ICRA 2024 Workshop

Autonomy in Robotic Surgery:
State of the Art, Technical and Regulatory Challenges for Clinical Application

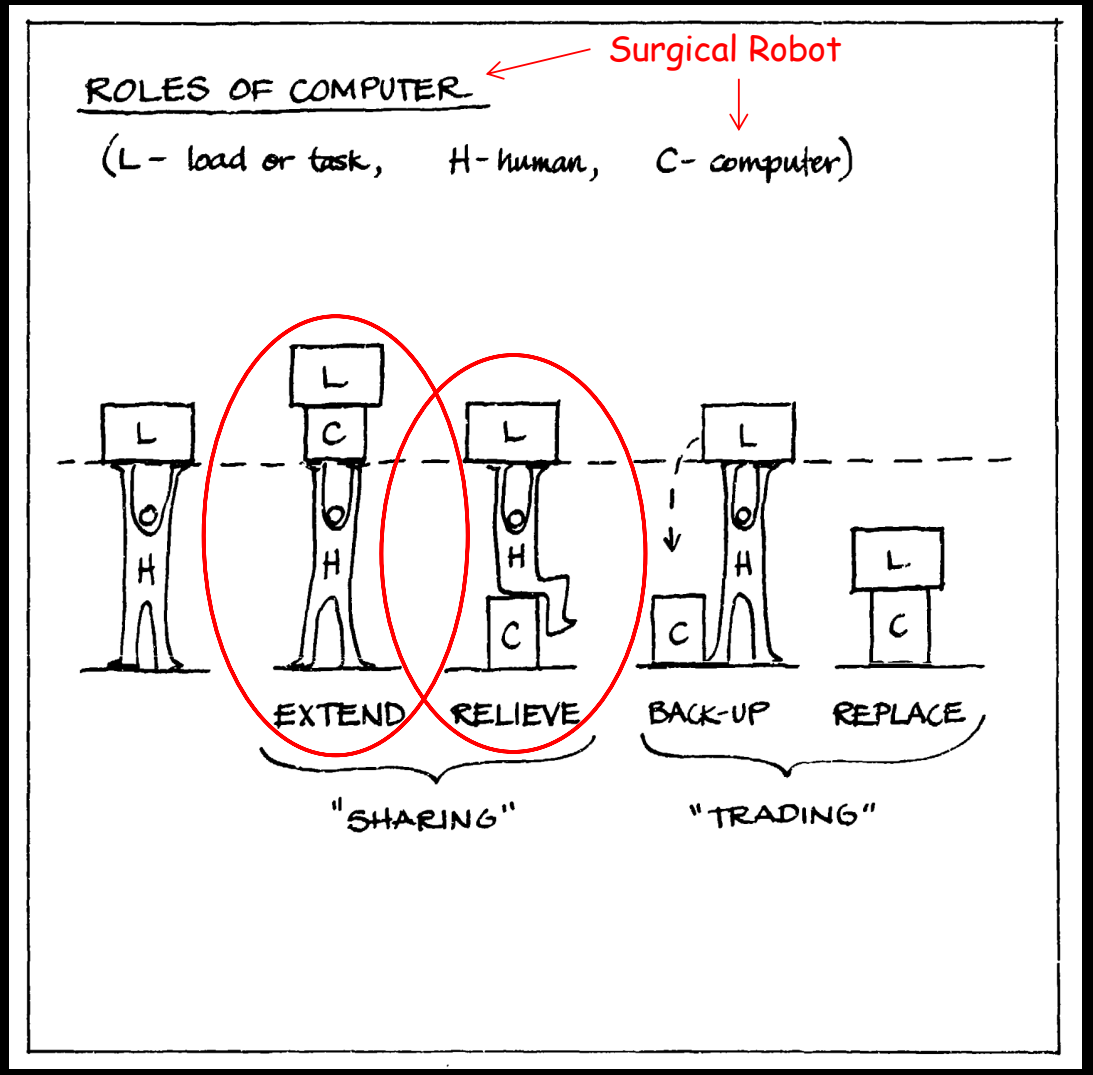
13th May 2024

Can Cognitive Architectures Support Autonomy in Robotic Surgery?

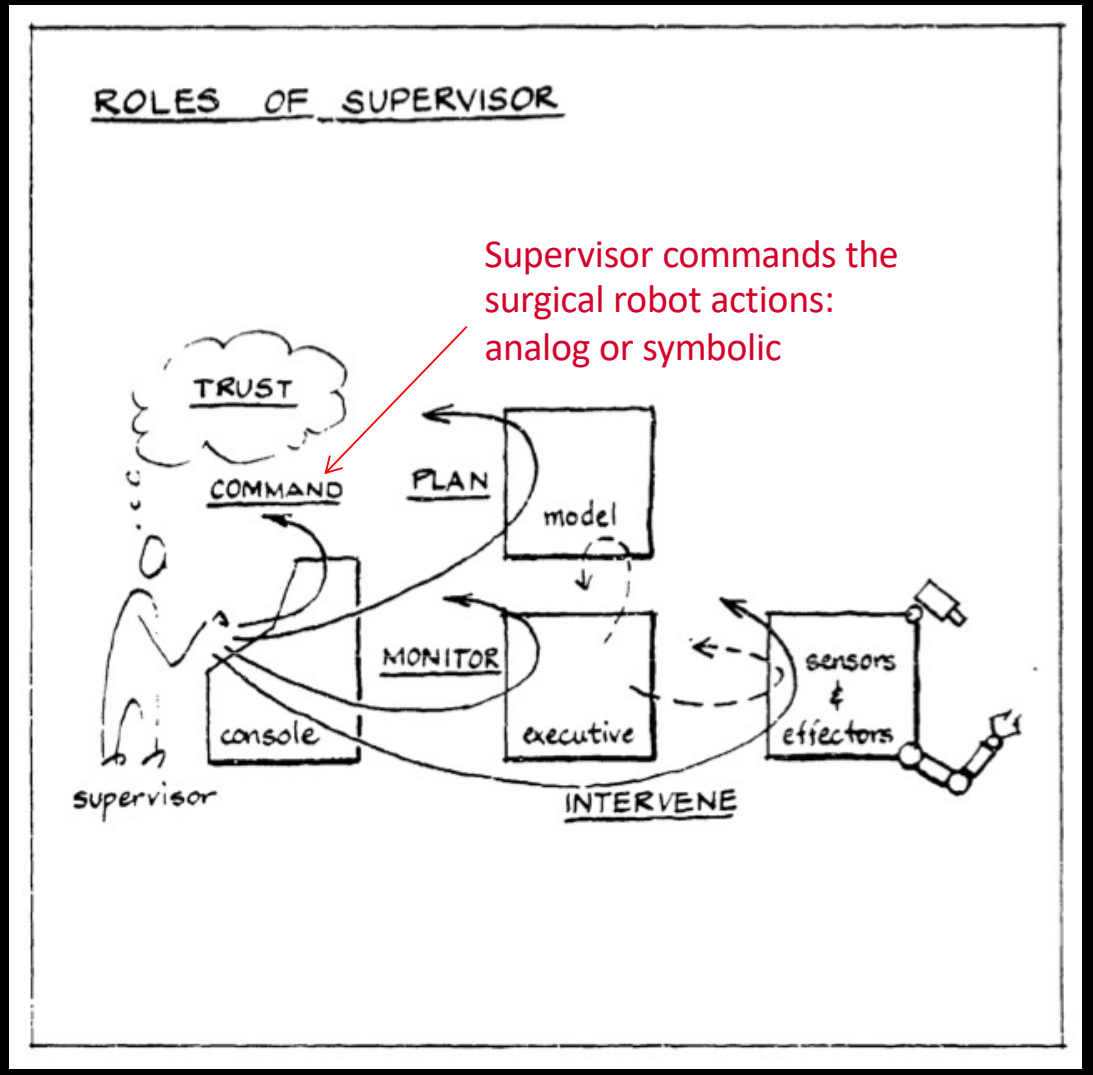
David Vernon

Carnegie Mellon University Africa

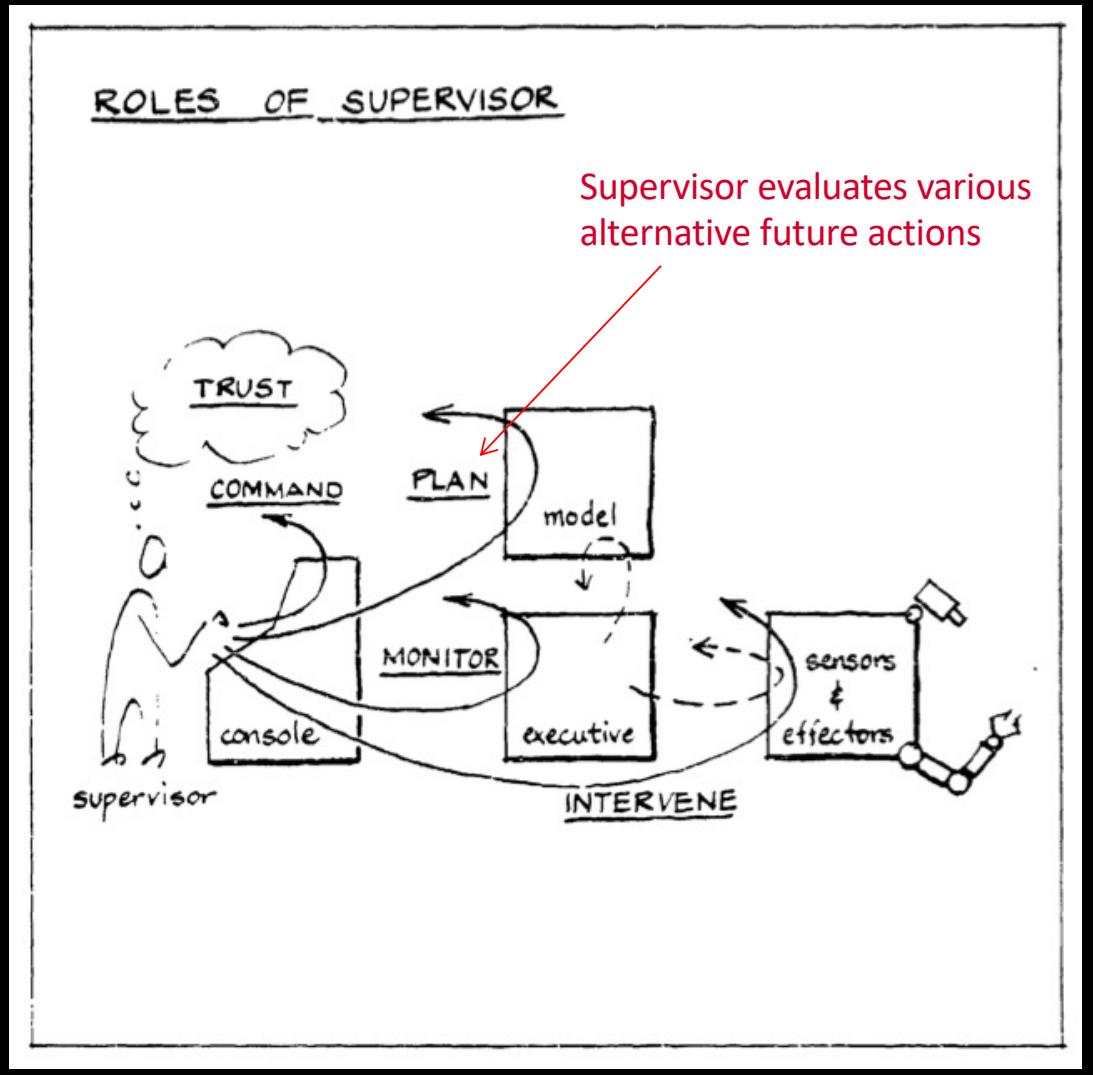
www.vernon.eu



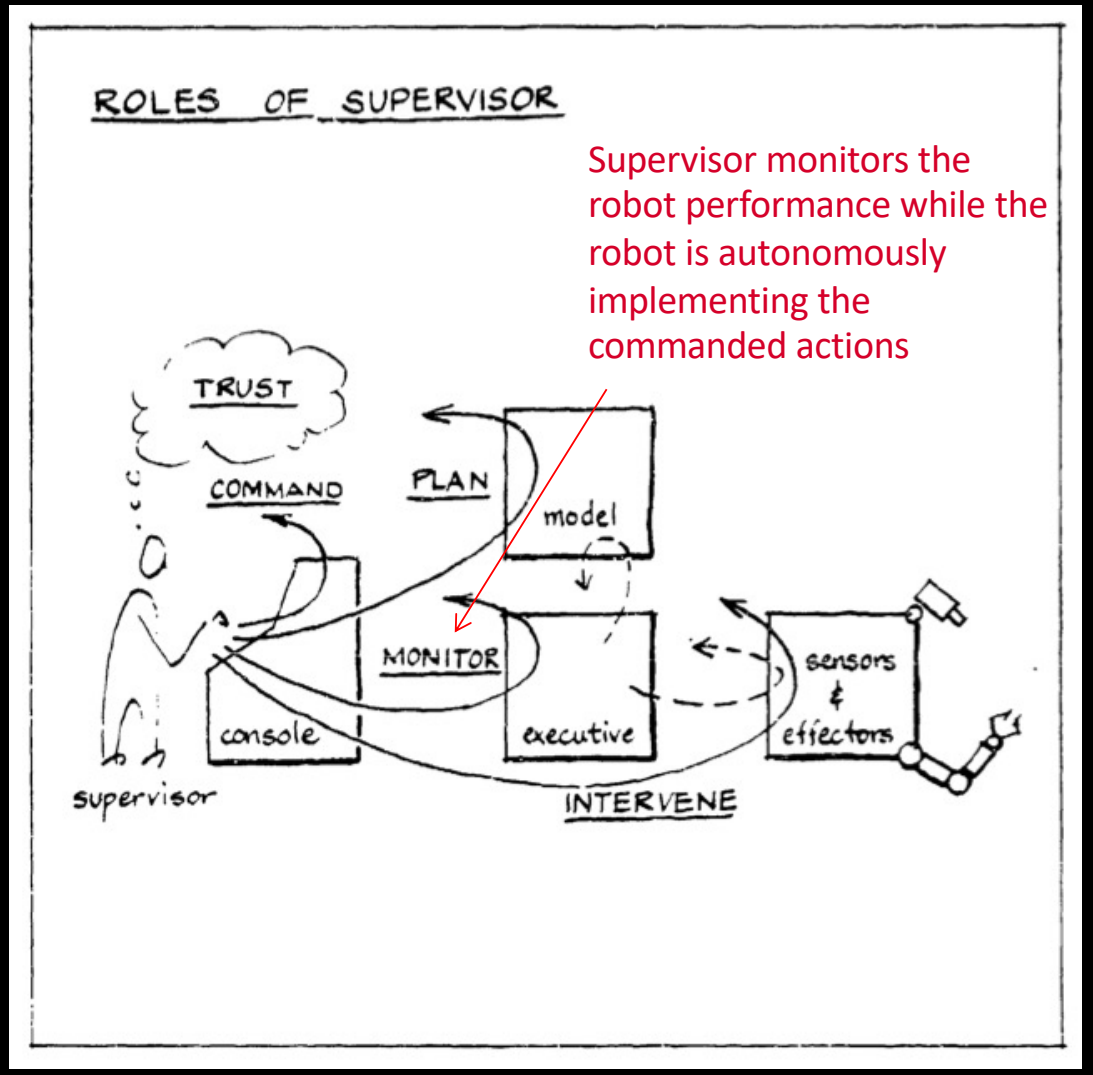
T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.



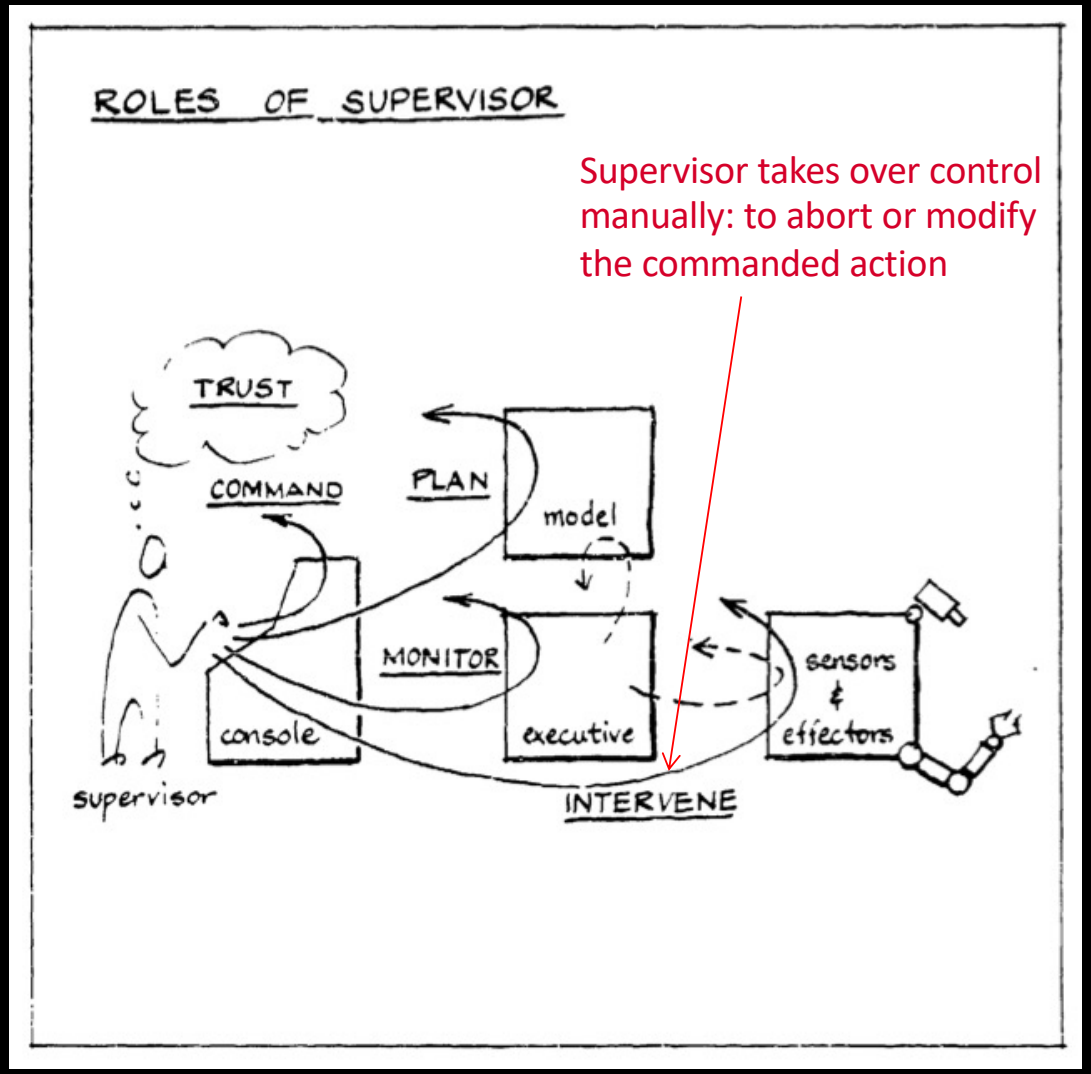
T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.



T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.



T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.

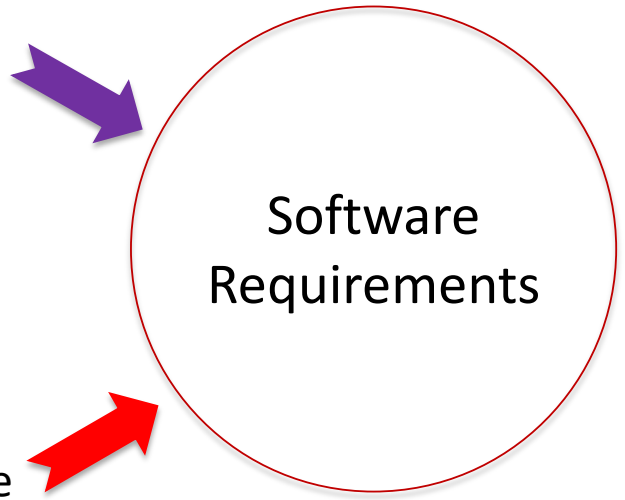


T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.

Autonomy in Robotic Surgery

A human supervisor must be able to ...

Understand the actions
and the basis for these actions
a surgical robot performs
autonomously

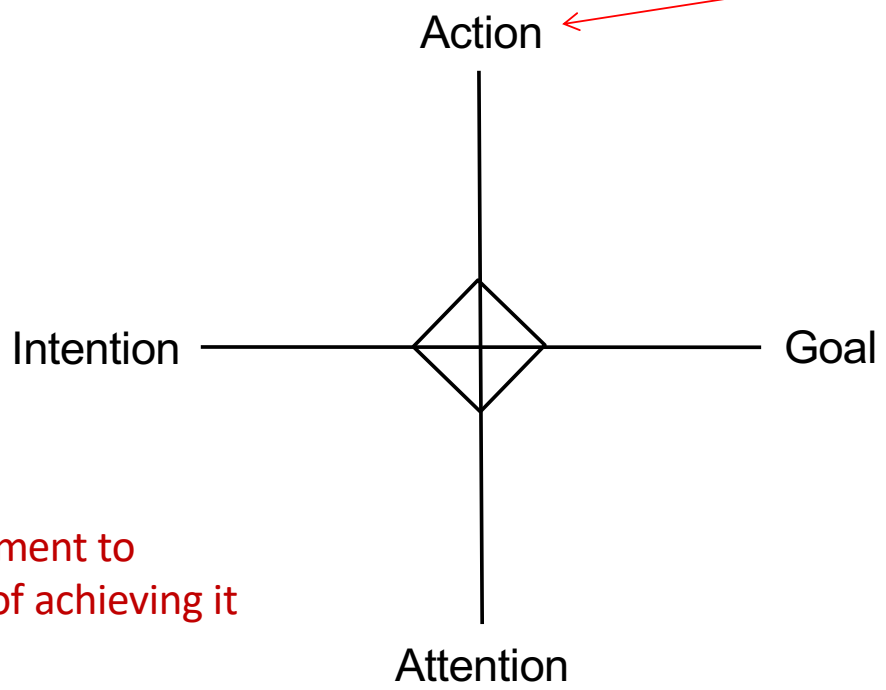


Surgical robot's system architecture
facilitates timely interaction with
the human supervisor

Let's unpack this

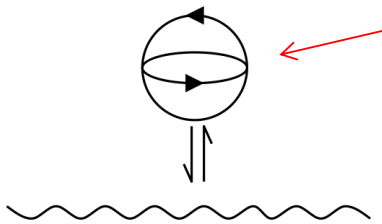
Intervene to stop or modify the
robot's intended action

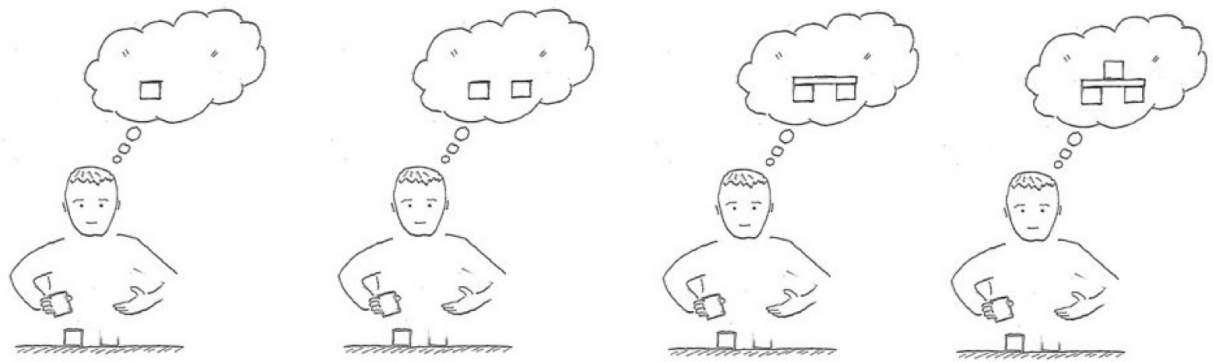
Guided by prospection



An intention implies a commitment to achieving a goal and a means of achieving it

An individual autonomous agent interacting with its environment





The Future

Cognition

"Cognition is the process by which an autonomous system

perceives its environment,

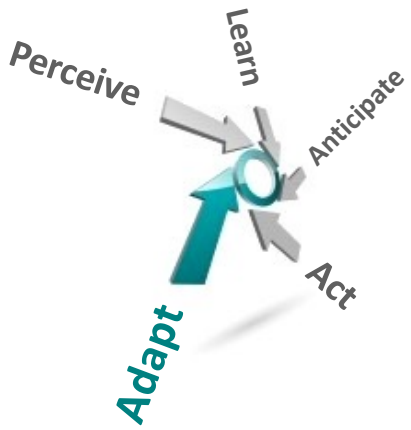
learns from experience,

anticipates the outcome of events,

acts to pursue goals, and

adapts to changing circumstances."

D. Vernon, Artificial Cognitive Systems – A Primer, MIT Press, 2014

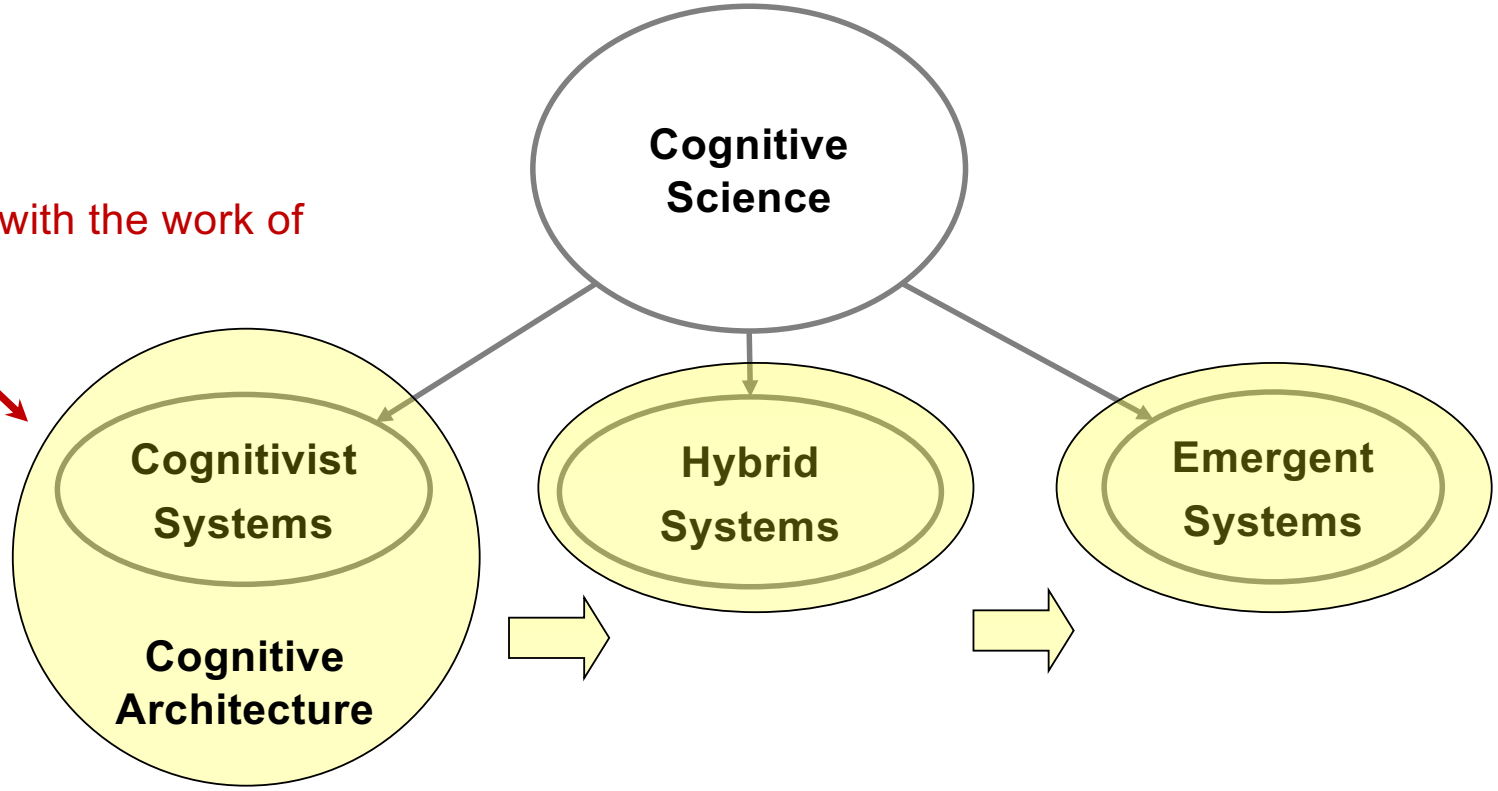


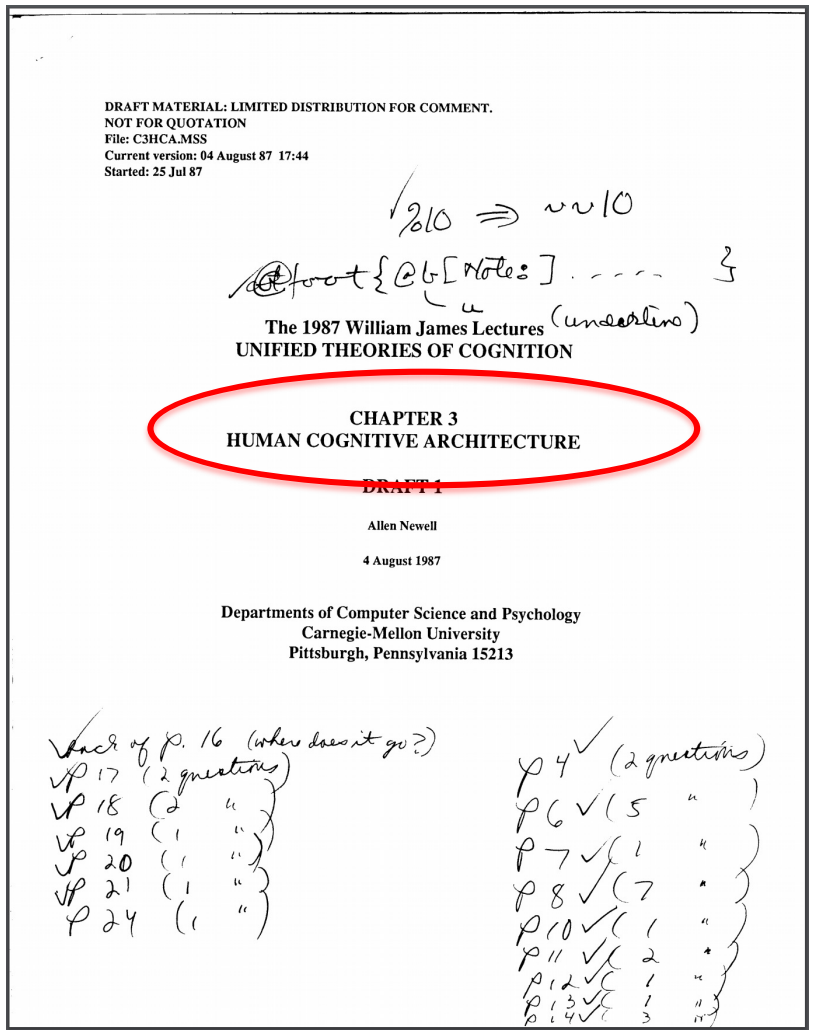
Doing this requires a **cognitive architecture** to orchestrate the core cognitive abilities:

Perception
Attention
Action selection
Memory
Learning
Reasoning
Meta-reasoning
Prospection

Cognitive Architectures

The term originated with the work of Allen Newell (1990)





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 (underline)
UNIFIED THEORIES OF COGNITION

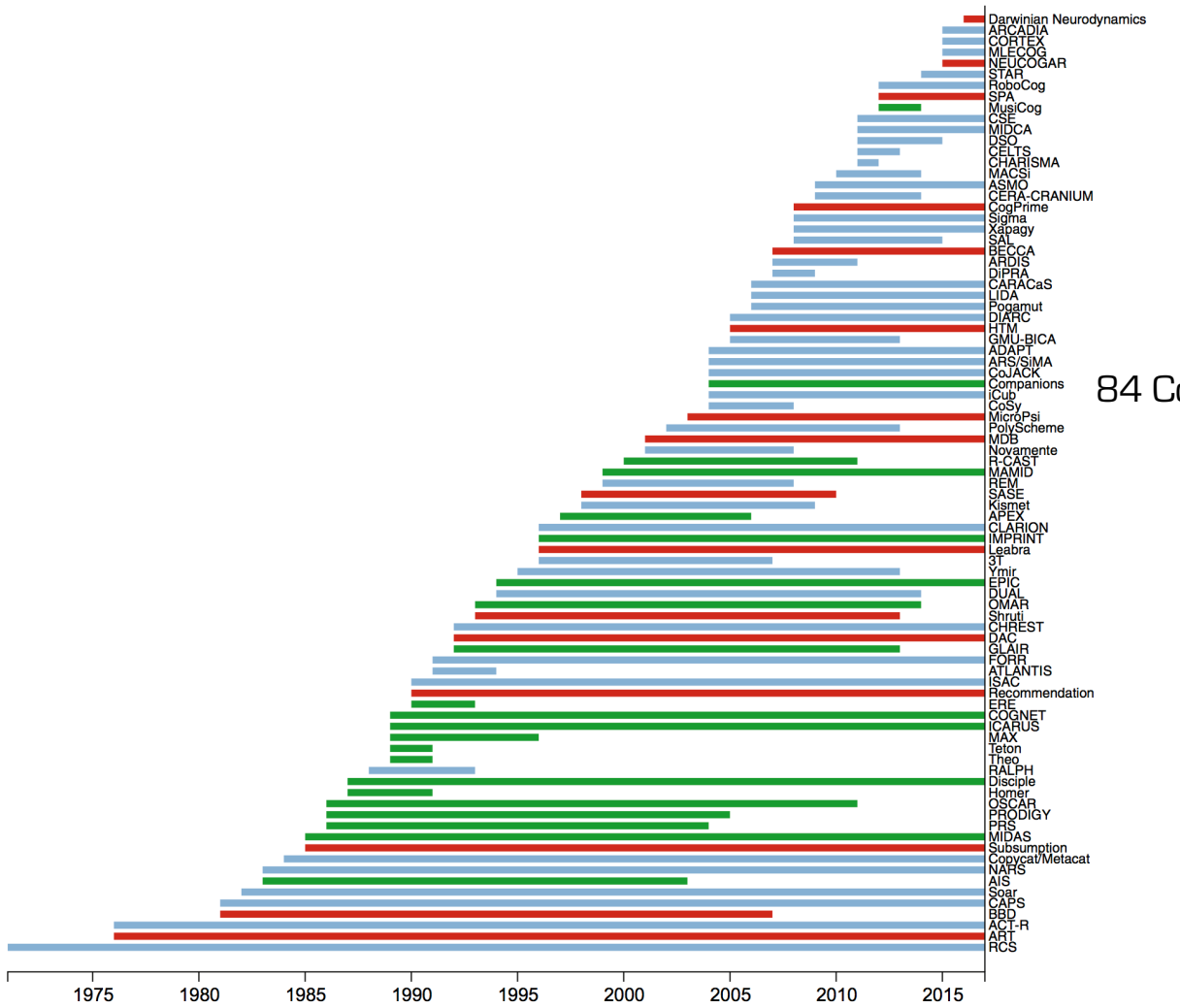
CHAPTER 3
HUMAN COGNITIVE ARCHITECTURE

DRAFT 1

Allen Newell

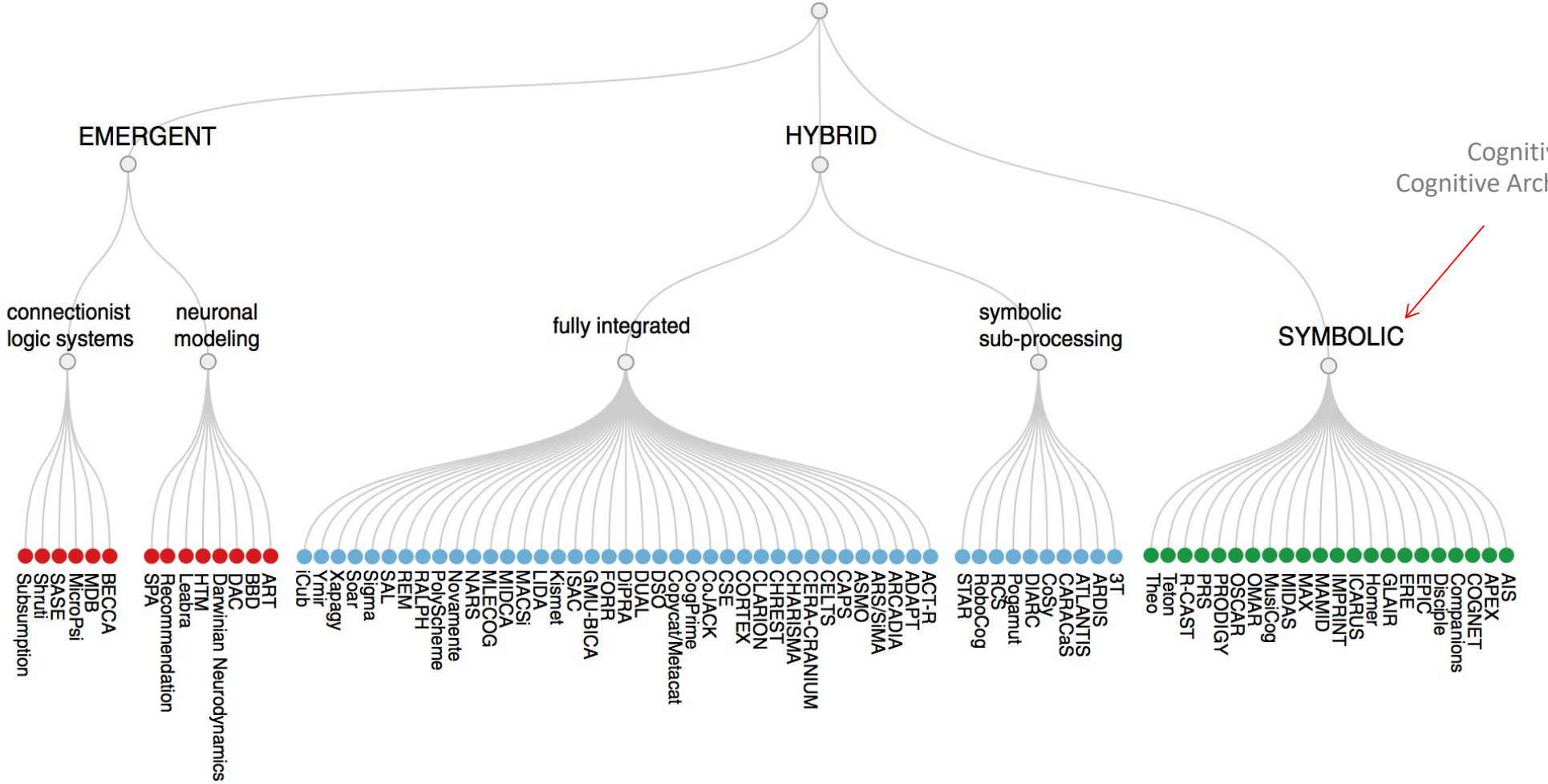
4 August 1987

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



84 Cognitive Architectures

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.

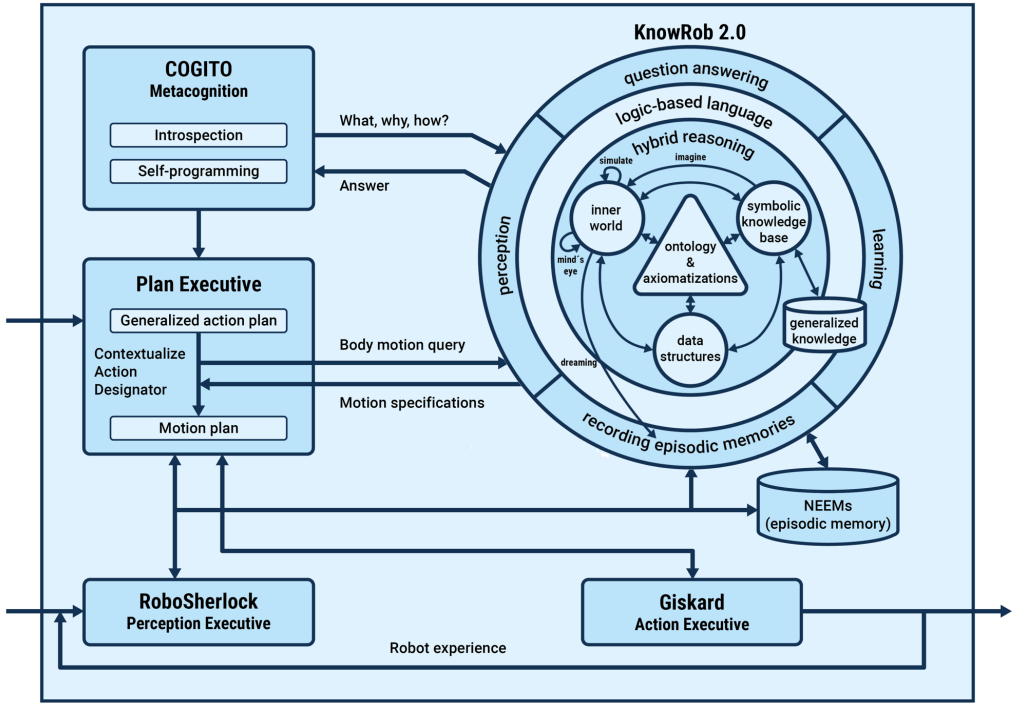


I. Kotseruba and J. Tsotsos. 40 years of cognitive architectures: core cognitive abilities and practical applications. Artificial Intelligence Review, Vol. 53, No. 1, pp. 17-94, 2020.

The CRAM Robot Cognitive Architecture

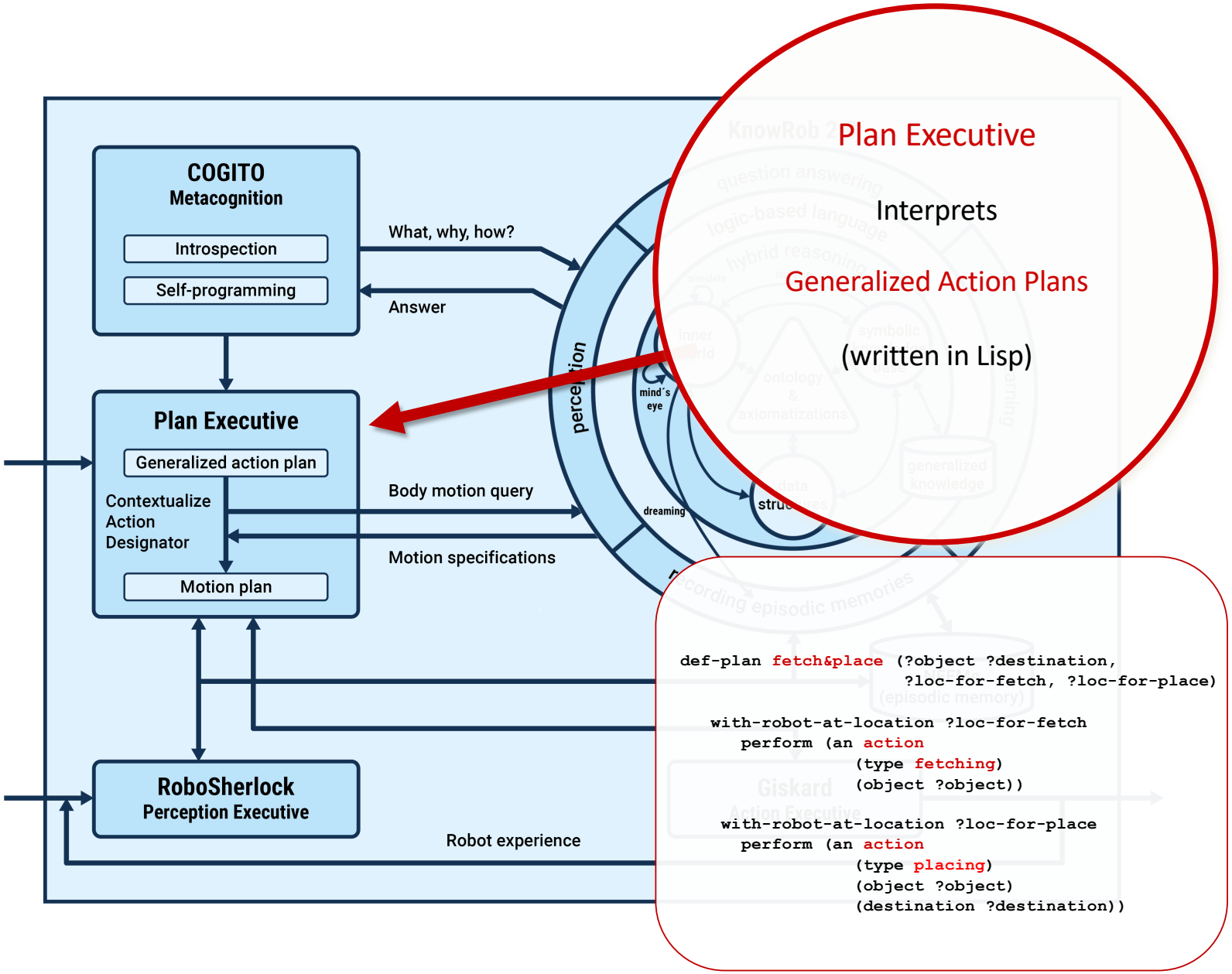
CRAM has five core elements:

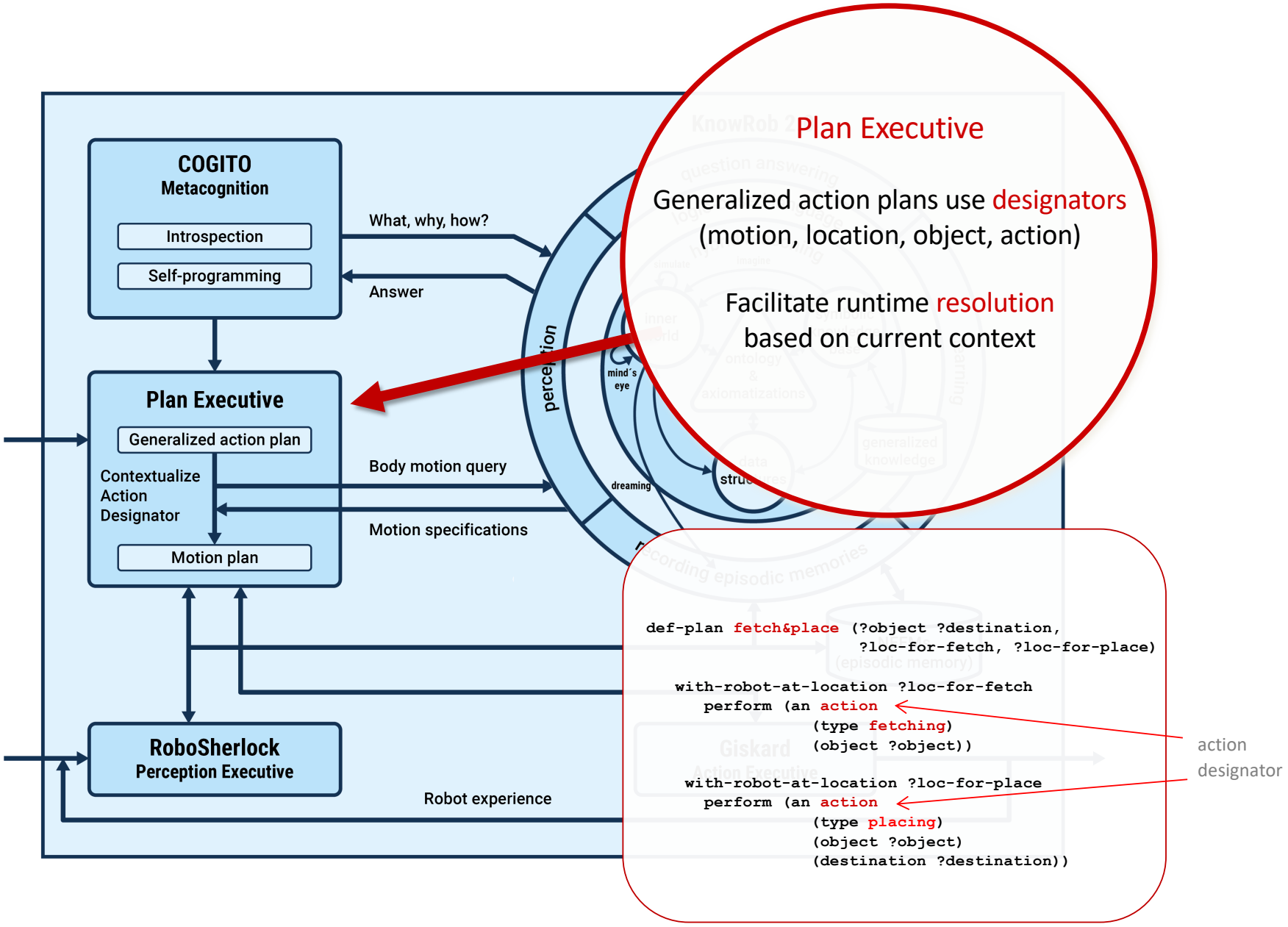
1. CRAM Plan Language (CPL) executive
2. KnowRob2.0 knowledge-bases and associated reasoning processes
3. RoboSherlock, the perception executive
4. Giskard, the action executive
5. COGITO, a metacognition system

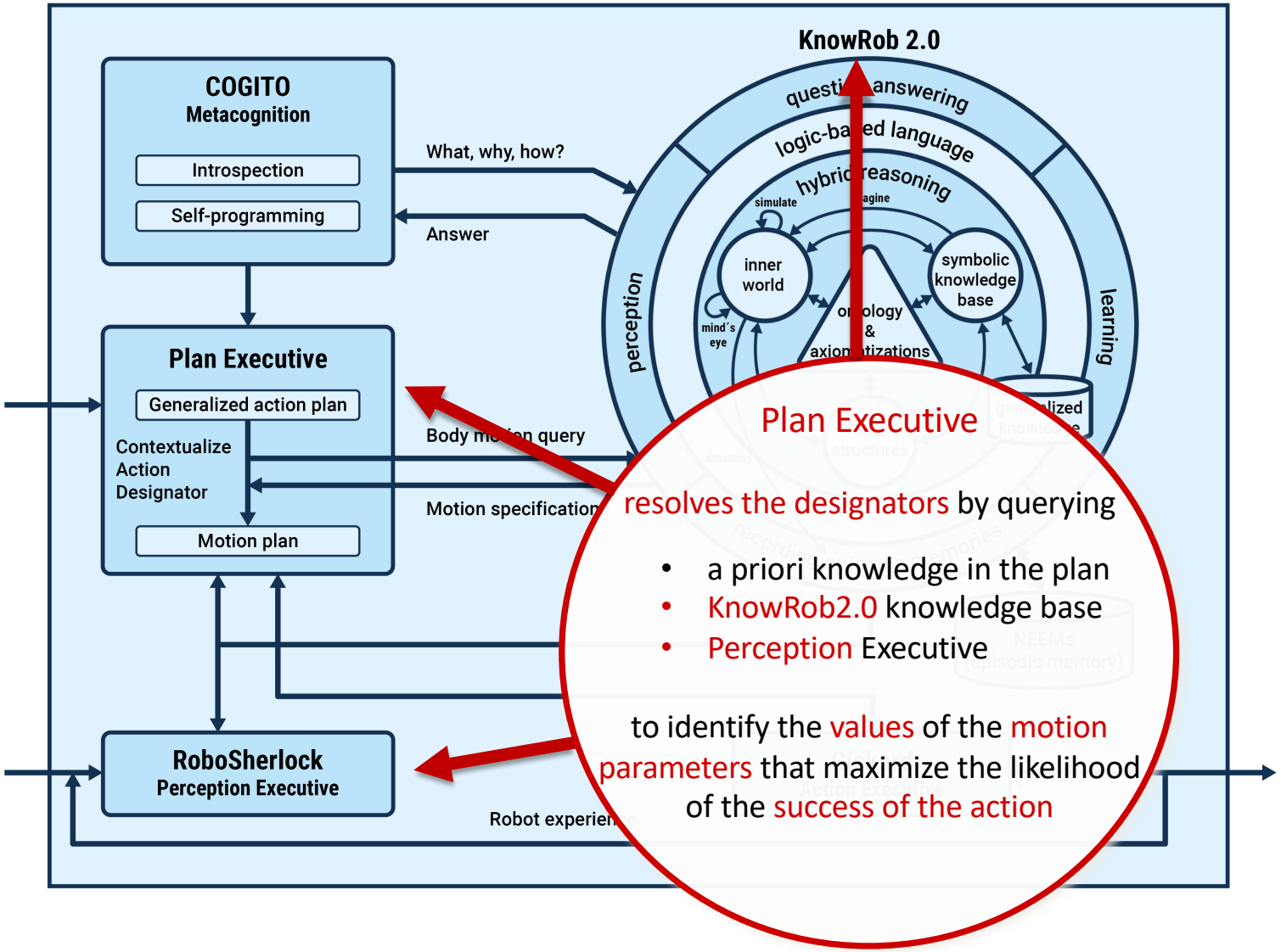


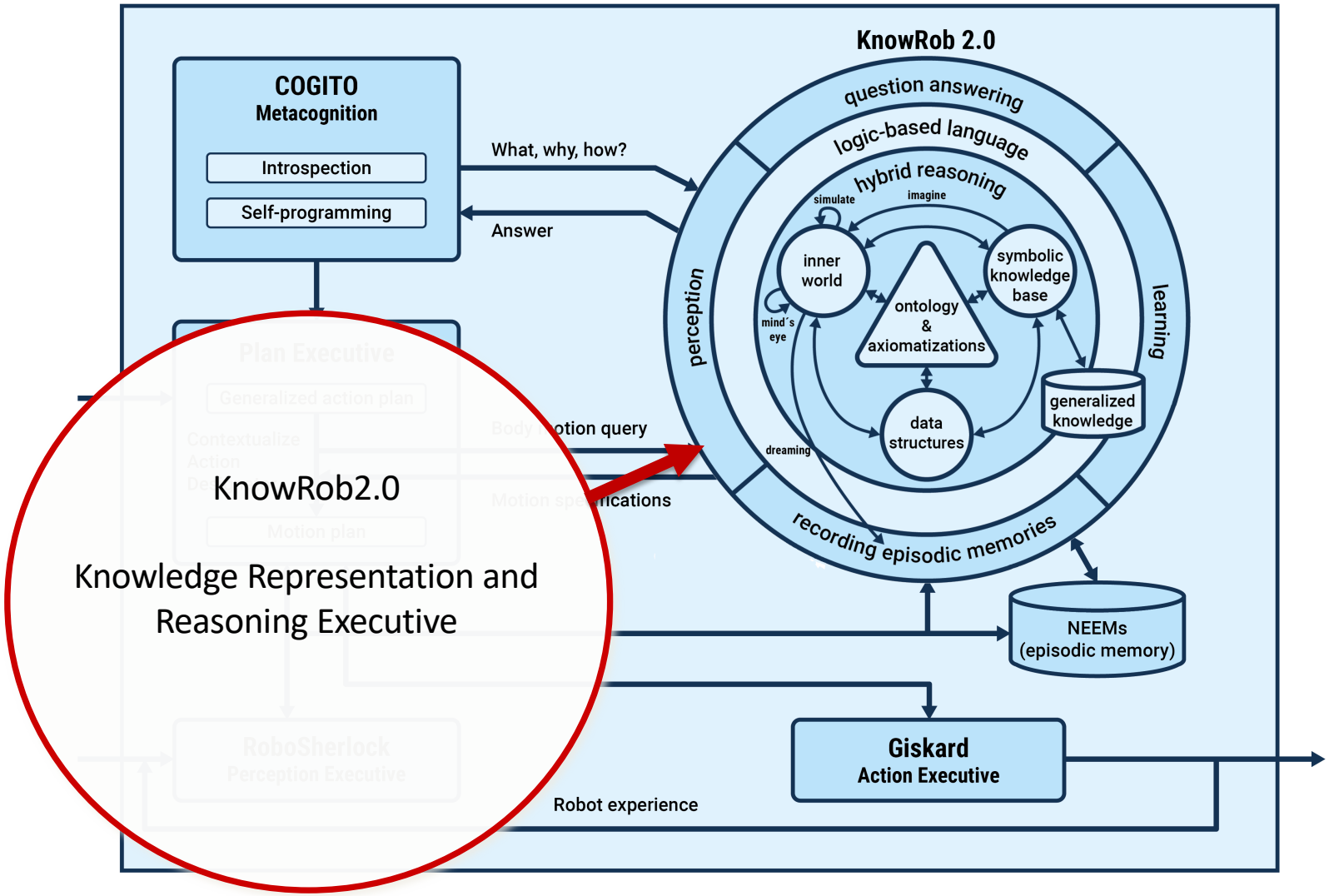
Beetz, M., Mösenlechner, L., and Tenorth, M. 2010. “CRAM—a Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments.” In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, 1012–1017. New York: IEEE.

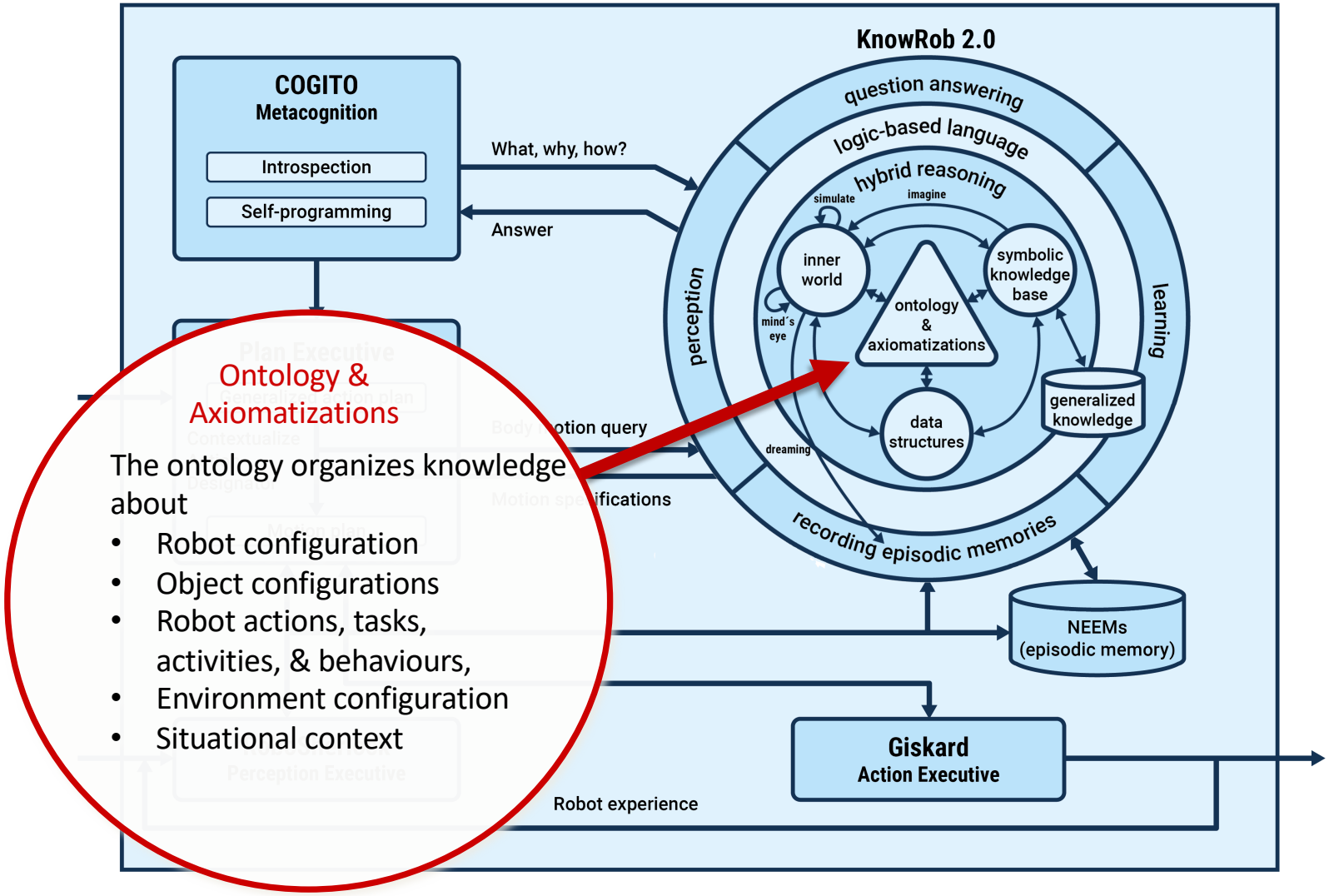
Beetz, M., Kazhoyan, GI, and Vernon, D. 2023. "The CRAM Cognitive Architecture for Robot Manipulation in Everyday Activities", arXiv 2304.14119 [cs.RO].

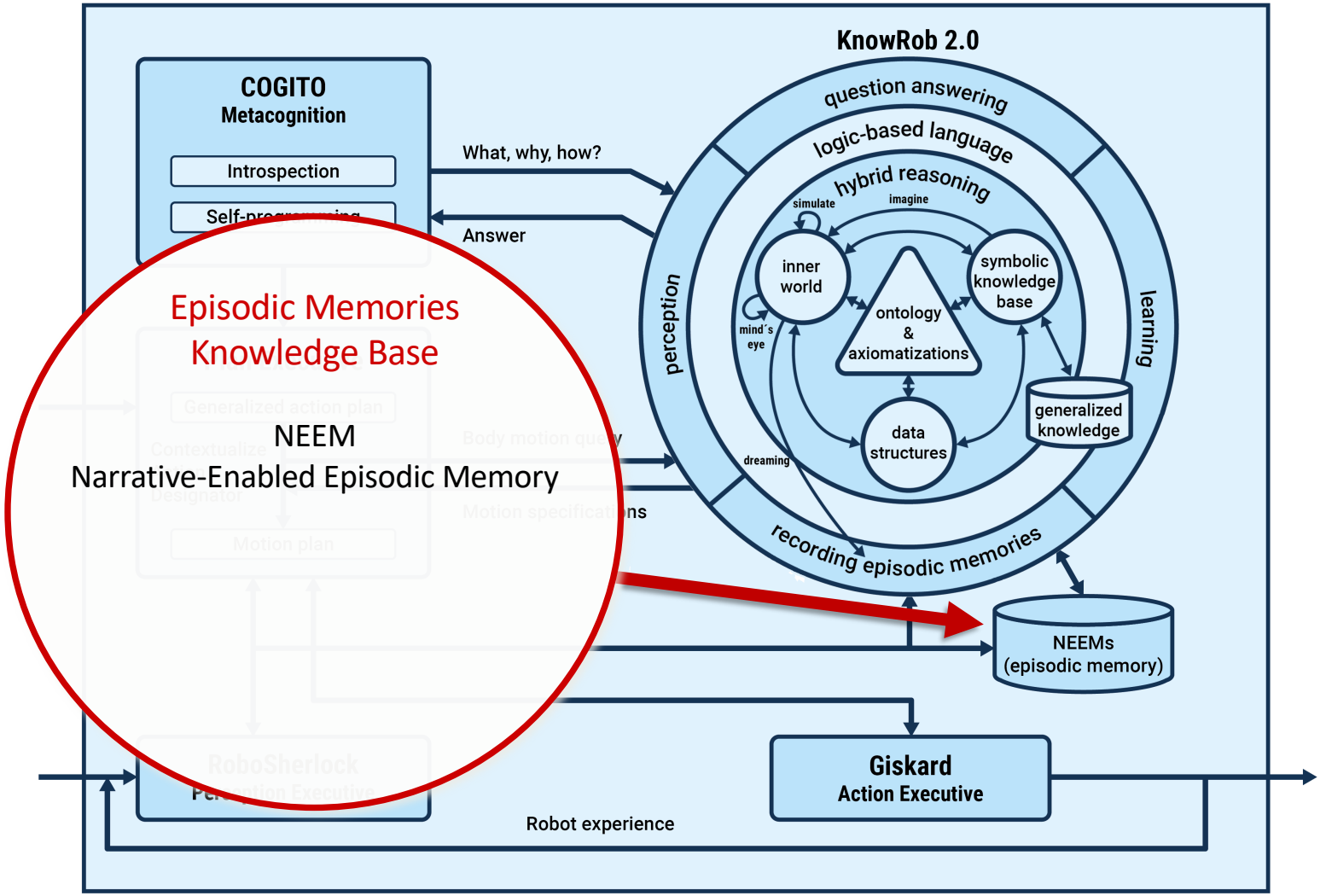

COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
Plan Executive
Generalized action plan
Contextualize
Action
Designator
Motion plan
Body motion query
Motion specifications
RoboSherlock
Perception Executive
Robot experience
Plan Executive
Interprets
Generalized Action Plans
(written in Lisp)
def-plan fetch&place (?object ?destination,
?loc-for-fetch, ?loc-for-place)
with-robot-at-location ?loc-for-fetch
perform (an action
(type fetching)
(object ?object))
with-robot-at-location ?loc-for-place
perform (an action
(type placing)
(object ?object)
(destination ?destination))

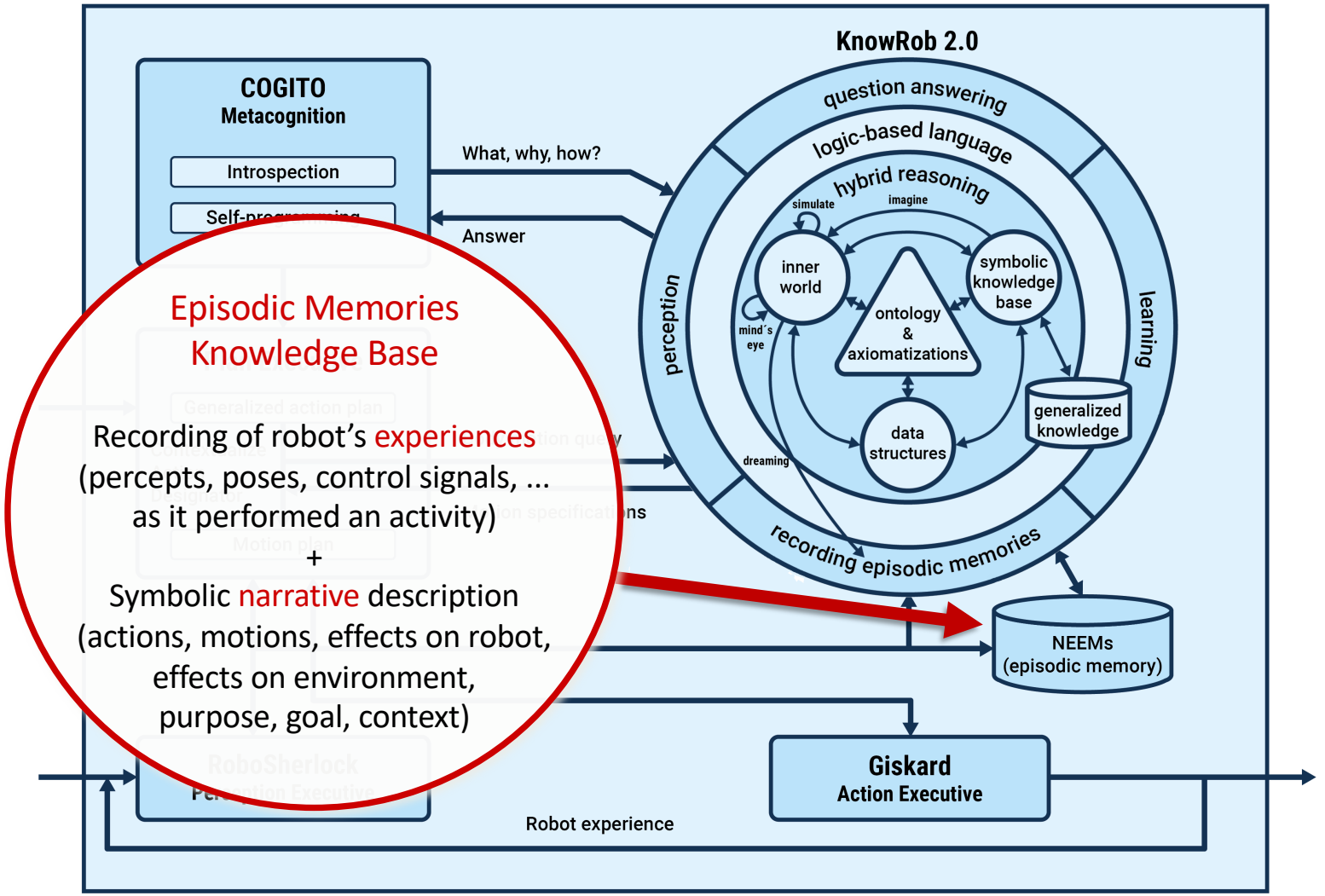

COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
Plan Executive
Generalized action plan
Contextualize
Action
Designator
Motion plan
Body motion query
Motion specifications
RoboSherlock
Perception Executive
Robot experience
Plan Executive
Generalized action plans use designators
(motion, location, object, action)
Facilitate runtime resolution
based on current context
def-plan fetch&place (?object ?destination,
?loc-for-fetch, ?loc-for-place)
with-robot-at-location ?loc-for-fetch
perform (an action
(type fetching)
(object ?object))
with-robot-at-location ?loc-for-place
perform (an action
(type placing)
(object ?object)
(destination ?destination))
action
designator

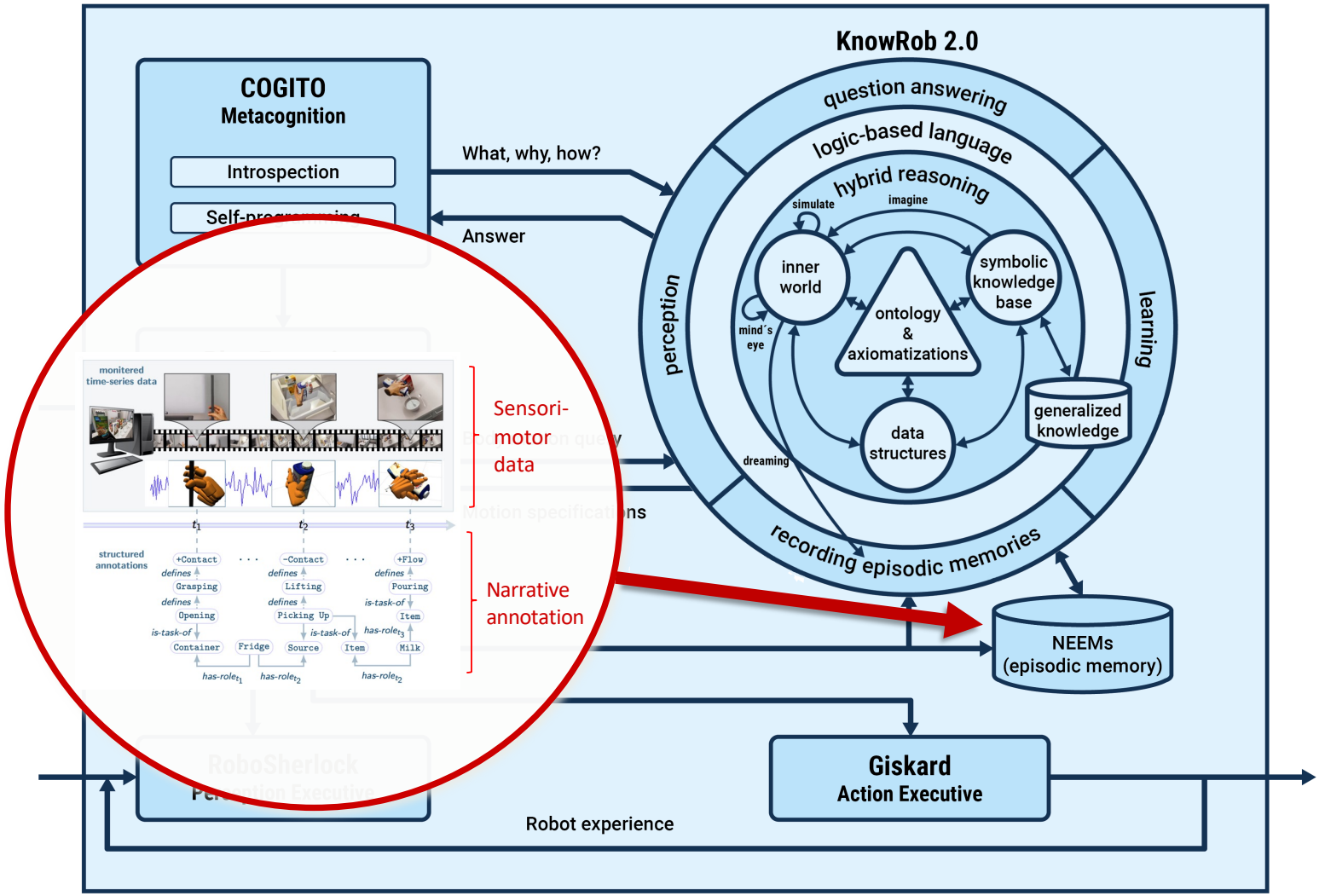

KnowRob 2.0
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
Plan Executive
Generalized action plan
Contextualize
Action
Designator
Motion plan
Body motion query
Motion specification
RoboSherlock
Perception Executive
Robot experience
perception
learning
inner world
symbolic knowledge base
simulate
mind´s eye
Plan Executive
resolves the designators by querying
a priori knowledge in the plan
KnowRob2.0 knowledge base
Perception Executive
to identify the values of the motion parameters that maximize the likelihood of the success of the action

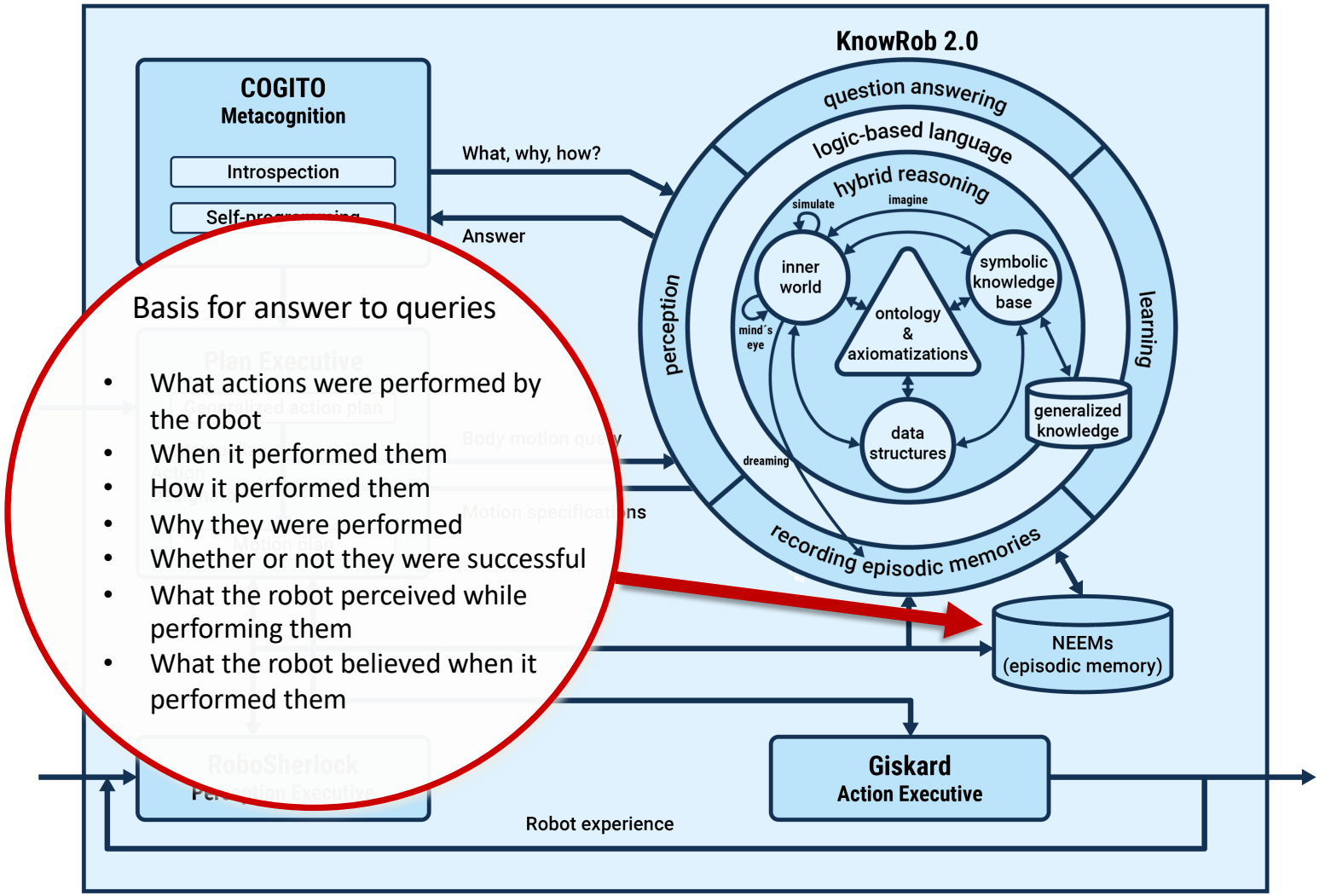

KnowRob 2.0
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
KnowRob2.0
Knowledge Representation and Reasoning Executive

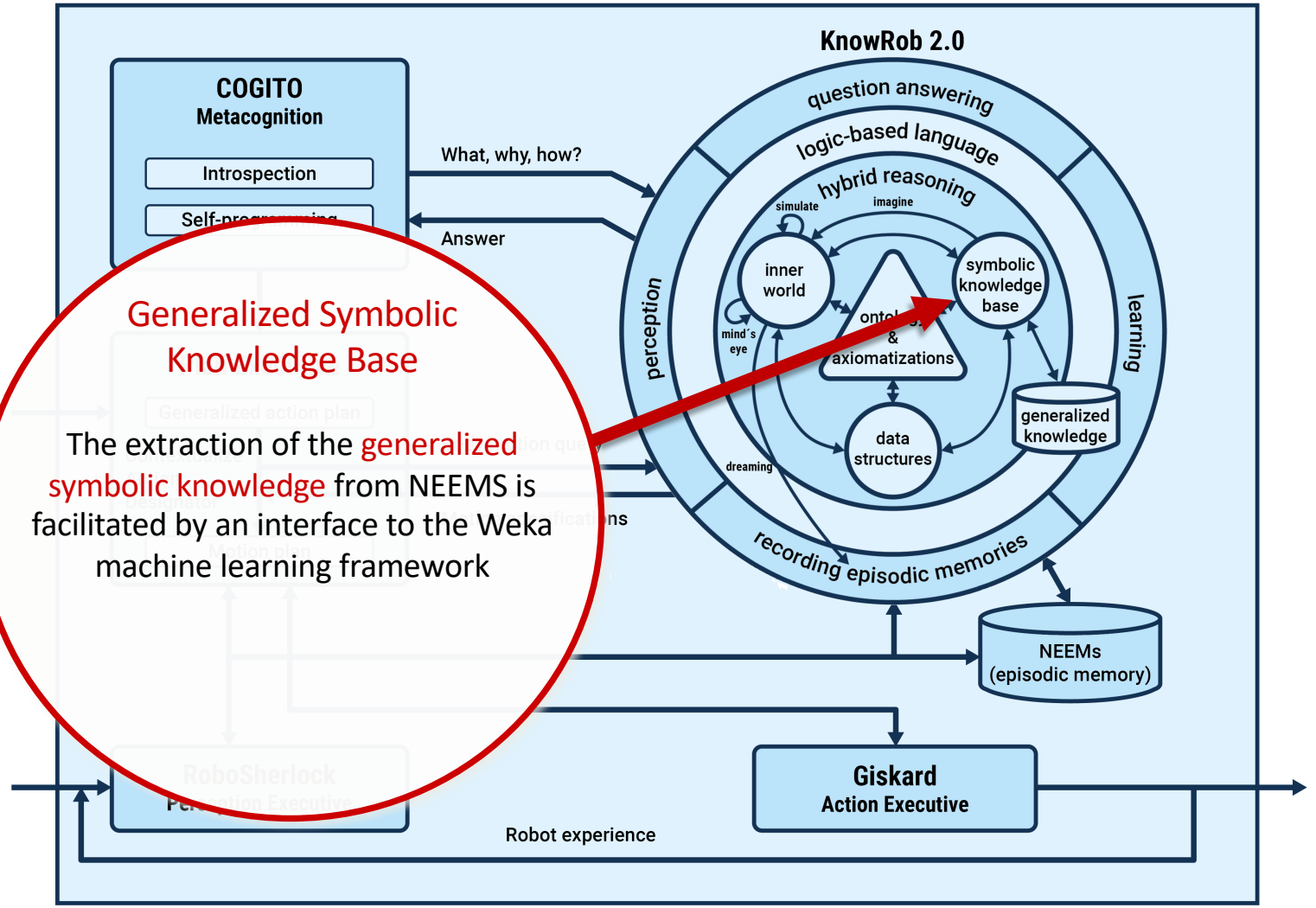

KnowRob 2.0
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
Ontology & Axiomatizations
The ontology organizes knowledge about
• Robot configuration
• Object configurations
• Robot actions, tasks, activities, & behaviours,
• Environment configuration
• Situational context

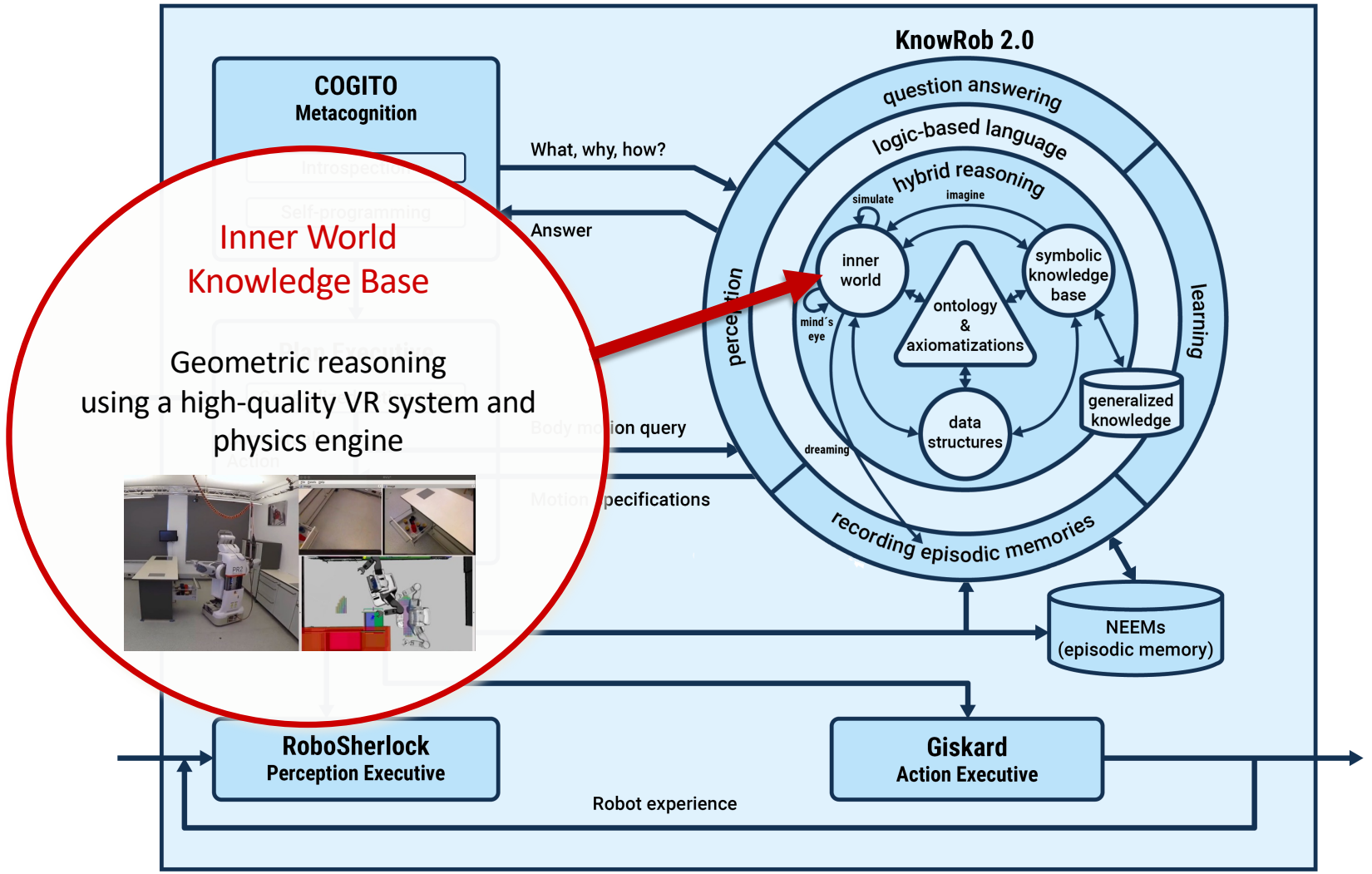

KnowRob 2.0
COGITO
Metacognition
Introspection
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
Episodic Memories Knowledge Base
NEEM
Narrative-Enabled Episodic Memory
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience

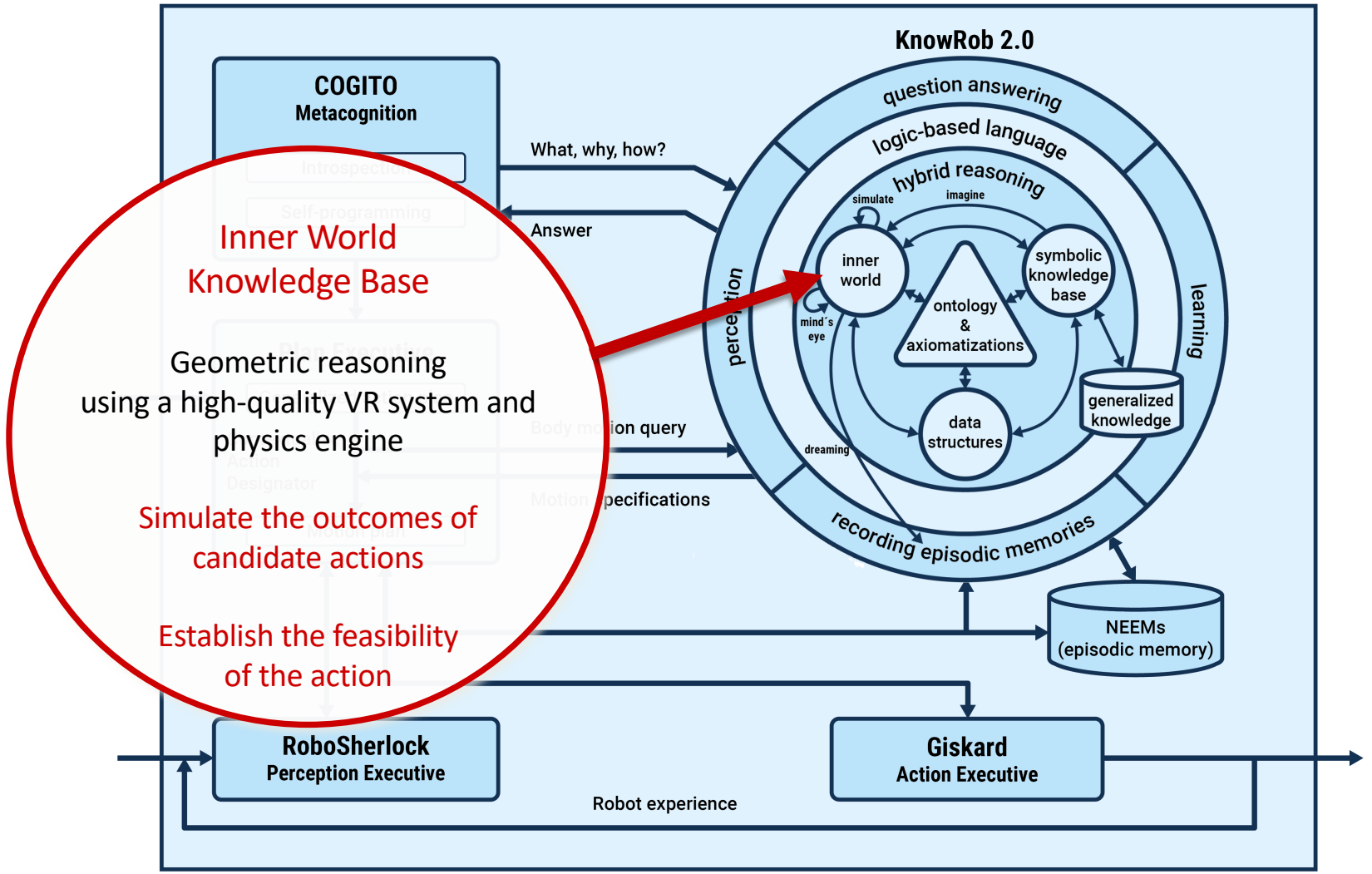

KnowRob 2.0
COGITO
Metacognition
Introspection
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
Episodic Memories Knowledge Base
Recording of robot's experiences (percepts, poses, control signals, ... as it performed an activity)
+
Symbolic narrative description (actions, motions, effects on robot, effects on environment, purpose, goal, context)

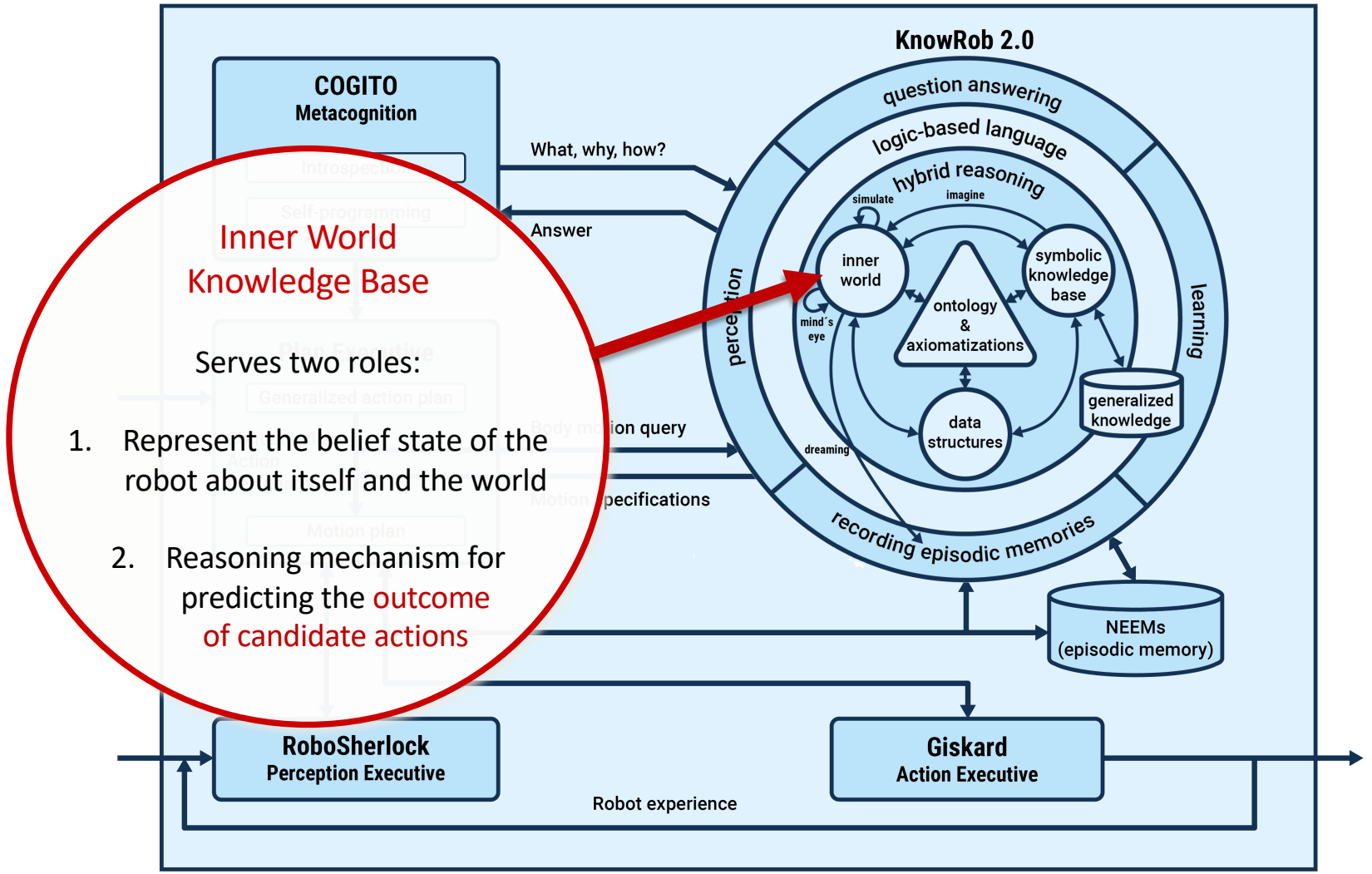

KnowRob 2.0
COGITO
Metacognition
Introspection
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
monitered time-series data
Sensori-motor data
structured annotations
Narrative annotation
Giskard
Action Executive
Robot experience

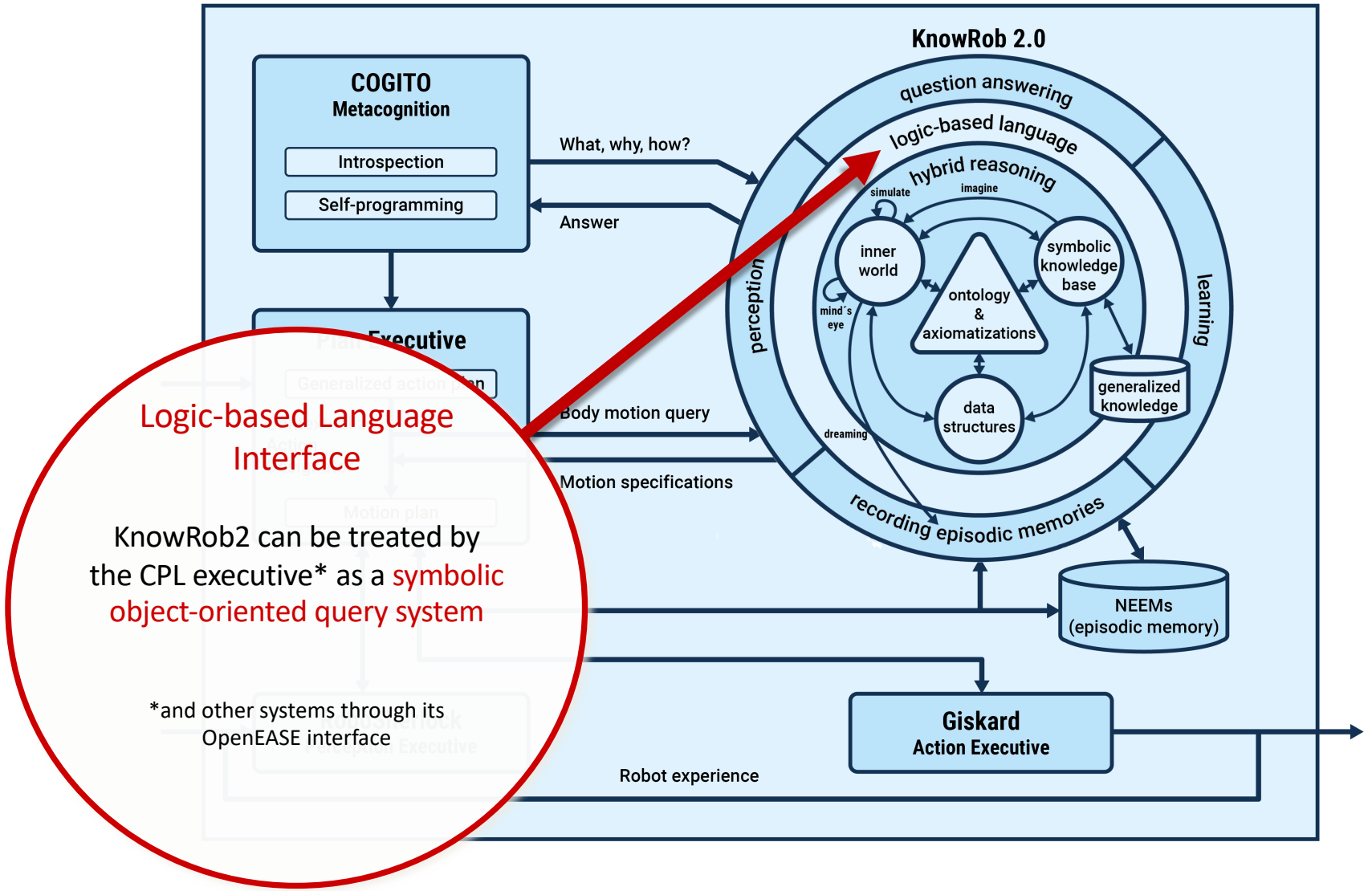

KnowRob 2.0
COGITO
Metacognition
Introspection
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
Basis for answer to queries
What actions were performed by the robot
When it performed them
How it performed them
Why they were performed
Whether or not they were successful
What the robot perceived while performing them
What the robot believed when it performed them

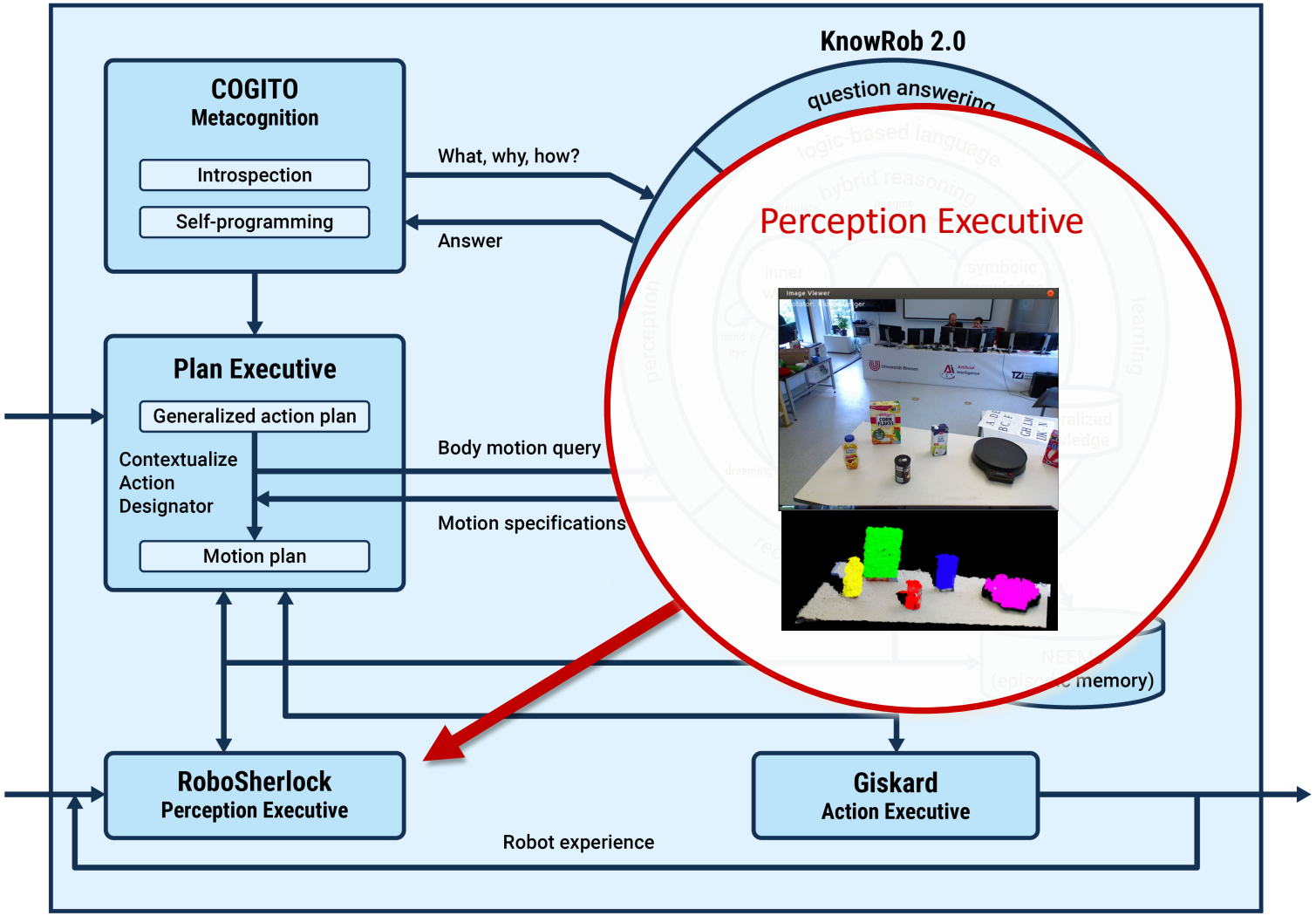

KnowRob 2.0
COGITO
Metacognition
Introspection
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
mind´s eye
ontologies & axiomatizations
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
Generalized Symbolic Knowledge Base
The extraction of the generalized symbolic knowledge from NEEMS is facilitated by an interface to the Weka machine learning framework

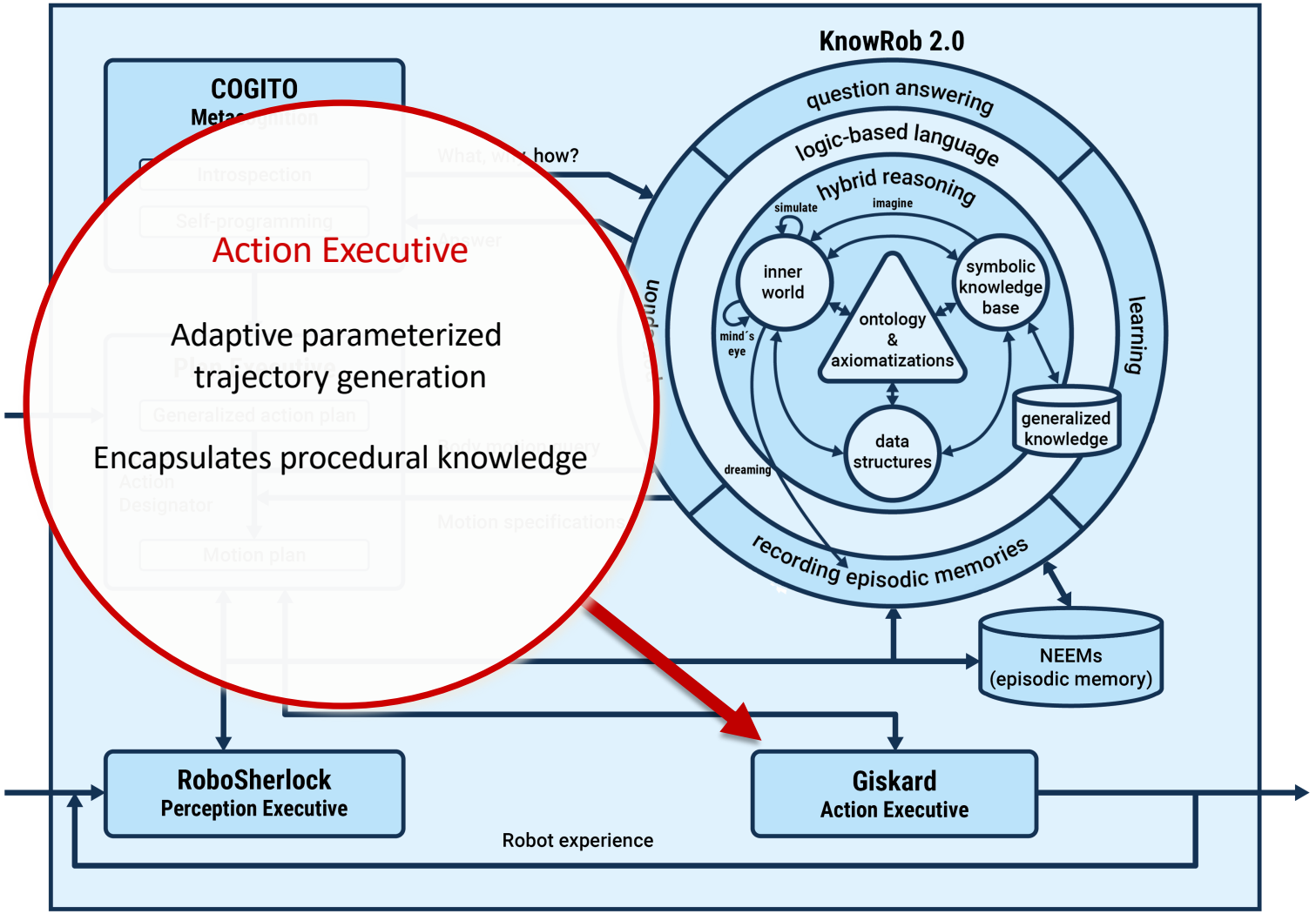

KnowRob 2.0
COGITO
Metacognition
What, why, how?
Answer
Inner World
Knowledge Base
Geometric reasoning
using a high-quality VR system and
physics engine
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
Body motion query
Motion specifications
NEEMs
(episodic memory)
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience

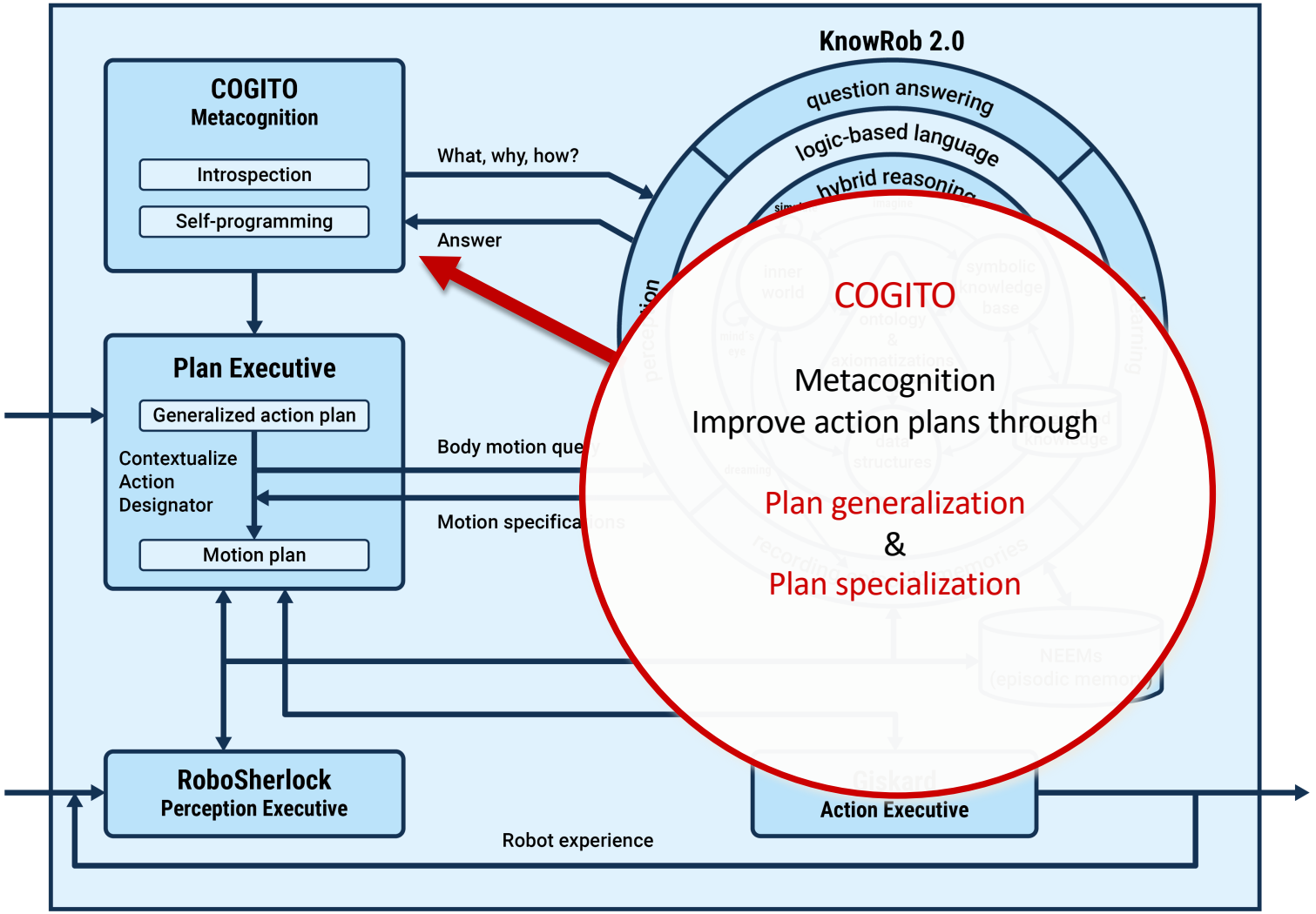

KnowRob 2.0
COGITO
Metacognition
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
Body motion query
Motion specifications
NEEMs (episodic memory)
Inner World Knowledge Base
Geometric reasoning using a high-quality VR system and physics engine
Simulate the outcomes of candidate actions
Establish the feasibility of the action
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience

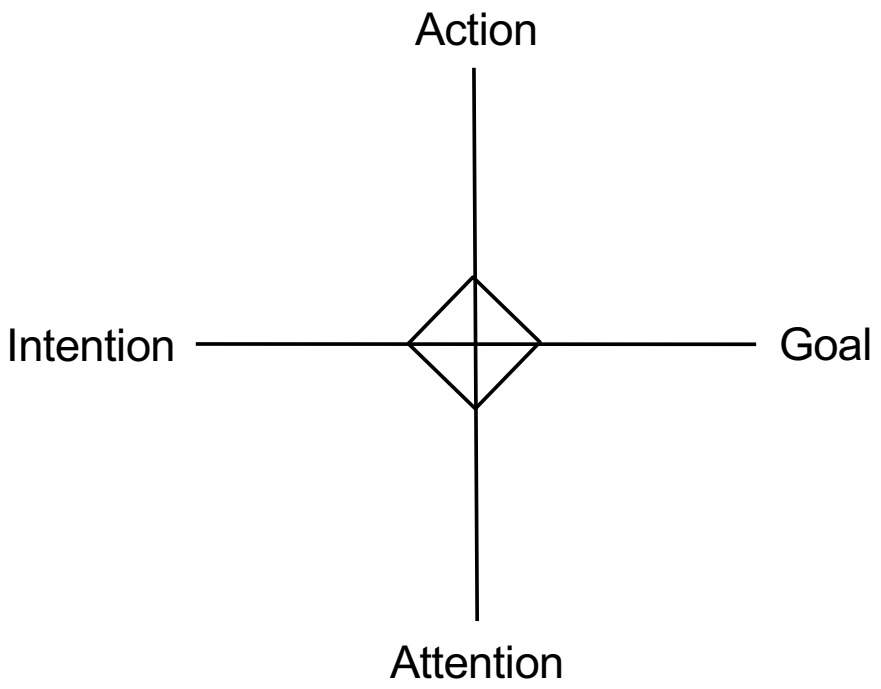

KnowRob 2.0
COGITO
Metacognition
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
Body motion query
Motion specifications
NEEMs (episodic memory)
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience
Inner World Knowledge Base
Serves two roles:
1. Represent the belief state of the robot about itself and the world
2. Reasoning mechanism for predicting the outcome of candidate actions

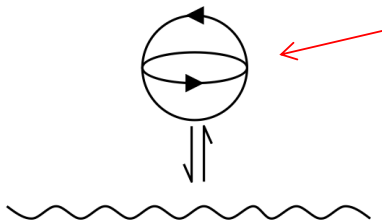

KnowRob 2.0
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
Executive
Body motion query
Motion specifications
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
Logic-based Language Interface
KnowRob2 can be treated by the CPL executive* as a symbolic object-oriented query system
*and other systems through its OpenEASE interface


KnowRob 2.0
question answering
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
Perception Executive
Plan Executive
Generalized action plan
Contextualize
Action
Designator
Motion plan
Body motion query
Motion specifications
memory)
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience

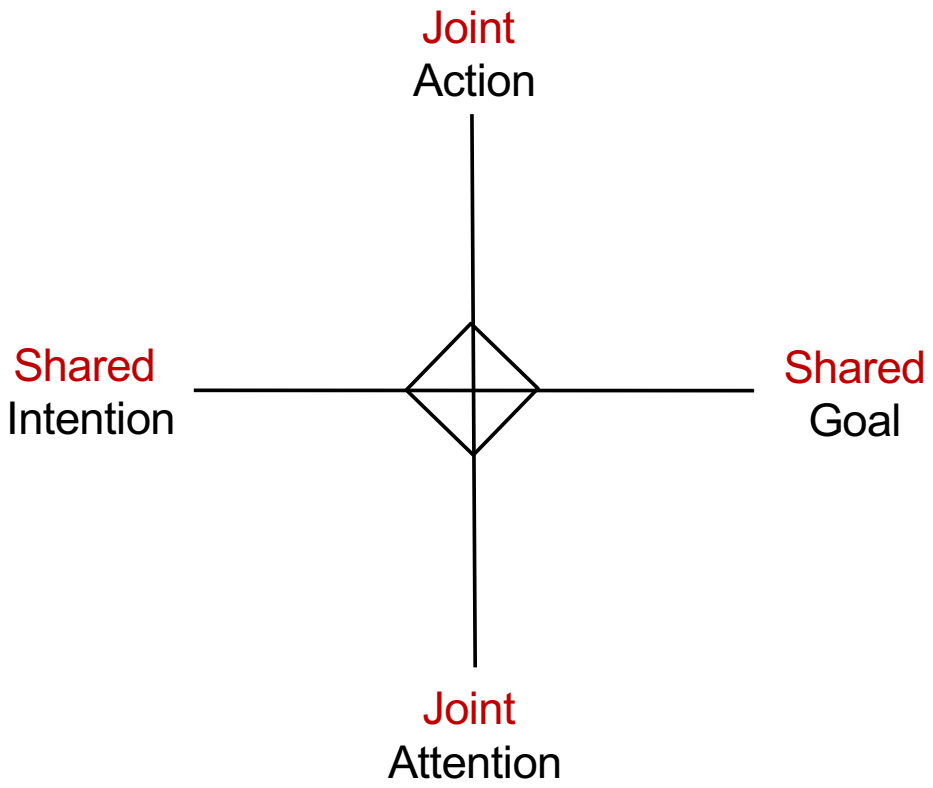

KnowRob 2.0
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
learning
recording episodic memories
NEEMs (episodic memory)
COGITO
Action Executive
Adaptive parameterized trajectory generation
Encapsulates procedural knowledge
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience

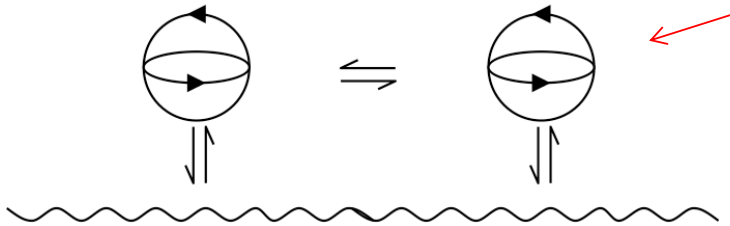

KnowRob 2.0
question answering
logic-based language
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Answer
Plan Executive
Generalized action plan
Contextualize
Action
Designator
Motion plan
Body motion que
Motion specifica
RoboSherlock
Perception Executive
Action Executive
Robot experience
COGITO
Metacognition
Improve action plans through
Plan generalization
&
Plan specialization



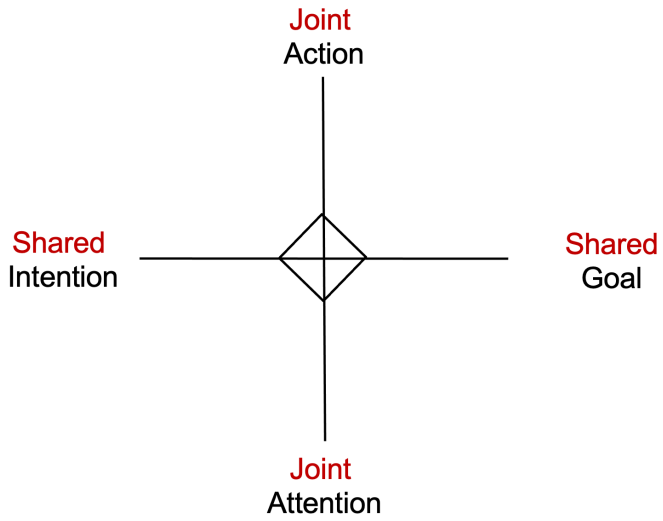


An individual autonomous agent interacting with its environment

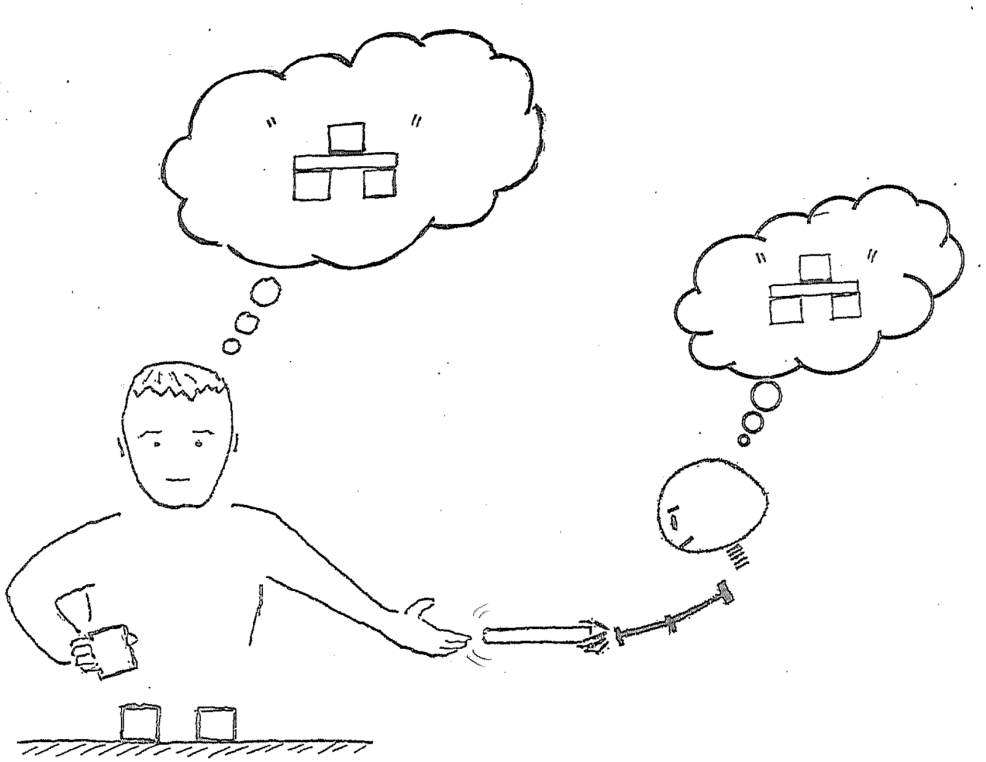


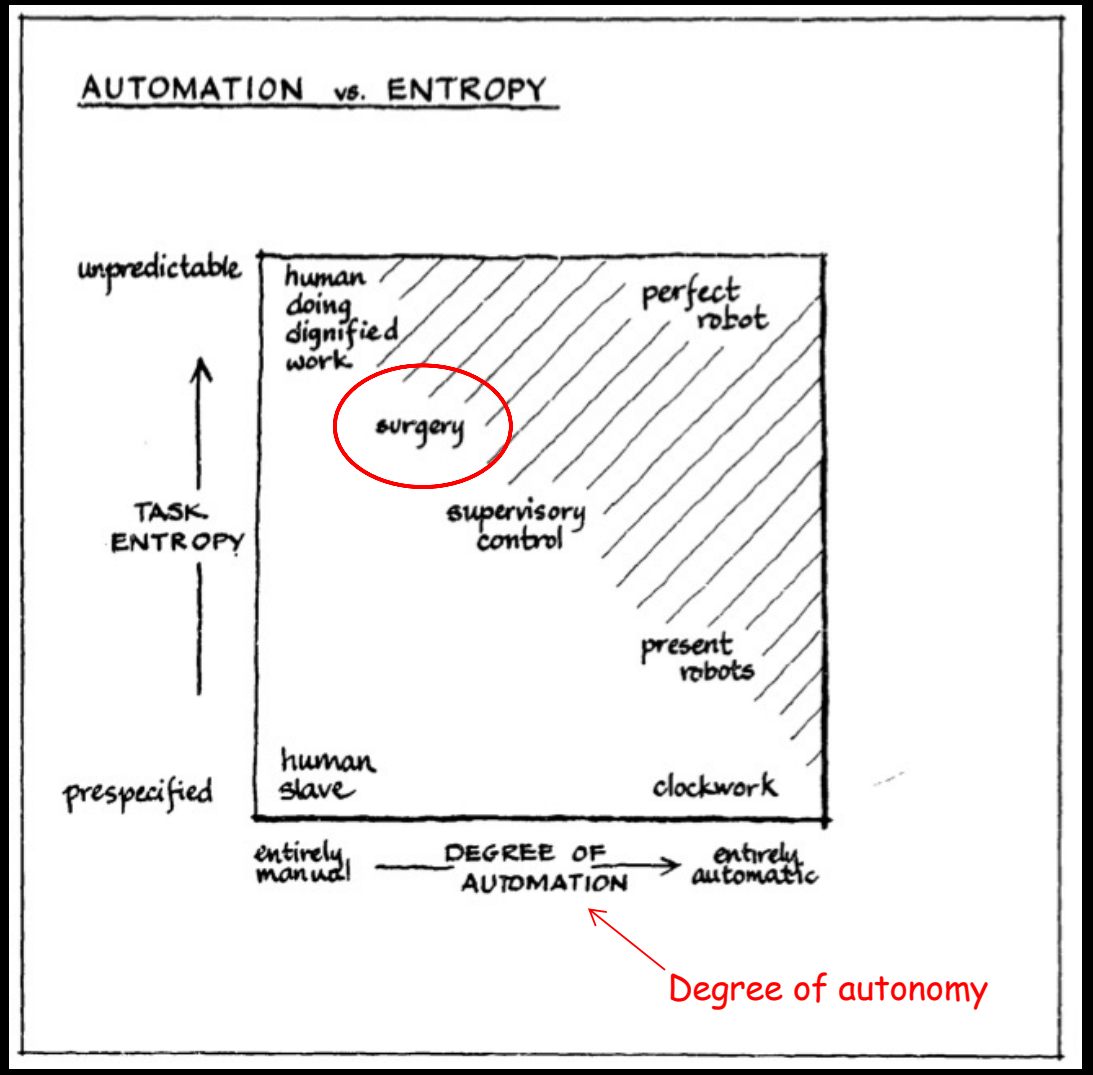


Two autonomous agents interacting with their environment and each other



Blackbox peer-to-peer collaboration
by two autonomous agents
⇒
Mutual Theory of Mind





T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.

Degrees of Autonomy

1. Human does the whole job up to the point of turning it over to the computer to implement

2. Computer helps by determining options

...

5. Computer selects an action and implements it if human approves

6. Computer selects action, informs human in plenty of time to stop it

7. Computer does whole job and necessarily tells human what it did

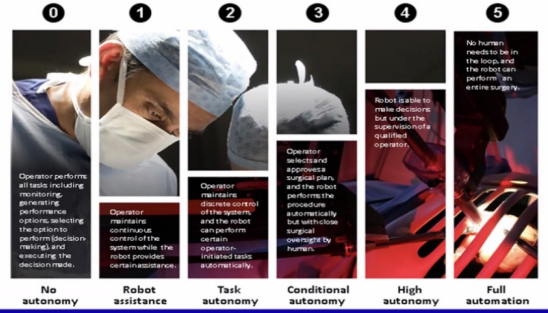
...

10. Computer does whole job, if it decides it should be done, and if so, tells human, if it decides she should be told

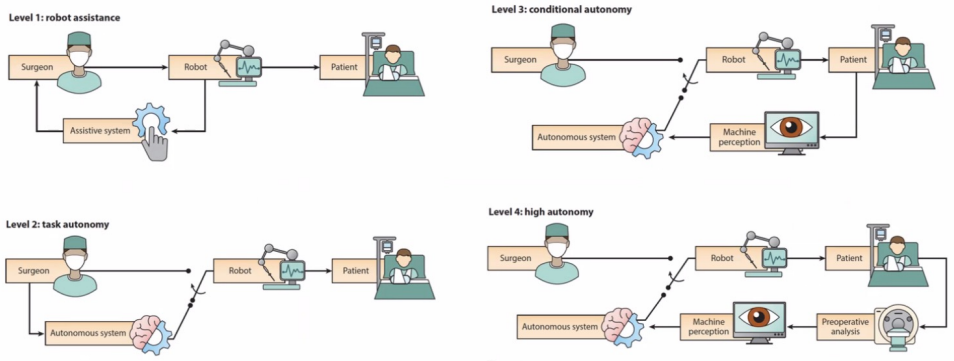
T. B. Sheridan and W. L. Verplank. Human and computer control for undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.

Autonomy in Surgery

Referring to SAE and ISO/IEC frameworks, Yang and Haidegger presented two six-level classifications for surgical robots,



Visual Representation of Levels of Autonomy

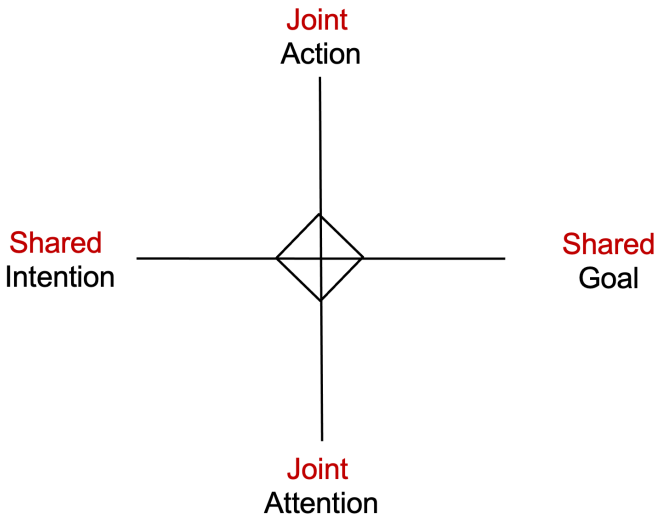


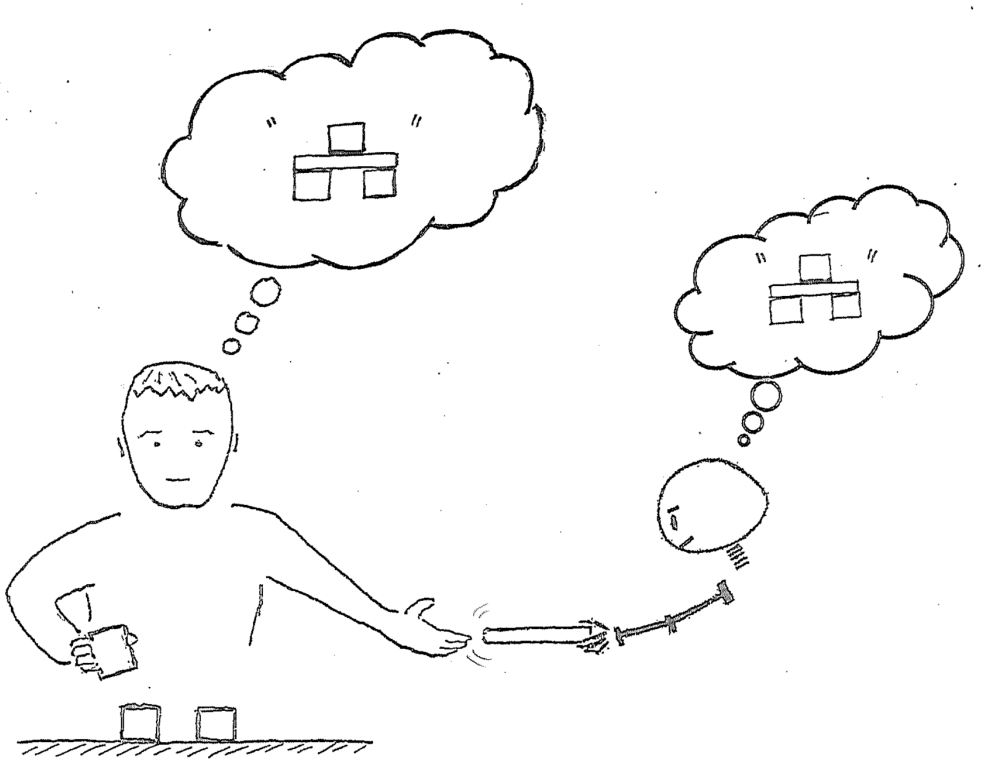
P. Fiorini, Fundamentals of Autonomy, ICRA 2024 Workshop, Autonomy in Robotic Surgery: State of the Art, Technical and Regulatory Challenges for Clinical Application, 2024.

Blackbox peer-to-peer collaboration by two fully autonomous agents

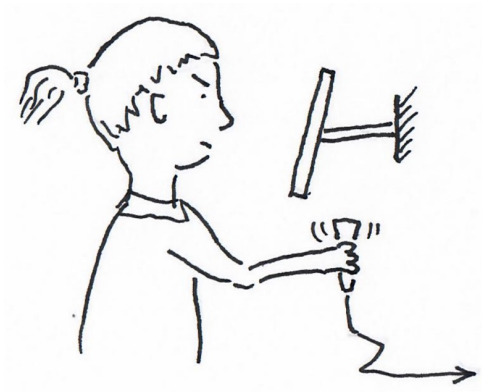
⇒

Mutual theory of mind

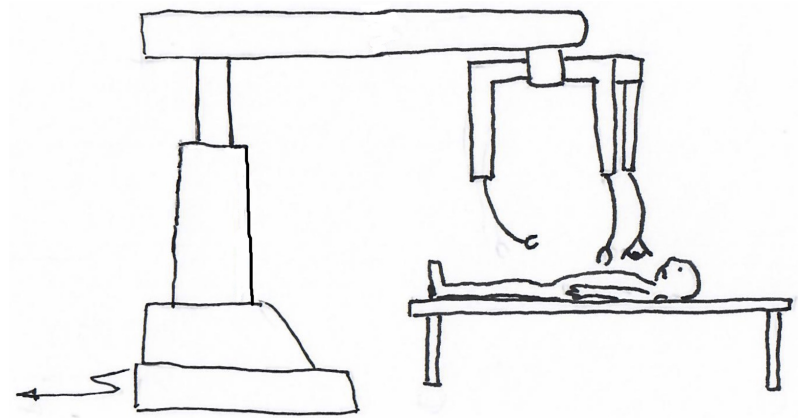




Blackbox preemptive collaboration
by the human supervisor

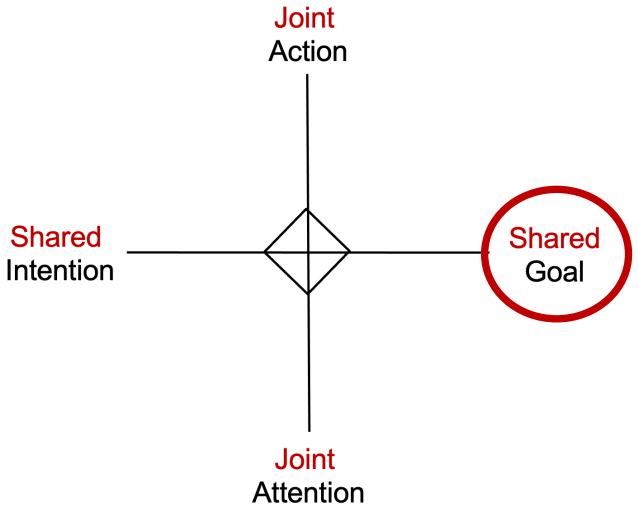


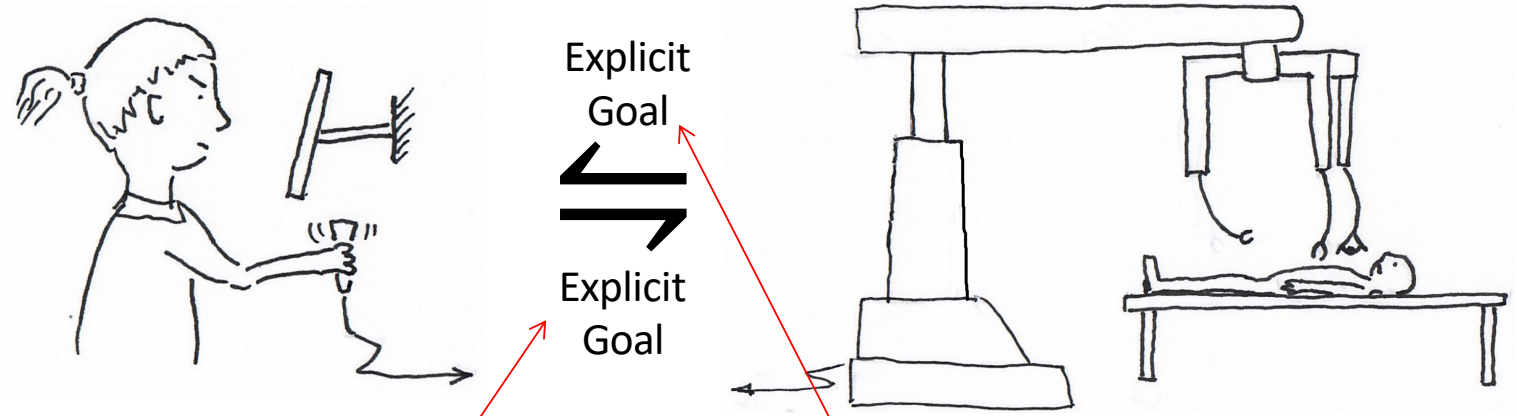
Whitebox subservient collaboration
by the surgical robot



Blackbox preemptive collaboration
by the human supervisor
⇒
Explicitly shared goal

Whitebox subservient collaboration
by the surgical robot
⇒
Explicitly shared goal



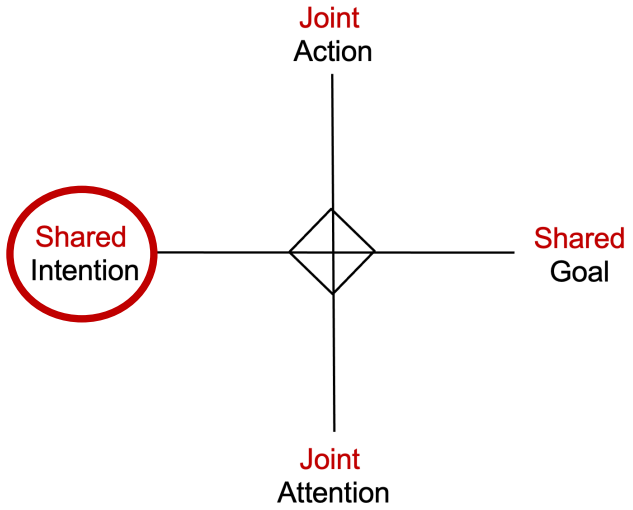


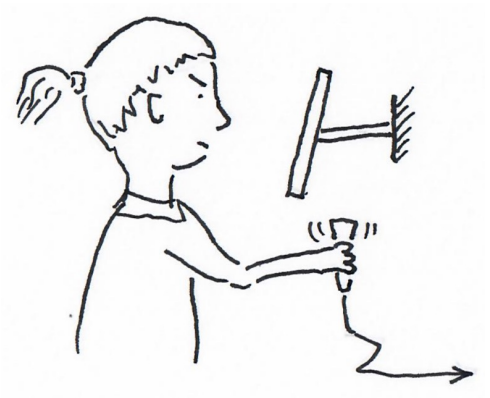
Encapsulated in the action plan

Encapsulated in the action plan

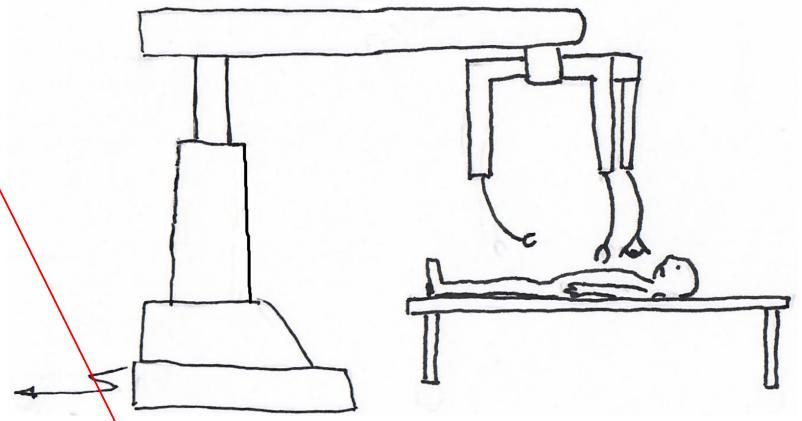
Blackbox preemptive collaboration
by the human supervisor
⇒
Explicitly shared intention

Whitebox subservient collaboration
by the surgical robot
⇒
Explicitly shared intention





Explicit
Intention
⇋
Explicit
Intention

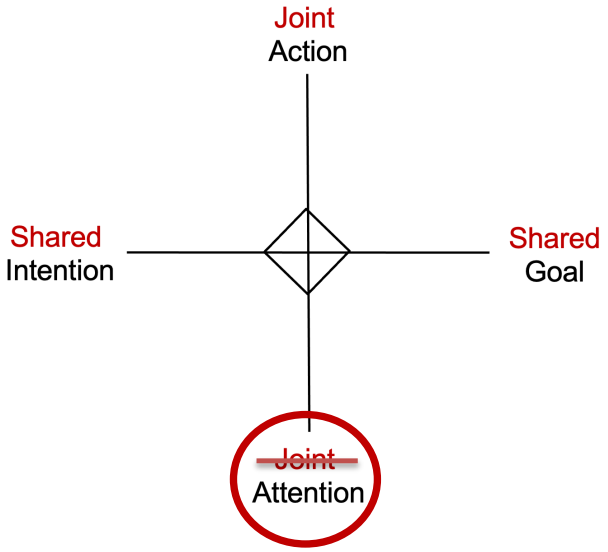


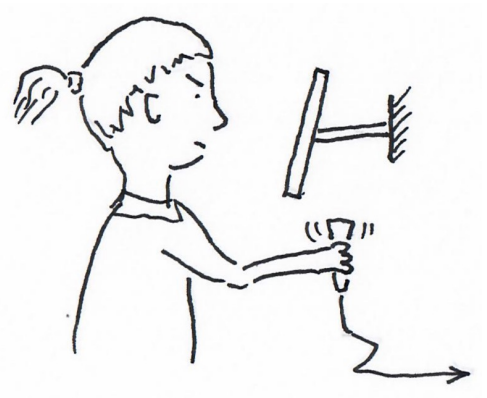
Encapsulated in the action plan

Revealed by internal simulation of action
& prospective reasoning

Blackbox preemptive collaboration
by the human supervisor
⇒
Mutually exclusive control

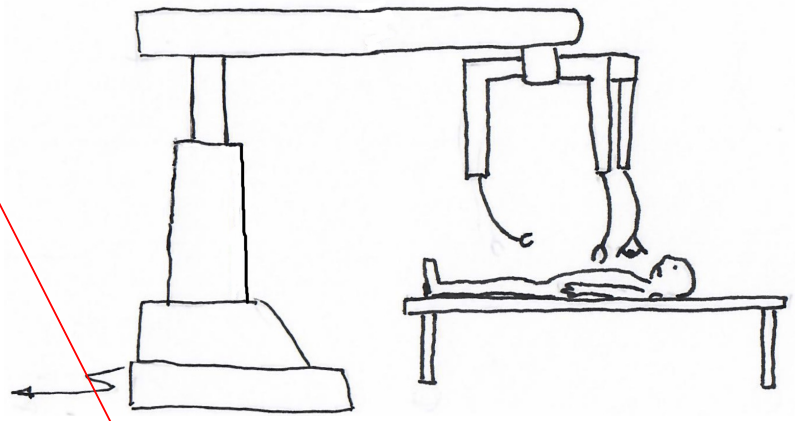
Whitebox subservient collaboration
by the surgical robot
⇒
Mutually exclusive control





Single
Attention

Single
Attention

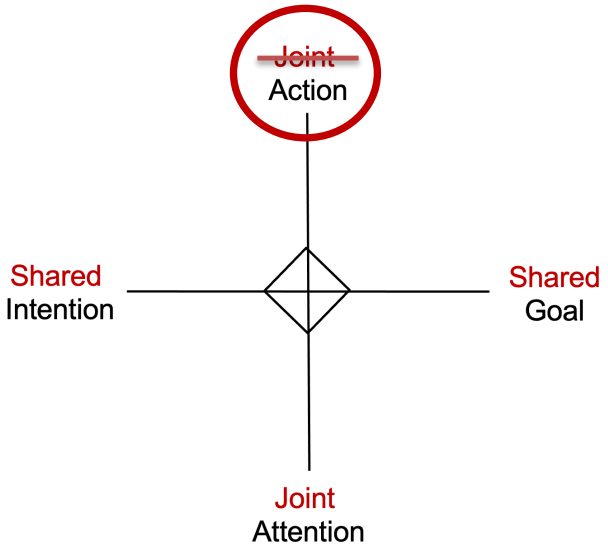


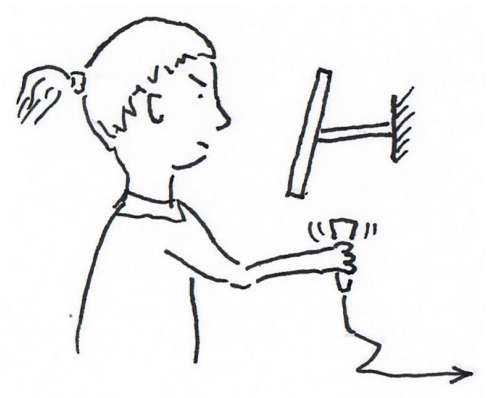
Human supervisor is either
passive or fully in control

Suspended or
Exhibiting Degree 6 autonomy

Blackbox preemptive collaboration
by the human supervisor
⇒
Mutually exclusive control

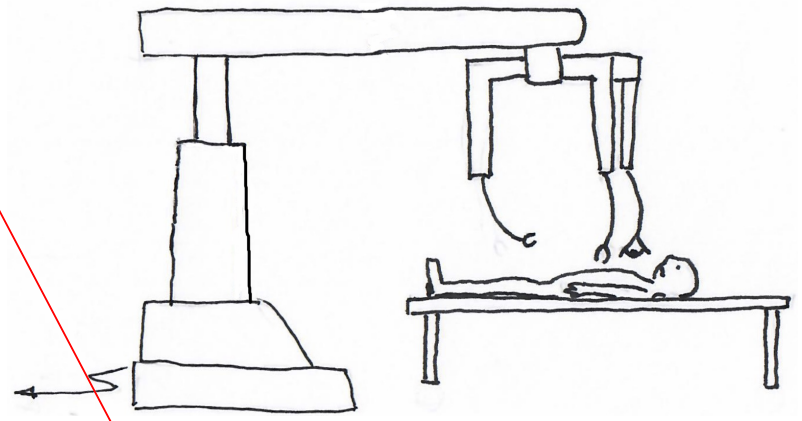
Whitebox subservient collaboration
by the surgical robot
⇒
Mutually exclusive control





Single
Action

Single
Action

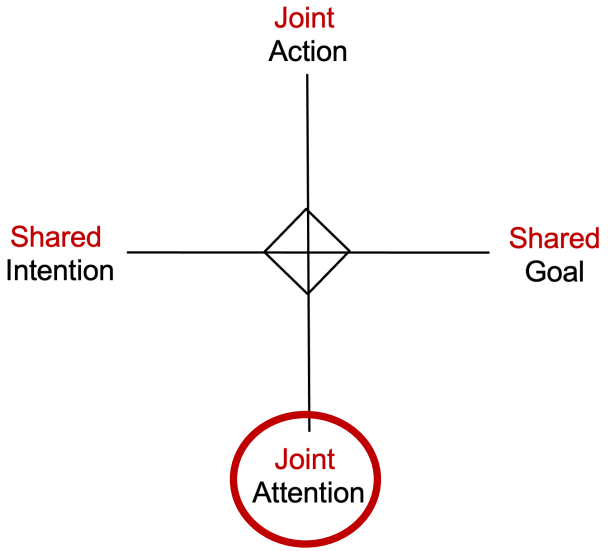


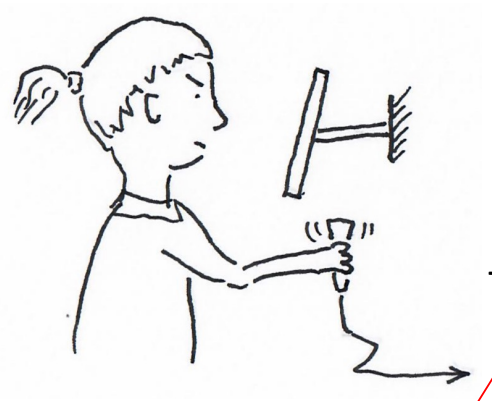
Human supervisor is either
passive or fully in control

Suspended or
Exhibiting Degree 6 autonomy

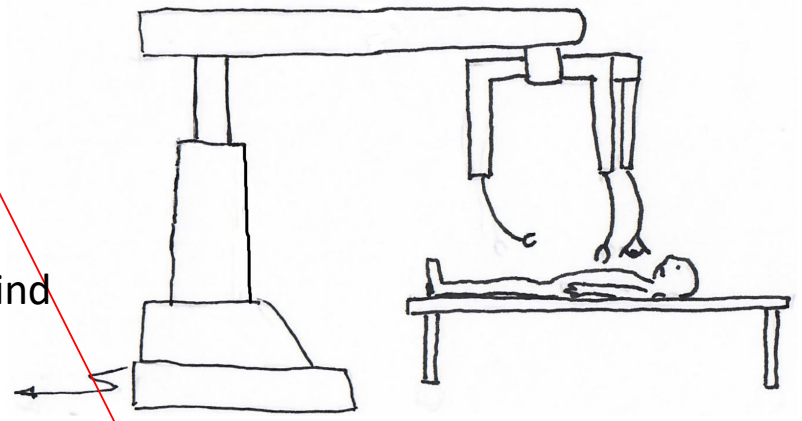
Blackbox preemptive collaboration
by the human supervisor
⇒
Collaborative control

Whitebox subservient collaboration
by the surgical robot
⇒
Collaborative control





Explicit
Attention
⇋
Theory of Mind

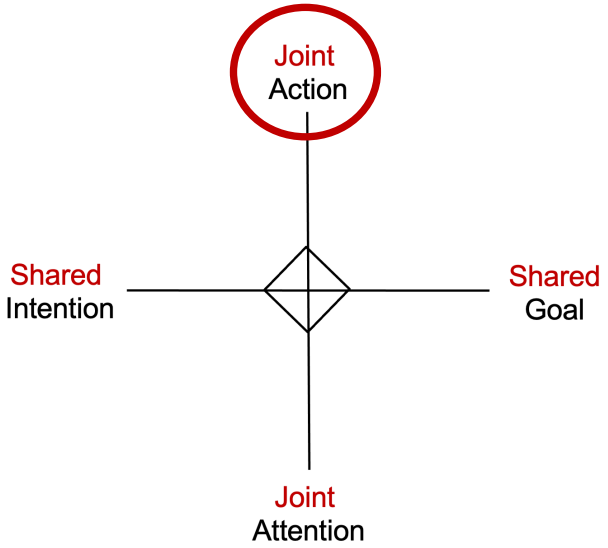


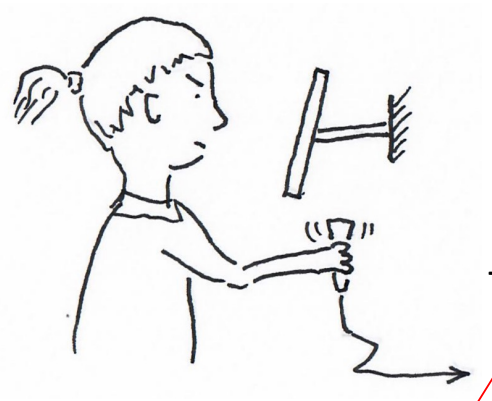
Topical research challenge:
how does a cognitive surgical
robot anticipate the actions of
a human supervisor?

Revealed by internal simulation of action

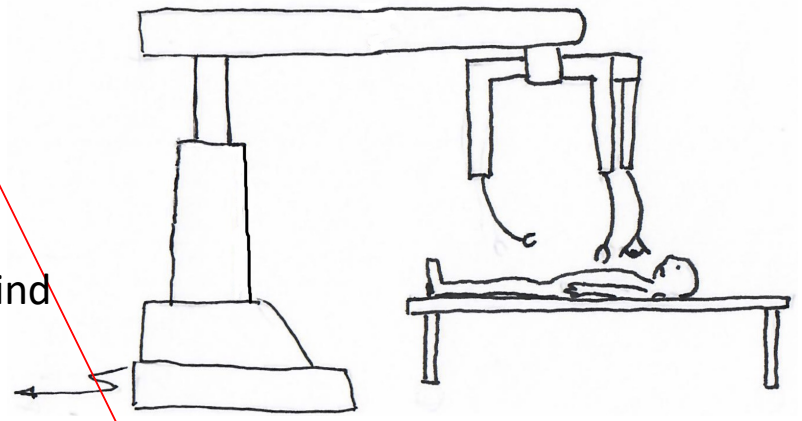
Blackbox preemptive collaboration
by the human supervisor
⇒
Collaborative control

Whitebox subservient collaboration
by the surgical robot
⇒
Collaborative control





Explicit
Action
⇋
Theory of Mind



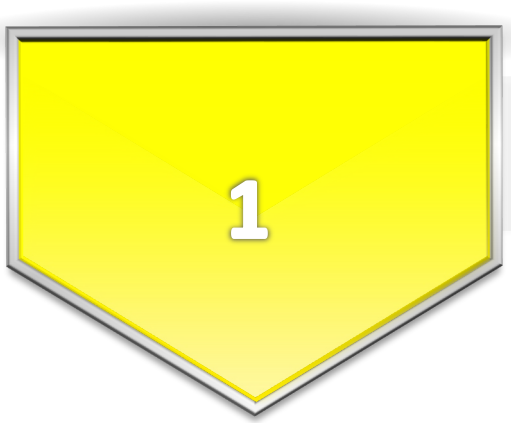
Topical research challenge:
how does a cognitive surgical
robot anticipate the actions of
a human supervisor?

Revealed by internal simulation of action

Can Cognitive Architectures Support Autonomy in Robotic Surgery?

Yes!

Implications



Surgical robot system architecture should be a cognitive architecture

Plan actions

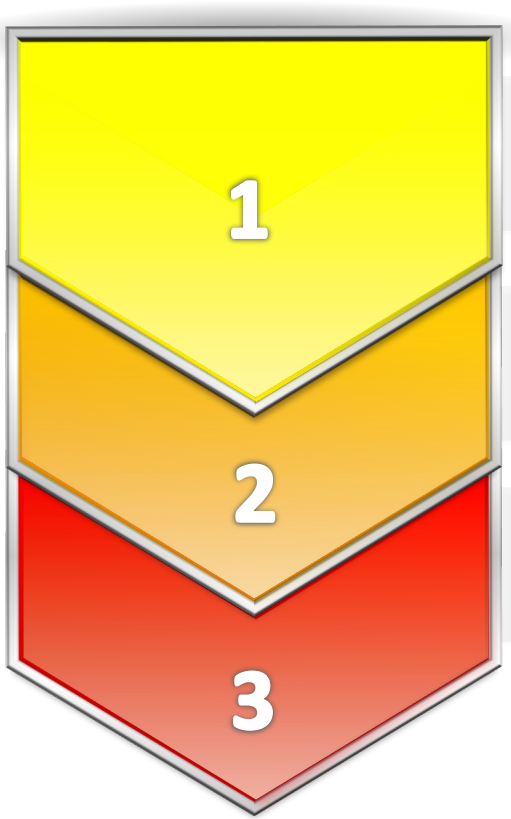
Anticipate outcome before acting

Implications

Surgical robot system architecture should be a cognitive architecture
Plan actions
Anticipate outcome before acting

Cognitive architecture must allow white box prospection & reasoning
Human supervisor does not have to form a Theory of Mind
Intentions and expected outcomes are revealed to the human supervisor

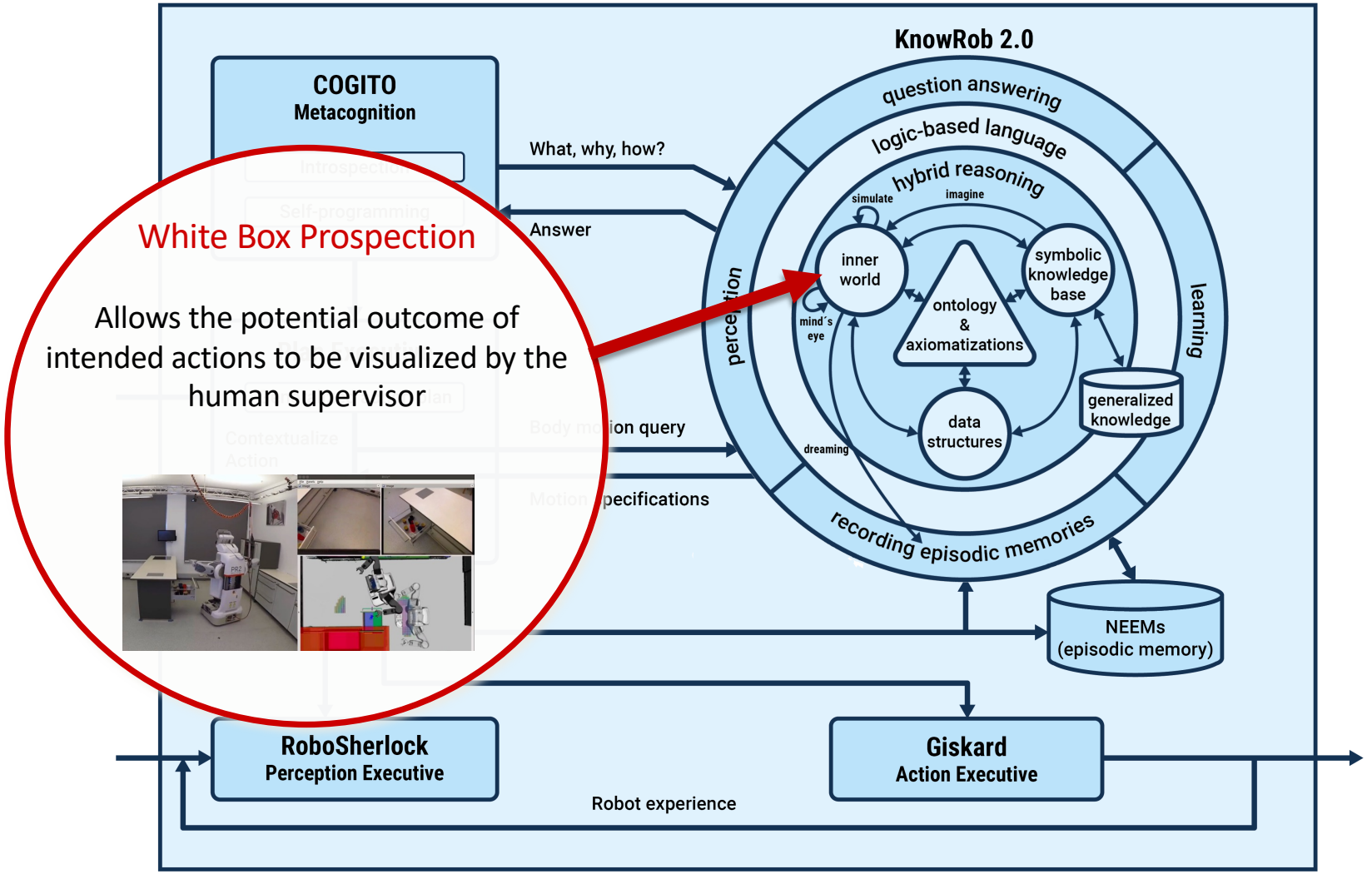
Implications

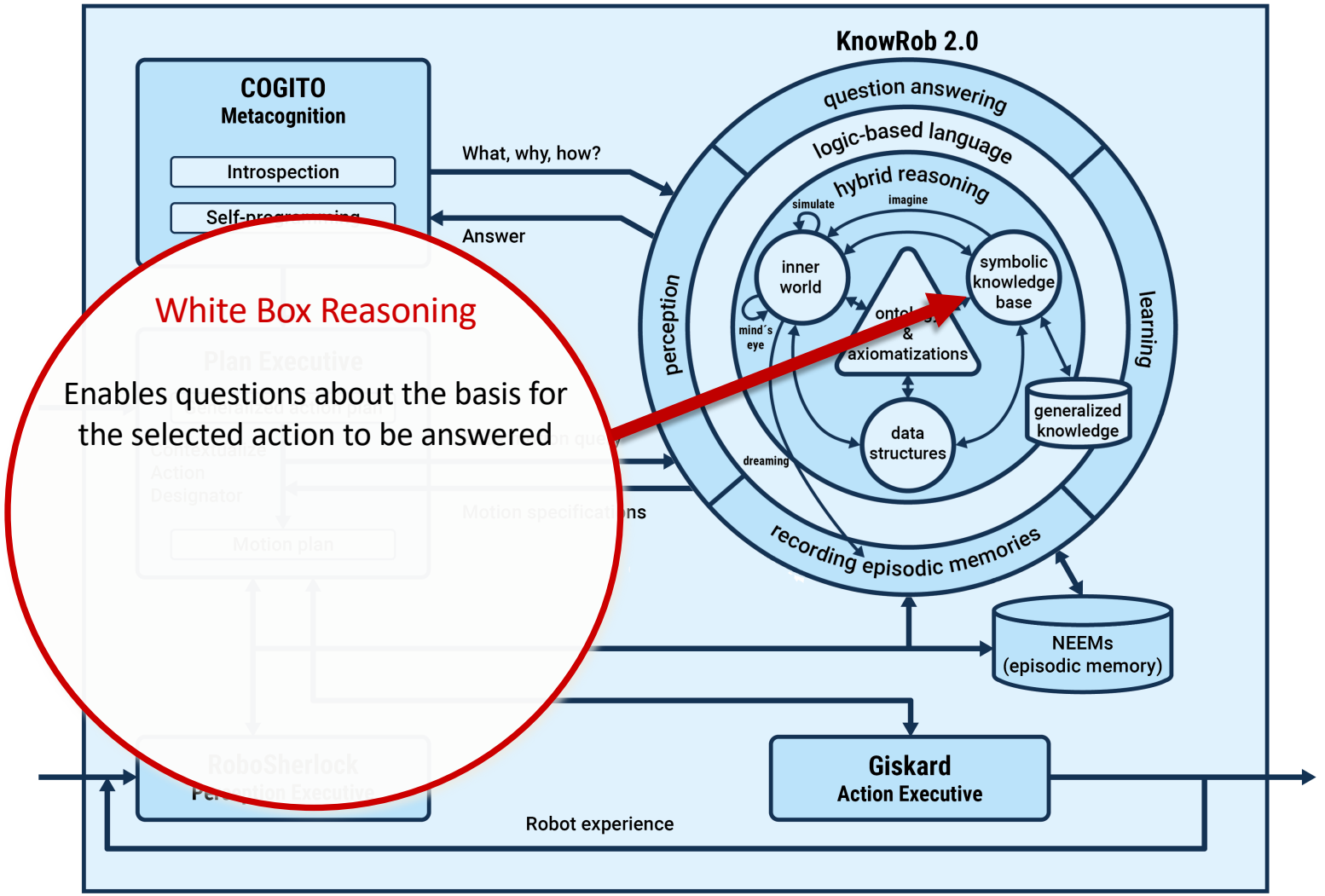


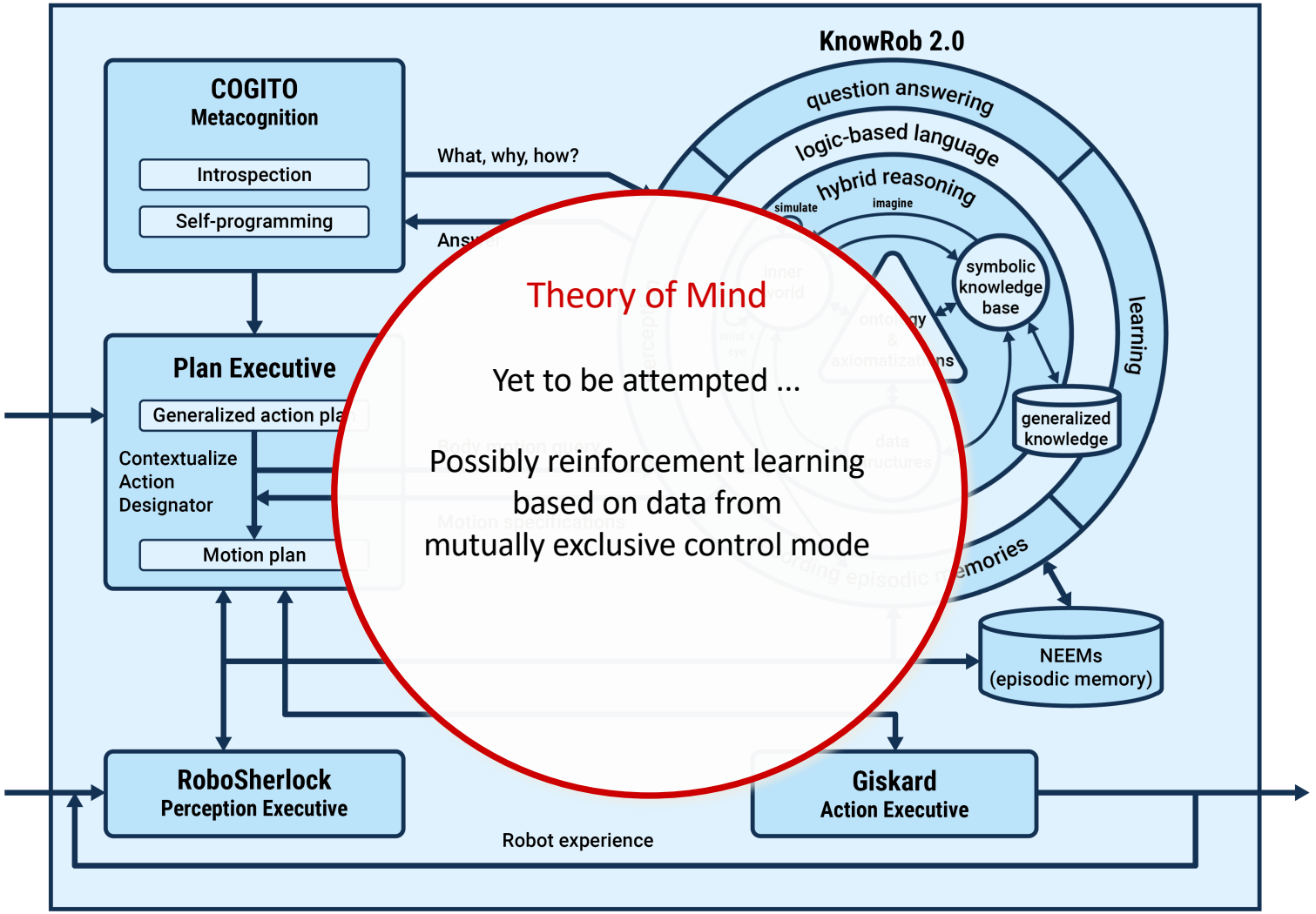
Surgical robot system architecture should be a cognitive architecture
Plan actions
Anticipate outcome before acting

Cognitive architecture must enable white box prospection & reasoning
Human supervisor does not have to form a Theory of Mind
Intentions and expected outcomes are revealed to the human supervisor

Cognitive architecture should ideally have a Theory of Mind
To anticipate interventions by human supervisor
To anticipate human supervisor's information needs


KnowRob 2.0
COGITO
Metacognition
What, why, how?
Answer
Body motion query
Motion specifications
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
ontology & axiomatizations
mind´s eye
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
White Box Prospection
Allows the potential outcome of intended actions to be visualized by the human supervisor
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience


KnowRob 2.0
COGITO
Metacognition
Introspection
What, why, how?
Answer
question answering
logic-based language
hybrid reasoning
simulate
imagine
inner world
symbolic knowledge base
mind´s eye
ontologies & axiomatizations
data structures
generalized knowledge
dreaming
perception
learning
recording episodic memories
NEEMs (episodic memory)
Giskard
Action Executive
Robot experience
White Box Reasoning
Enables questions about the basis for the selected action to be answered


KnowRob 2.0
question answering
logic-based language
hybrid reasoning
simulate
imagine
learning
symbolic knowledge base
generalized knowledge
NEEMs (episodic memory)
COGITO
Metacognition
Introspection
Self-programming
What, why, how?
Plan Executive
Generalized action pla
Contextualize
Action
Designator
Motion plan
Theory of Mind
Yet to be attempted ...
Possibly reinforcement learning based on data from mutually exclusive control mode
RoboSherlock
Perception Executive
Giskard
Action Executive
Robot experience

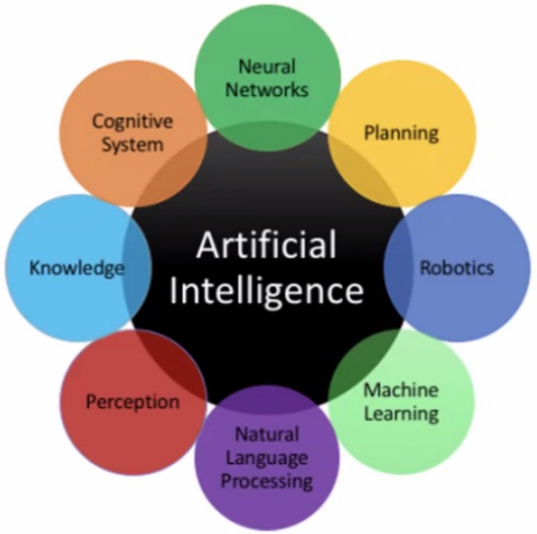
There are Many Types of AI

- Algorithmic
- Classifiers
- AI Symbolic
- AI Sub-symbolic
  - Neural Networks
  - Machine Learning
    - Supervised
    - Unsupervised
- AI Generative
- AI Explainable
- AI Trustworthy
- AI Regulations

Safe

Risky

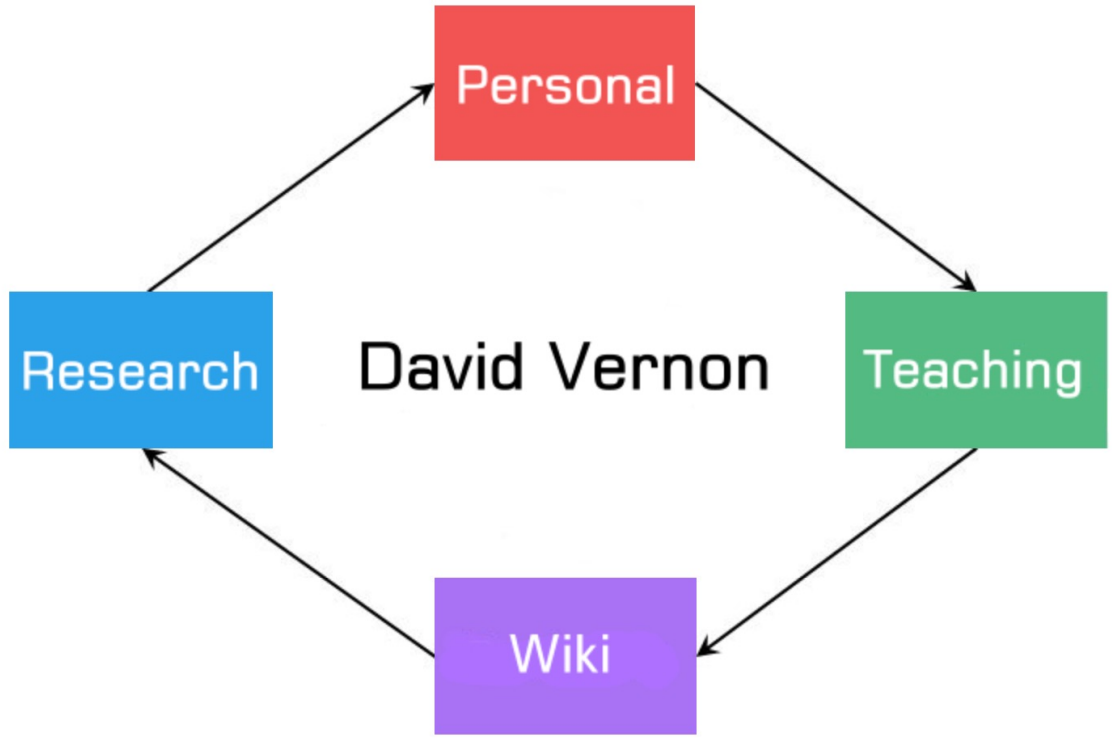
Desirable



P. Fiorini, Fundamentals of Autonomy, ICRA 2024 Workshop, Autonomy in Robotic Surgery: State of the Art, Technical and Regulatory Challenges for Clinical Application, 2024.


20
min


THANK
YOU!



www.vernon.eu/wiki/Talks_and_Presentations

david@vernon.eu
www.vernon.eu

ICRA 2024 Workshop

Autonomy in Robotic Surgery:
State of the Art, Technical and Regulatory Challenges for Clinical Application

13th May 2024

Can Cognitive Architectures Support Autonomy in Robotic Surgery?

David Vernon

Carnegie Mellon University Africa

www.vernon.eu