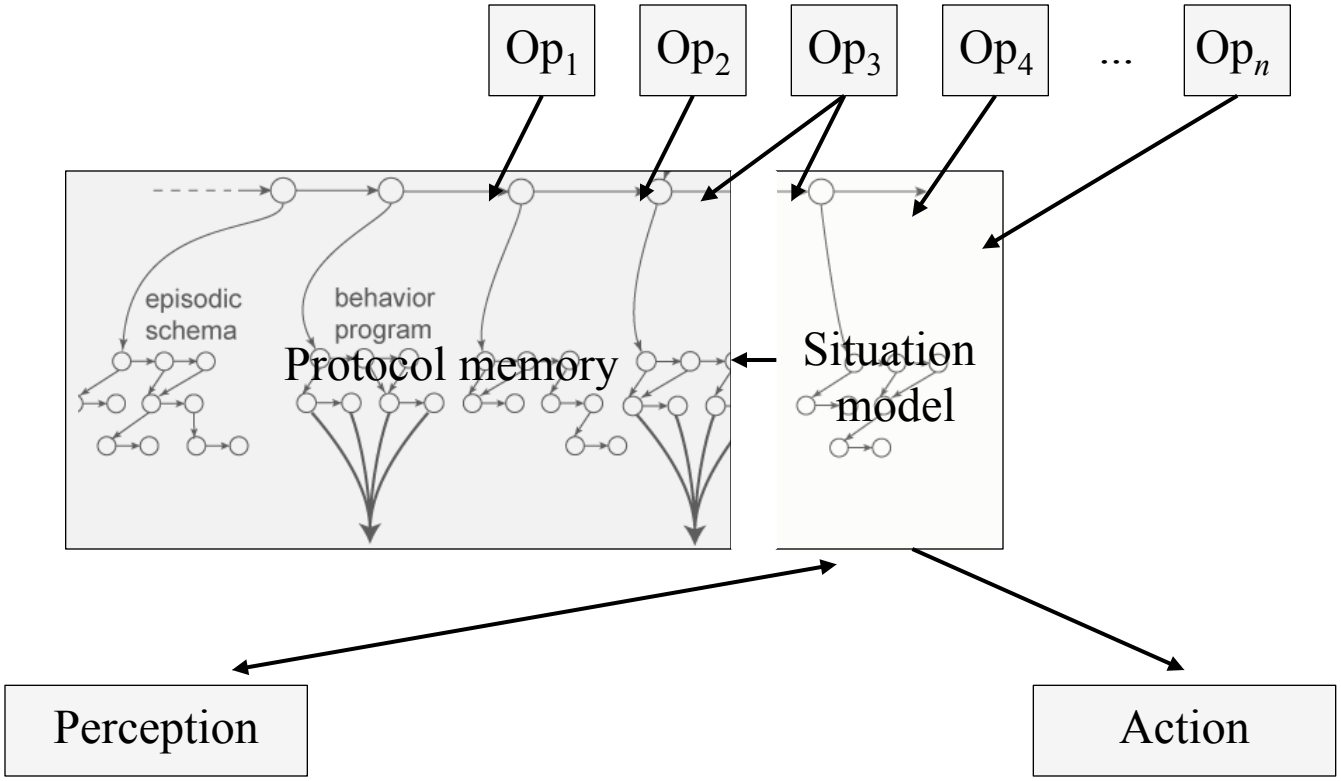
Components for Cognitive AI

- Perceptual grounding and action



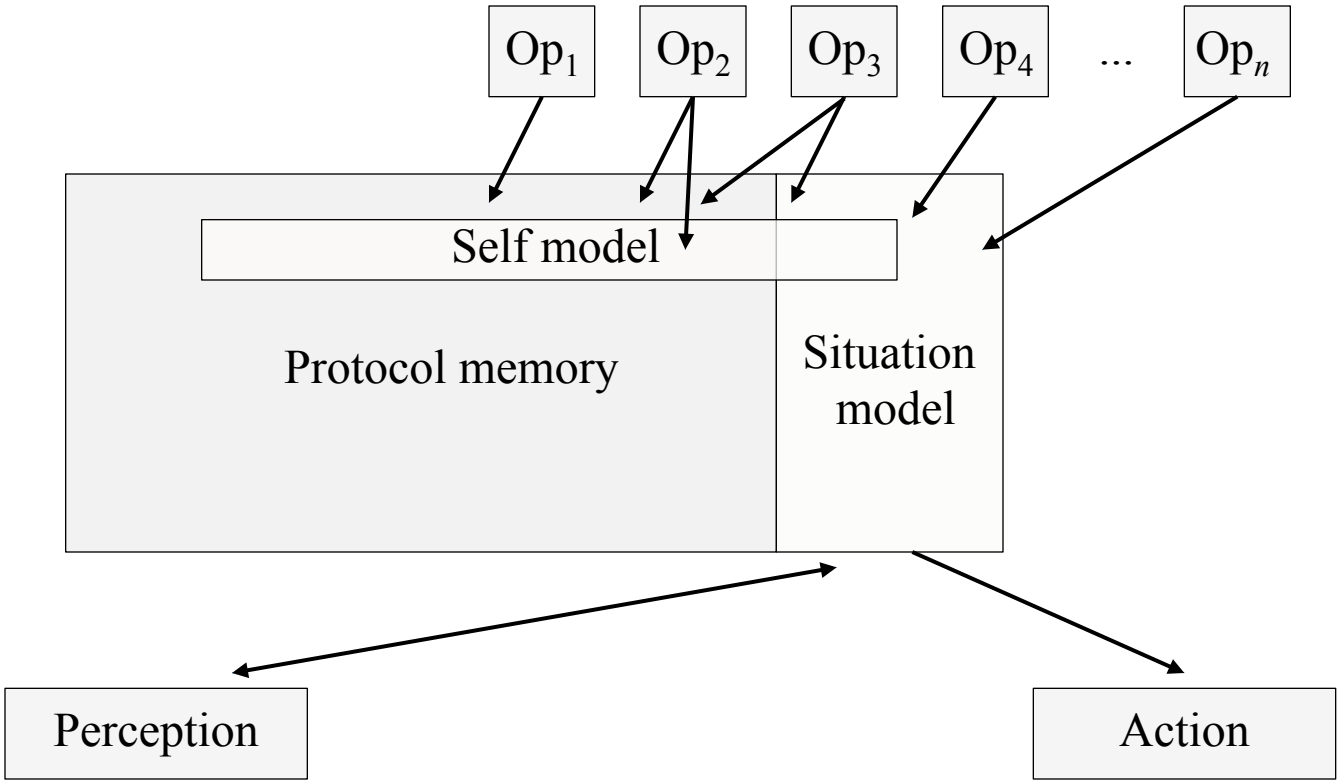
Components for Cognitive AI

- Model of current situation, and protocol of past situations



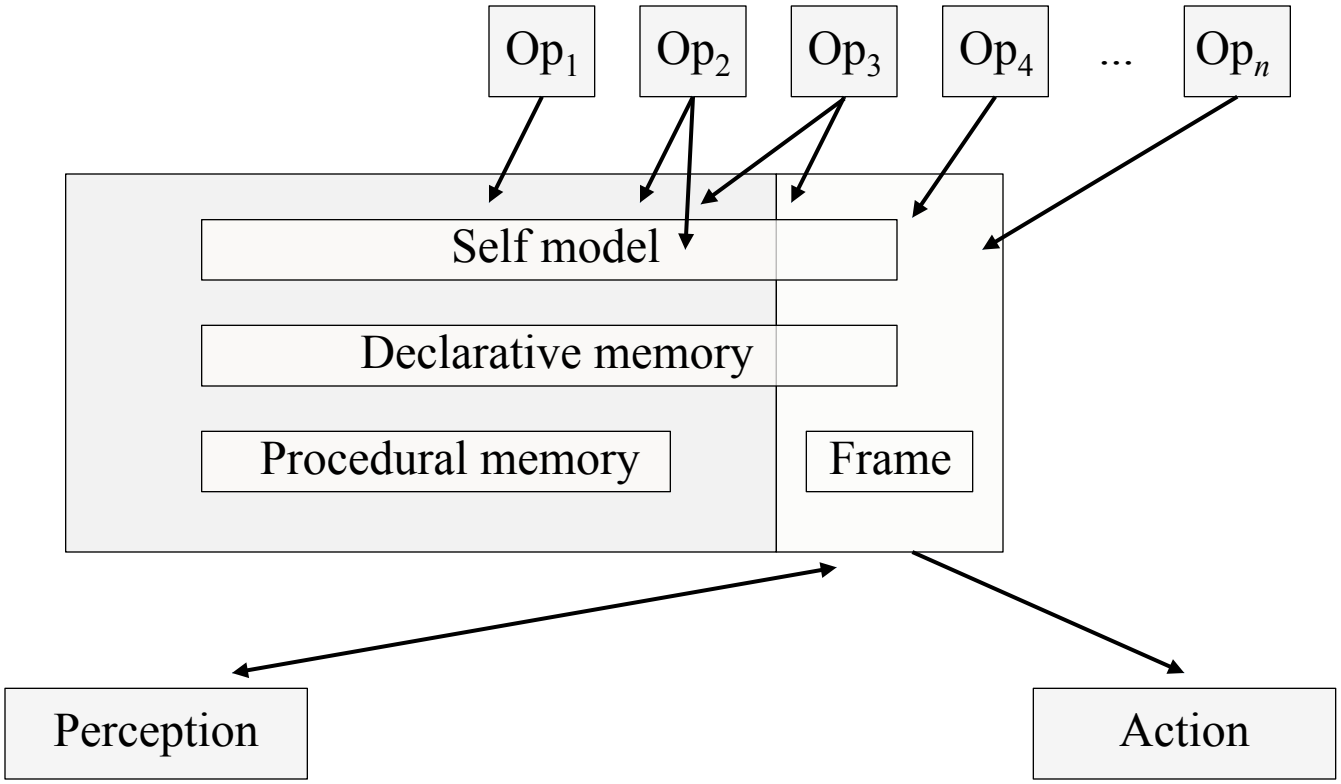
Components for Cognitive AI

- Model of self



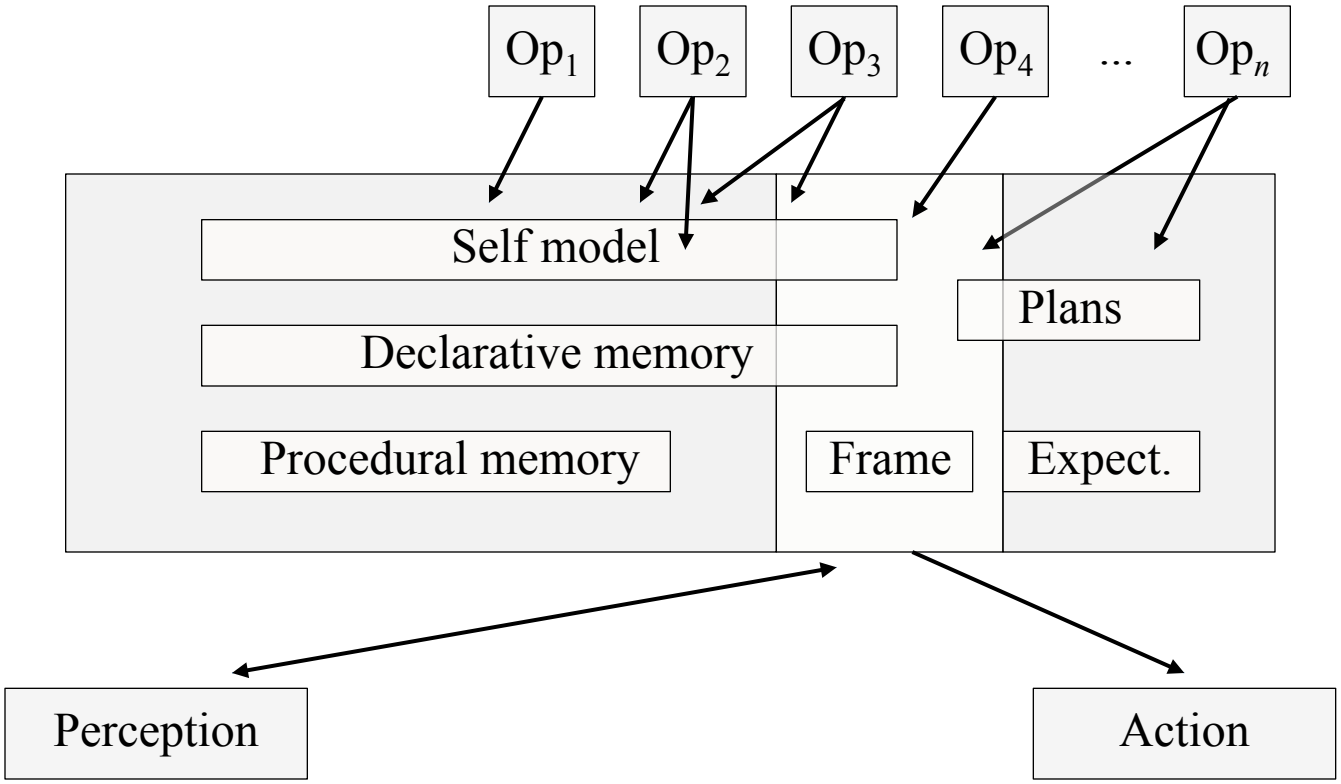
Components for Cognitive AI

- Abstractions of objects, episodes and types



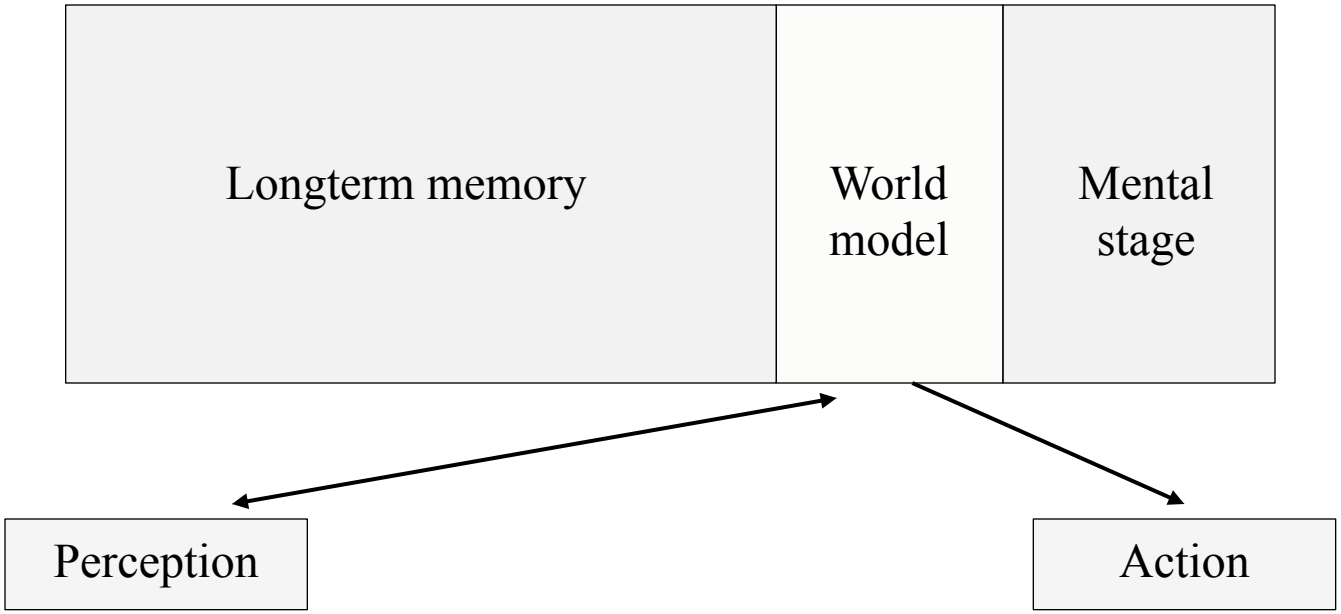
Components for Cognitive AI

- Anticipation of future developments



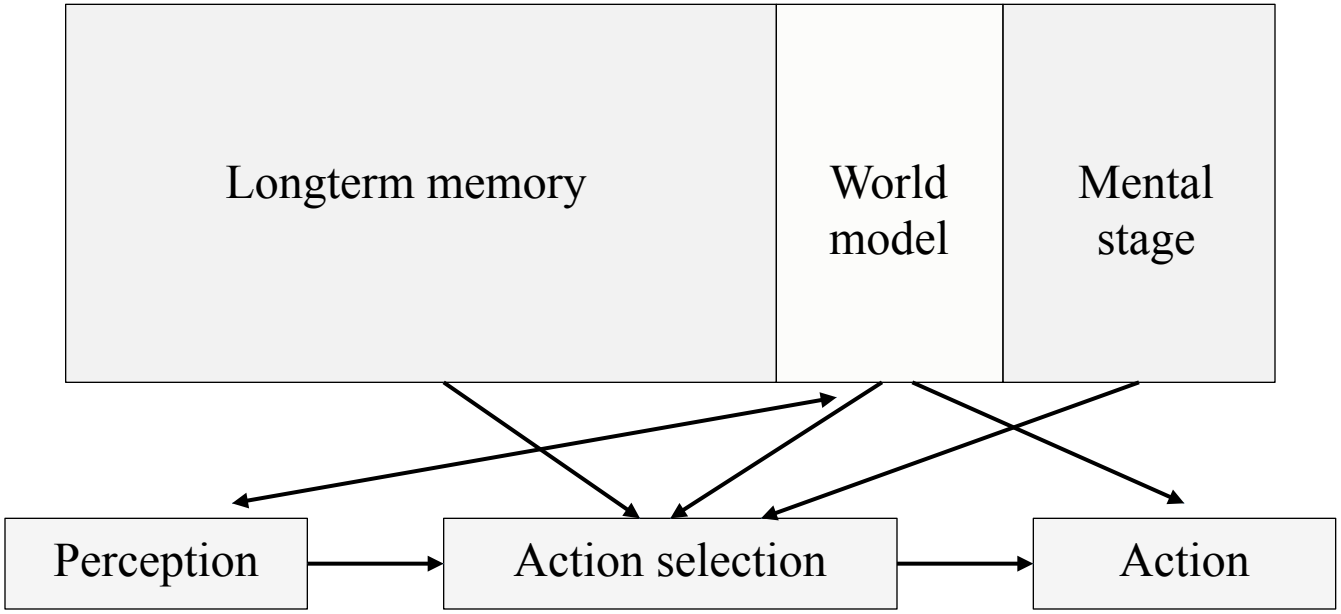
Components for Cognitive AI

- Action selection and executive control



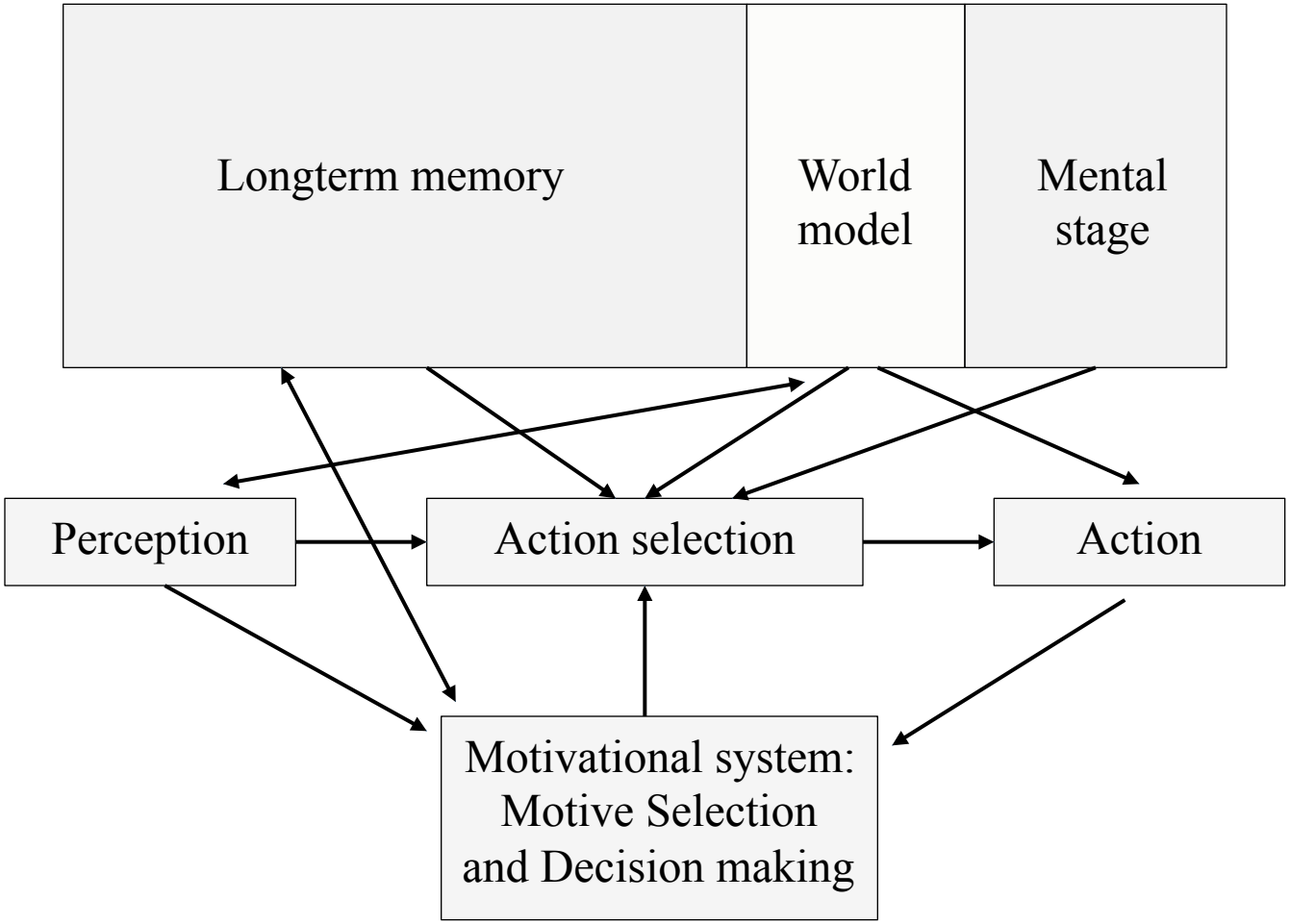
Components for Cognitive AI

- Action selection and executive control



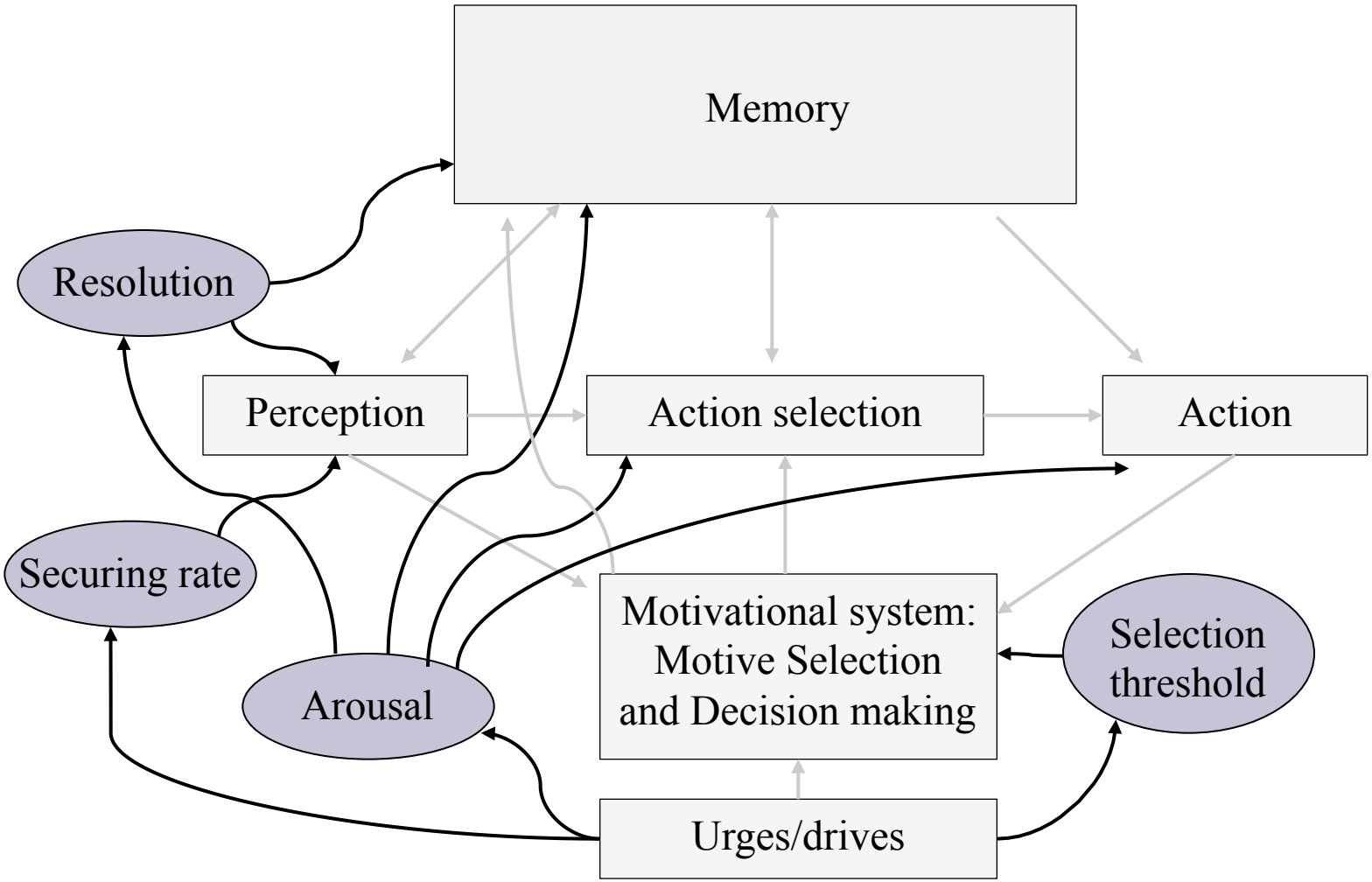
Components for Cognitive AI

- Universal motivation: autonomous identification of goals

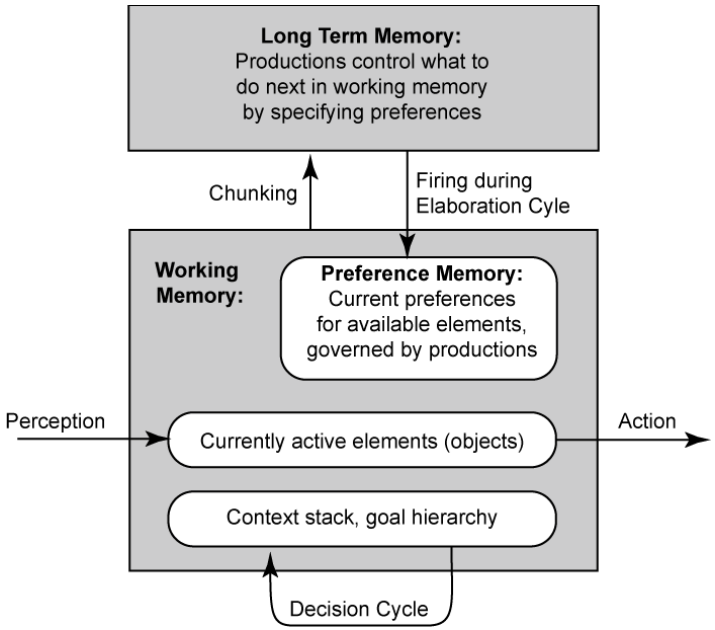


Components for Cognitive AI

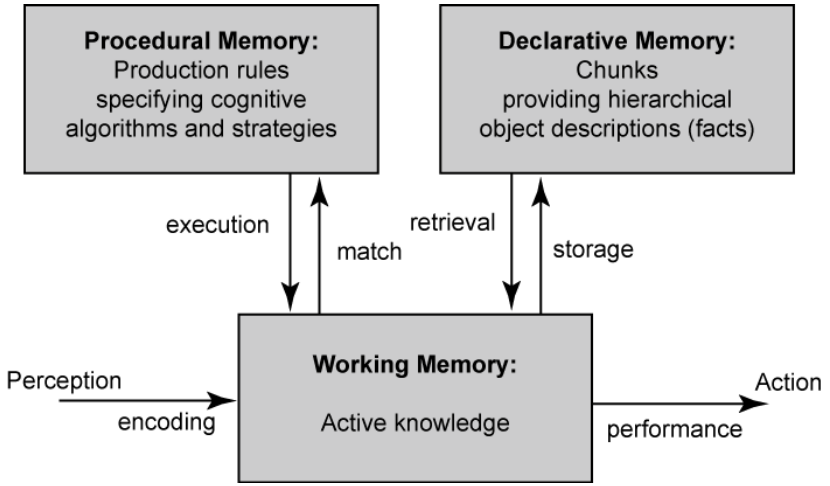
- Emotional modulation and affect



Cognitive Architectures



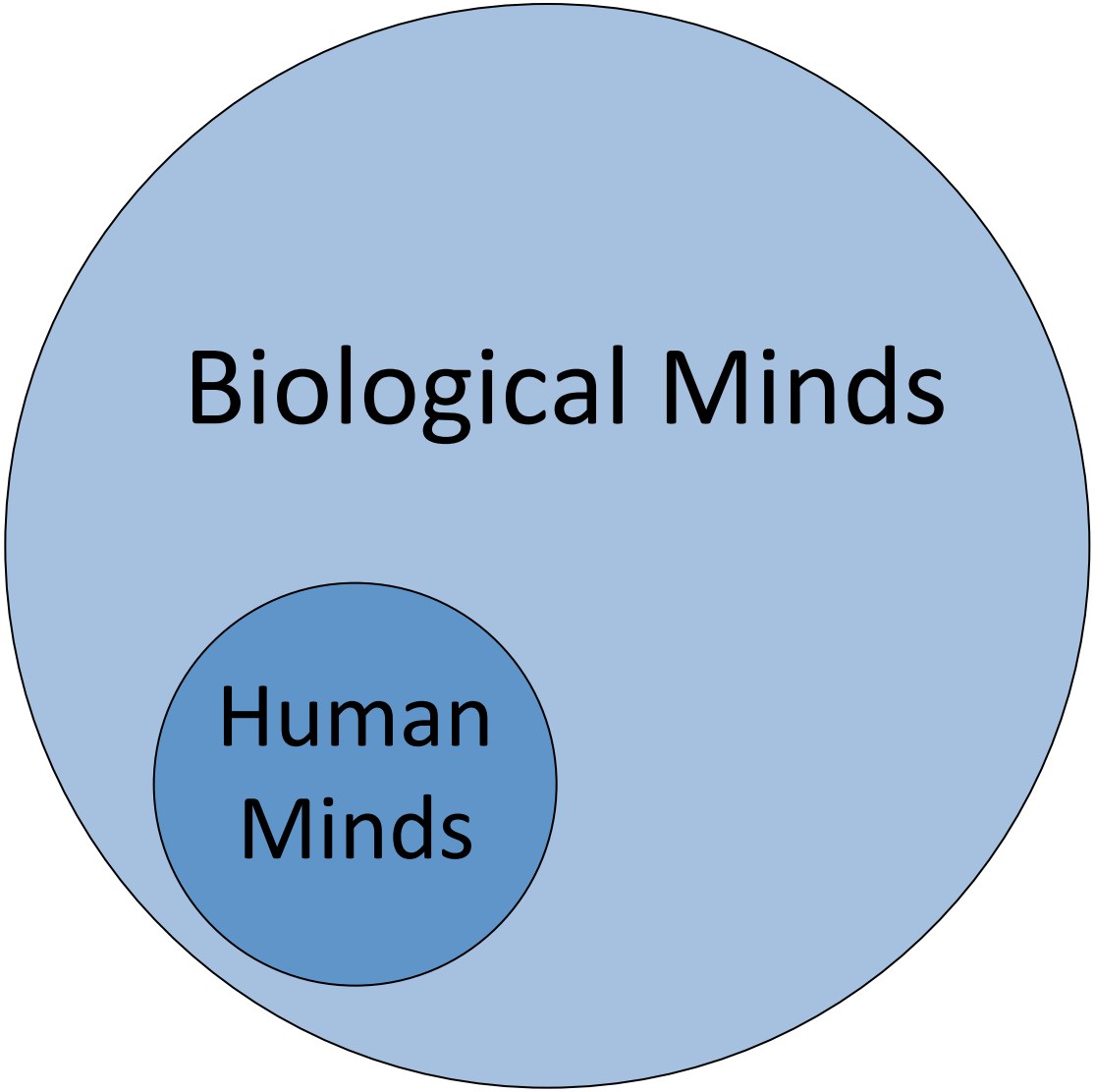
Soar



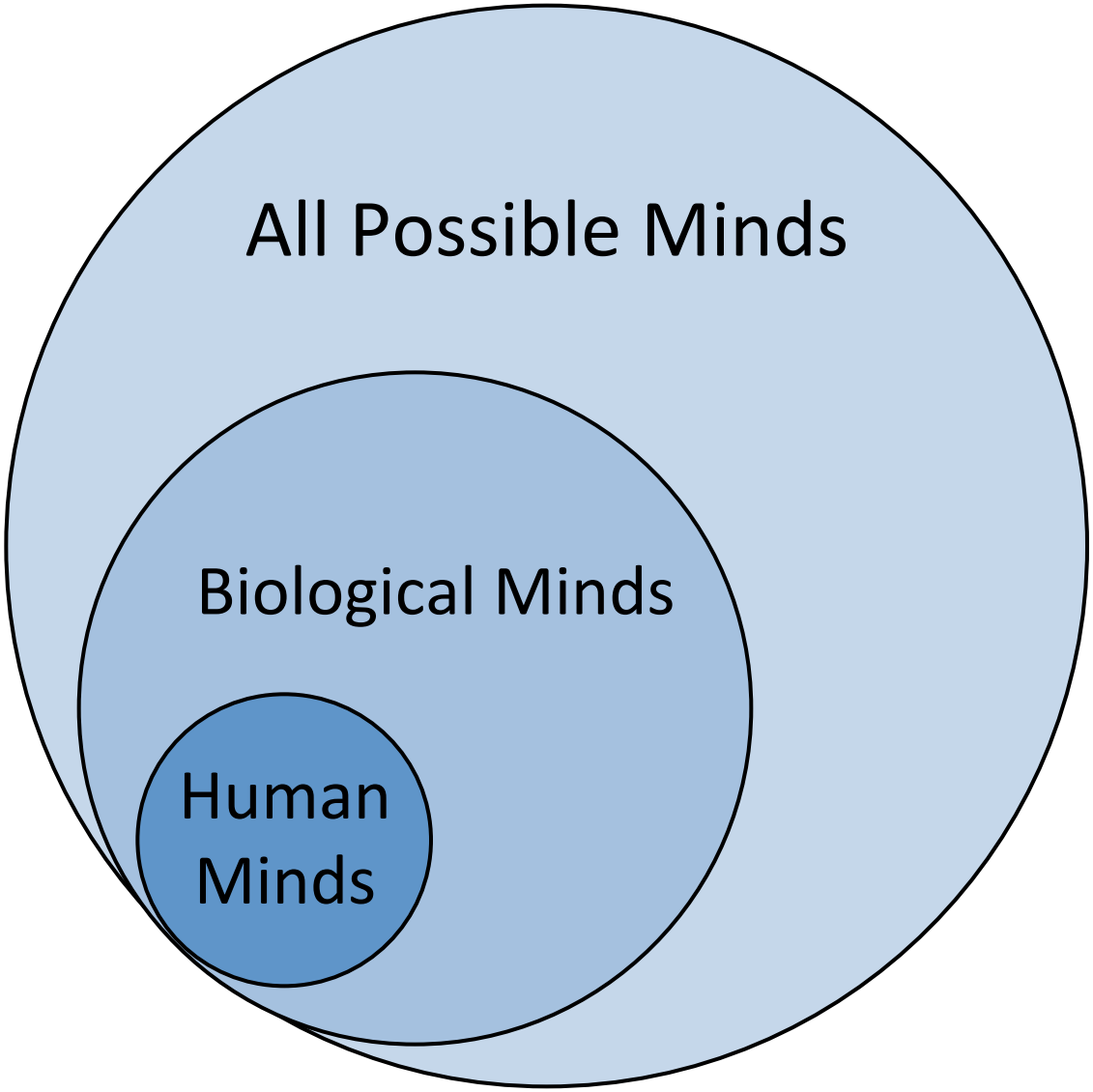
ACT-R

"Classical Cognitive Architectures" tend to focus on cognition as an isolated problem solving capability.

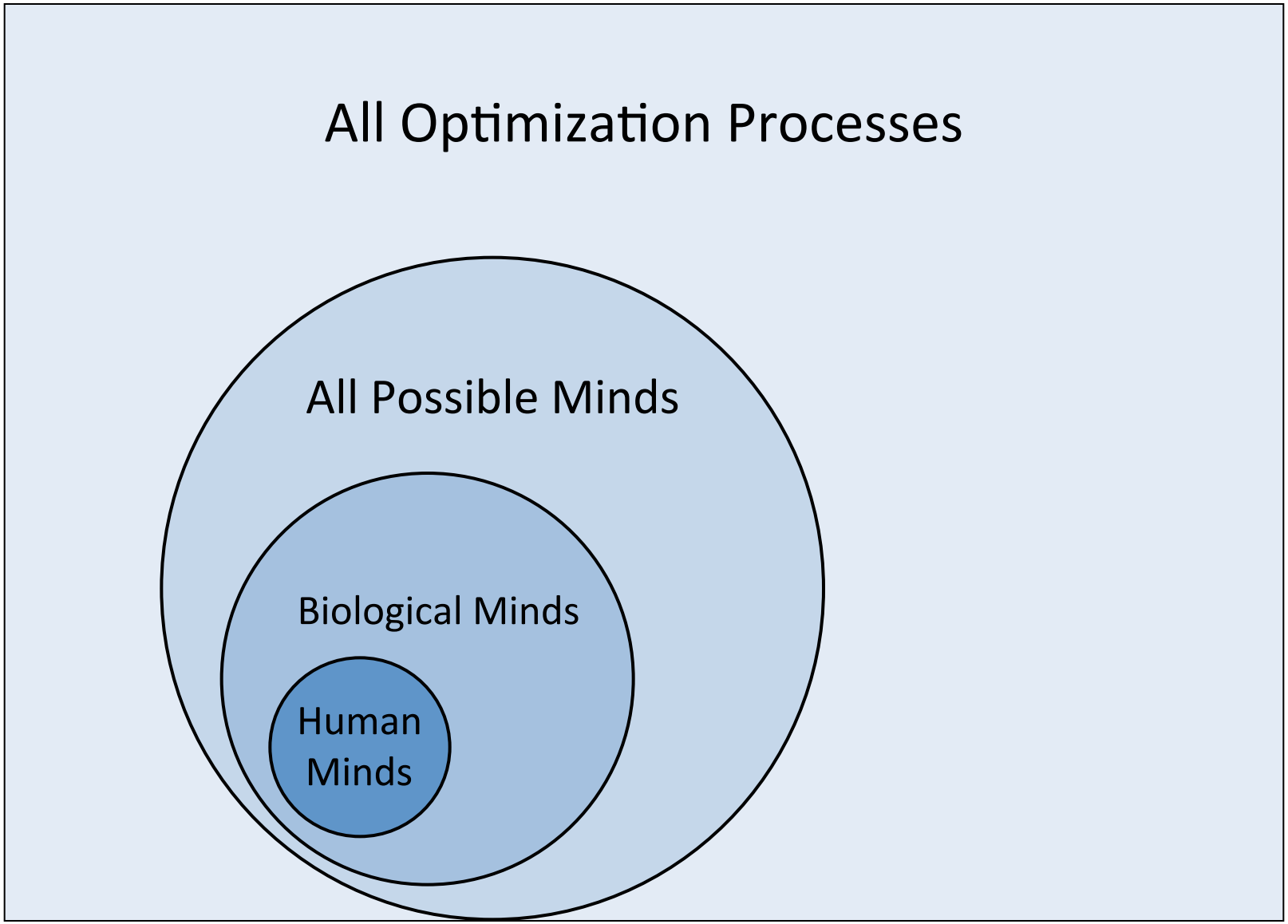
Possible Minds (Eliezer Yudkowsky)



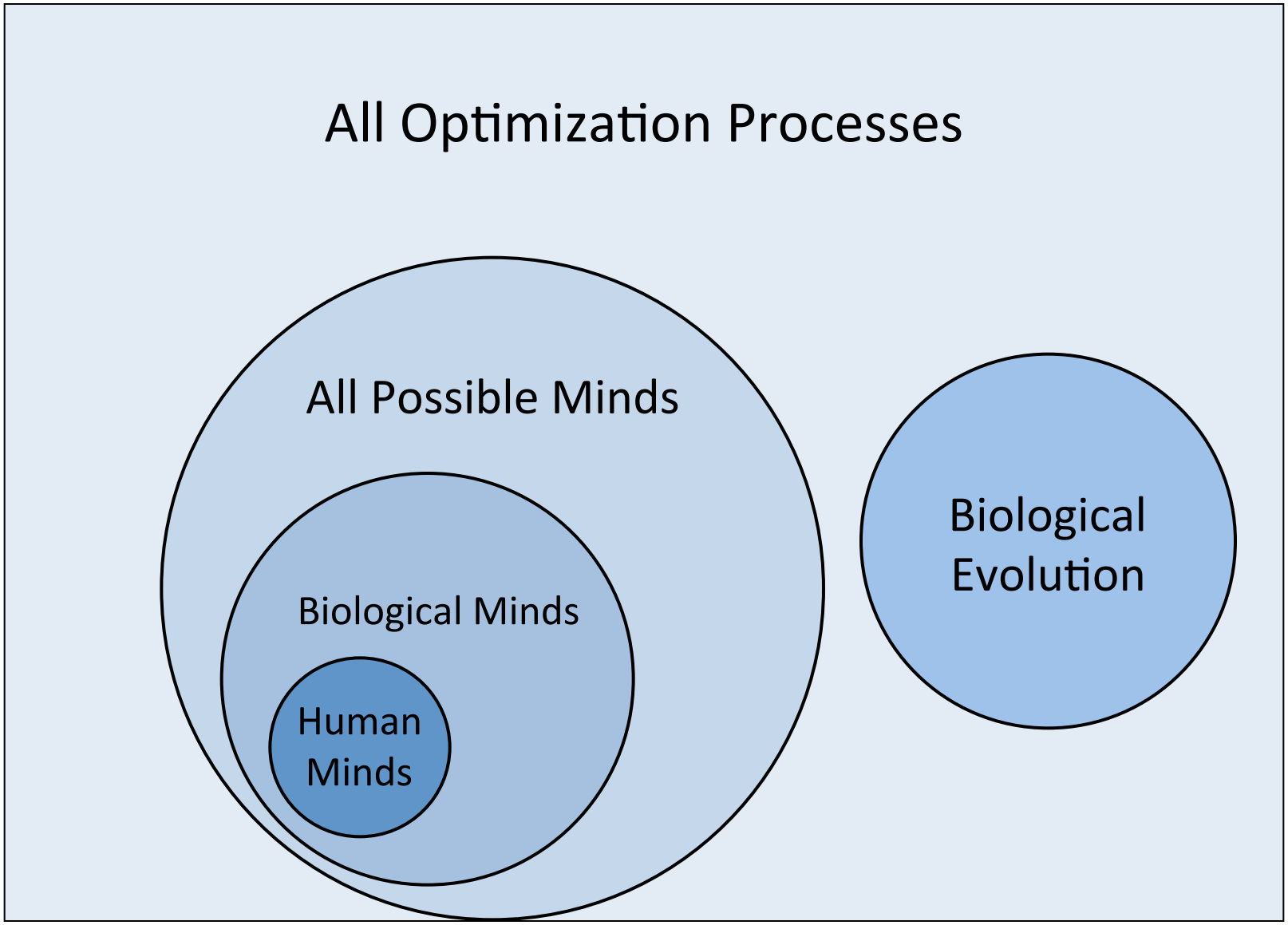
Possible Minds



Possible Minds



Possible Minds (not to scale)



Kinds of Minds (J. Storrs Hall)

- **Hypo-human:** infrahuman, less-than-human capacity.
- **Diahuman:** human-level capacities in some areas, but still not a general intelligence
- **Parahuman**: similar but not identical to humans, as for example, augmented humans.
- **Allohuman:** as capable as humans, but in different areas.
- **Epihuman:** slightly beyond the human level.
- **Hyperhuman:** much more powerful than human, superintelligent.

Space of Possible Minds (Aaron Sloman)

- Quantitative VS Structural;
- Continuous VS Discrete;
- Complexity of stored instructions;
- Serial VS Parallel;
- Distributed VS Fundamentally Parallel

Space of Possible Minds (Aaron Sloman)

- Connected to External Environment VS Not Connected;
- Moving VS Stationary;
- Capable of modeling others VS Not capable;
- Capable of logical inference VS Not Capable;
- Fixed VS Re-programmable;
- Goal consistency VS Goal Selection;
- Meta-Motives VS Motives;
- Able to delay goals VS Immediate goal following;
- Static Plans VS Dynamic Plans;
- Self-aware VS Not Self-Aware.

Generality of Human Minds

- ability to cope with *varied* objects in a domain
- ability to cope with *a variety* of domains of objects
- ability to perform *a variety* of tasks in relation to any object

- Self-improvement
- Graceful degradation

Specific Abilities of Human Minds

- Inference, reasoning under uncertainty, reasoning with nonlogical representations (maps, diagrams, models)
- hypothetical questions („What would happen if . . .?“) for plans, predictions, generalizations
- Insight and understanding rather than brute force
- Communication and co-operation with other intelligent agents
- Coping with inconsistent 'motivators', e.g. goals, tastes, preferences, ethical principles, constraints, etc.
- Self-reflection, ability to explain one's actions
- Generate, or appreciate, aesthetic objects
- Experience bodily sensations.
- Enjoy or dislike experiences, to be amused, angry, excited, irritated, hopeful, disgusted, etc.

Sessions: see futureai.media.mit.edu

9/21: Possibilities for artificial minds

9/28: Agents within agents. The Society of Mind (Minsky)

10/05: ?Vision as inverse rendering? (Poggio?)

10/12: Columbus Day (no session)

10/19: ?AI as engineering or AI as a science. The Norvig/ Chomsky debate

10/26: The Neocognitron and Deep Learning

11/2: Universal intelligence. From Solomonoff induction to AIXI

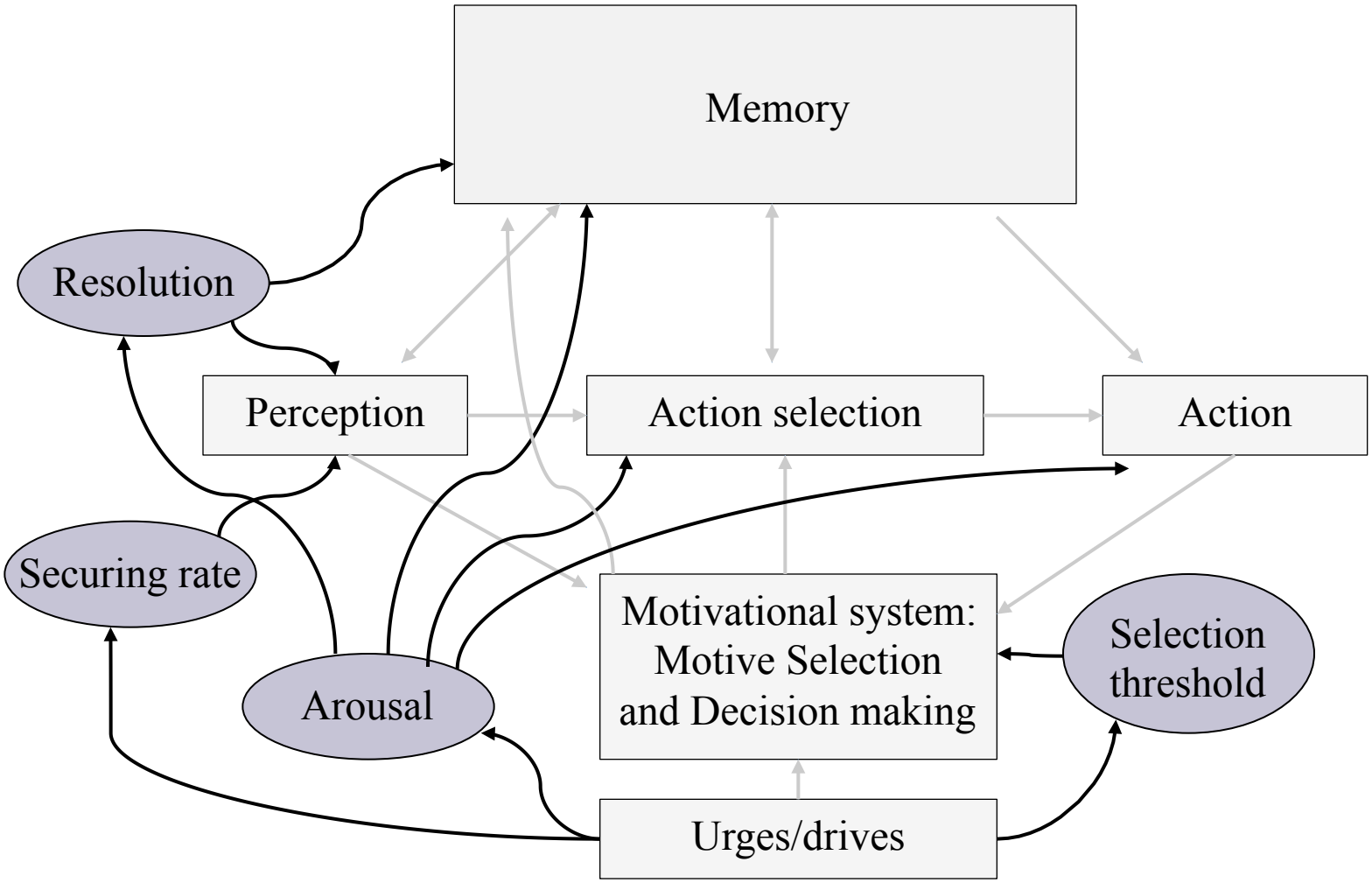
11/9: AI and Neuroscience

11/23: Affect and Motivation

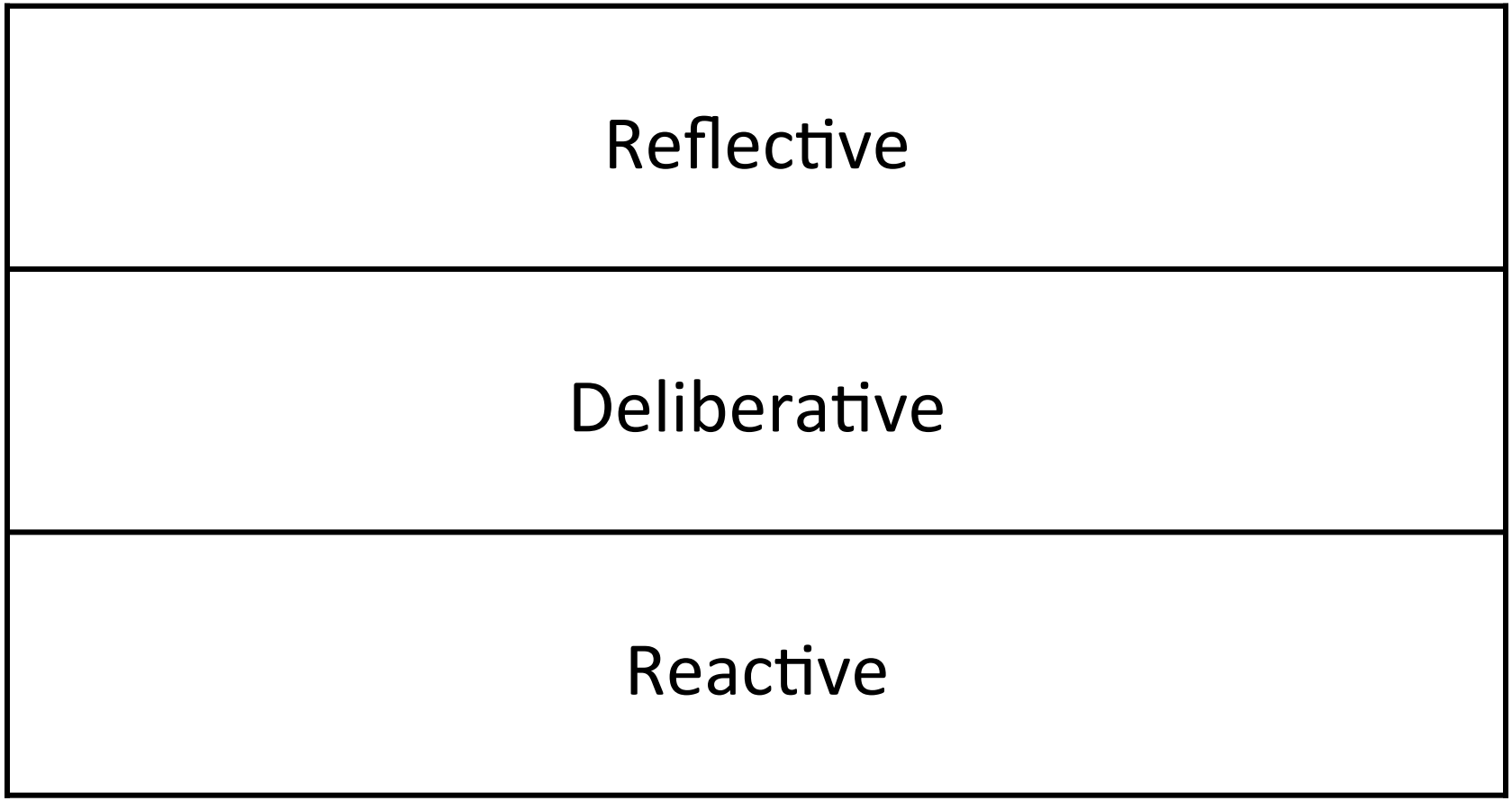
11/30: Measuring the Progress of AI. Benchmarks

12/7: Closing Discussion. Can we sketch a Map of Future AI research?

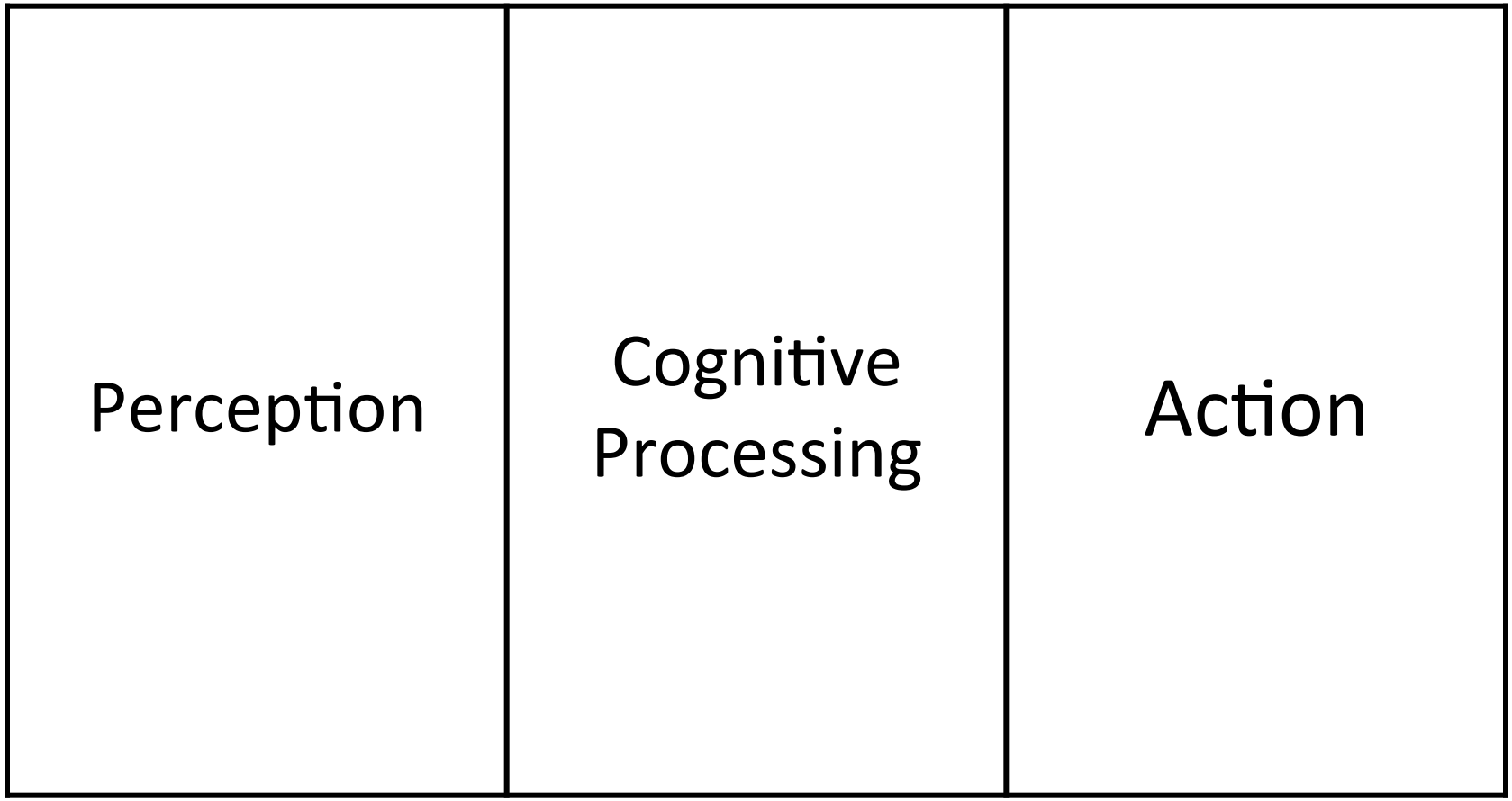
Components for Cognitive AI



Layers of Cognition



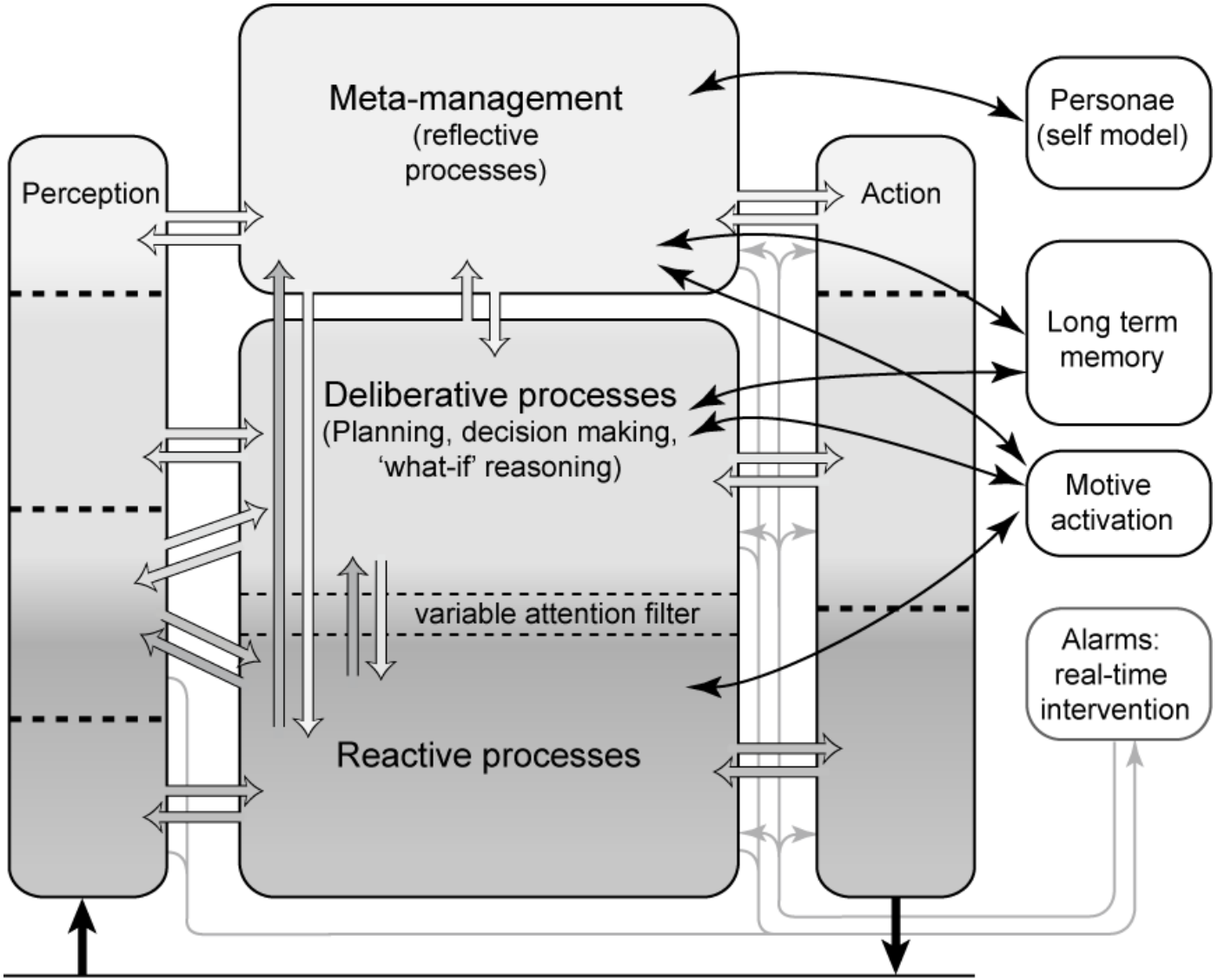
Columns of Cognition



Cognitive Grid

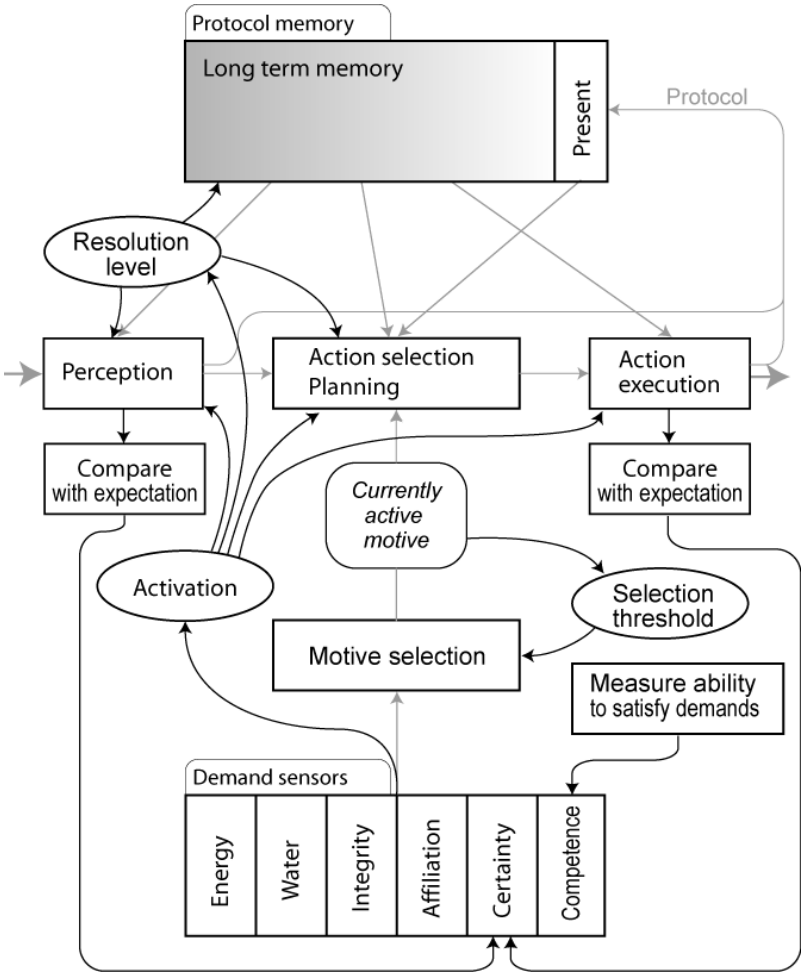
Reflexive Perception	Meta-Management	Management Action
Deliberative Perception	Planning, Reasoning	Deliberative Action
Reactive Perception	Reflexes	Reflexive Action

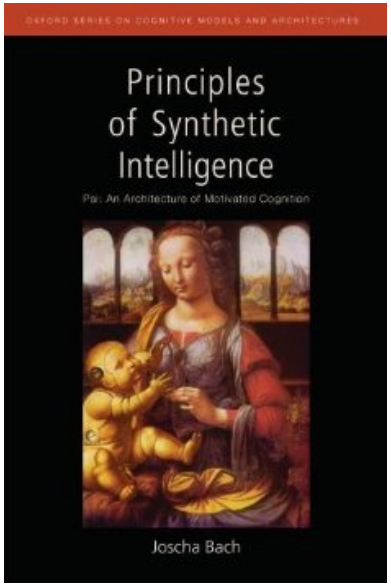
Conceptual Analysis: HCogAff (Sloman 2001)



Components for Cognitive AI

- Whole, testable architectures





PSI theory

Principles of Synthetic Intelligence

(Dörner 1999; Bach 2003, 2009)