

Artificial Cognitive Systems

Module 3: Cognitive Architectures

Lecture 4: Example cognitive architectures: CRAM

David Vernon
Carnegie Mellon University Africa

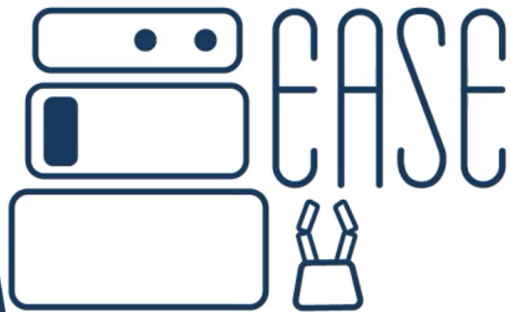
www.vernon.eu

CRAM

- CRAM: Cognitive Robot Abstract Machine
- Hybrid cognitive architecture (symbolic and sub-symbolic representations and processes)
- Introduced by Michael Beetz in 2010
 - developed significantly since then based on several research projects
- Designed to address robot manipulation tasks in everyday activities
 - tasks that would typically be carried out by people in household settings, e.g., in a kitchen.

The Robot Household Marathon aka the EASE Robot Day Demonstrator

Gayane Kazhoyan, Simon Stelter, Ferenc Balint-Benczedi,
Franklin Kenghagho Kenfack, Sebastian Koralewski and Michael Beetz

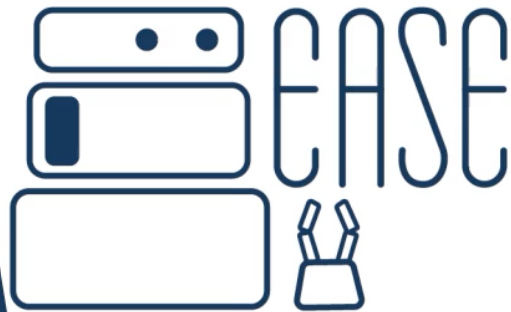


Everyday Activity Science and Engineering
www.ease-crc.org



The Robot Household Marathon aka the EASE Robot Day Demonstrator

Gayane Kazhoyan, Simon Stelter, Ferenc Balint-Benczedi,
Franklin Kenghagho Kenfack, Sebastian Koralewski and Michael Beetz



Design Principles

Implicit-to-explicit manipulation: “fetch the spoon and put it on the table”

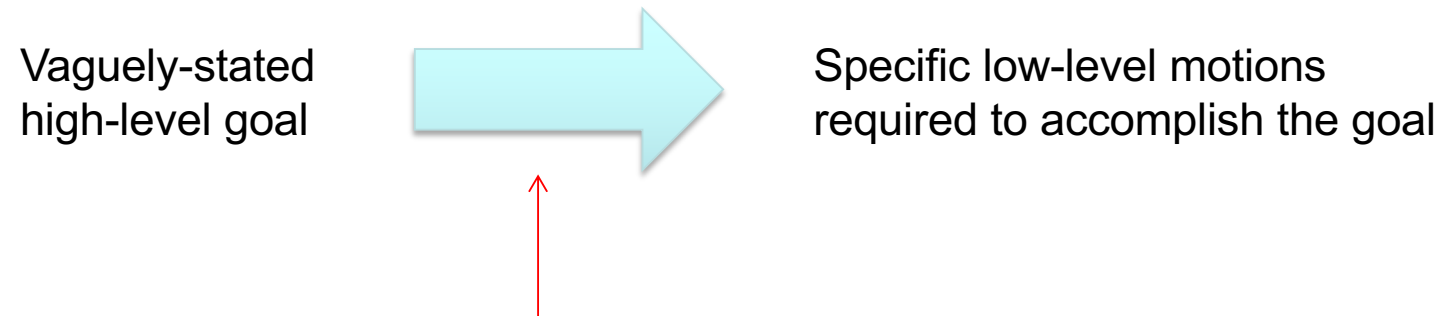
Vaguely-stated
high-level goal



Specific low-level motions
required to accomplish the goal

Design Principles

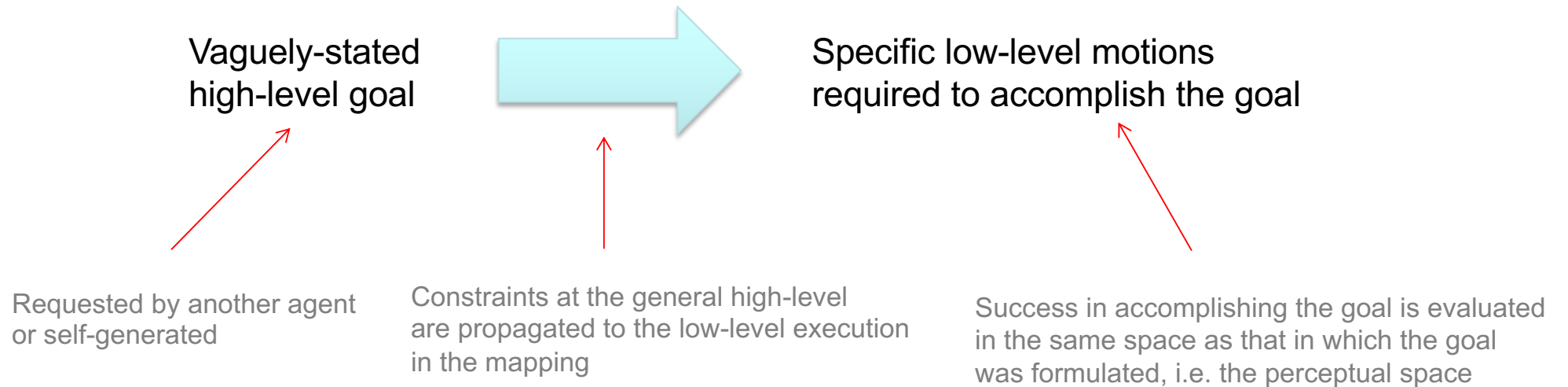
Implicit-to-explicit manipulation: “fetch the spoon and put it on the table”



Mapping is accomplished using a **generative model**:
Tightly-coupled **symbolic** and **sub-symbolic** knowledge representations
of the **robot**, **tasks** it is performing, **objects** it is acting on and **environment** in which it is operating

Design Principles

Implicit-to-**explicit** manipulation: “fetch the spoon and put it on the table”



Design Principles

CRAM focusses on **abstract specification of robot actions** that are **underdetermined**

- The action specifications are framed in terms **without** all knowledge required to complete the action

e.g “fetch the milk and pour it in the bowl”

- The knowledge required to complete the action is **resolved at run-time** during plan execution
- by querying in real-time a multi-element knowledge-base
 - Prior knowledge
 - Current world states
 - Robot’s sensorimotor state

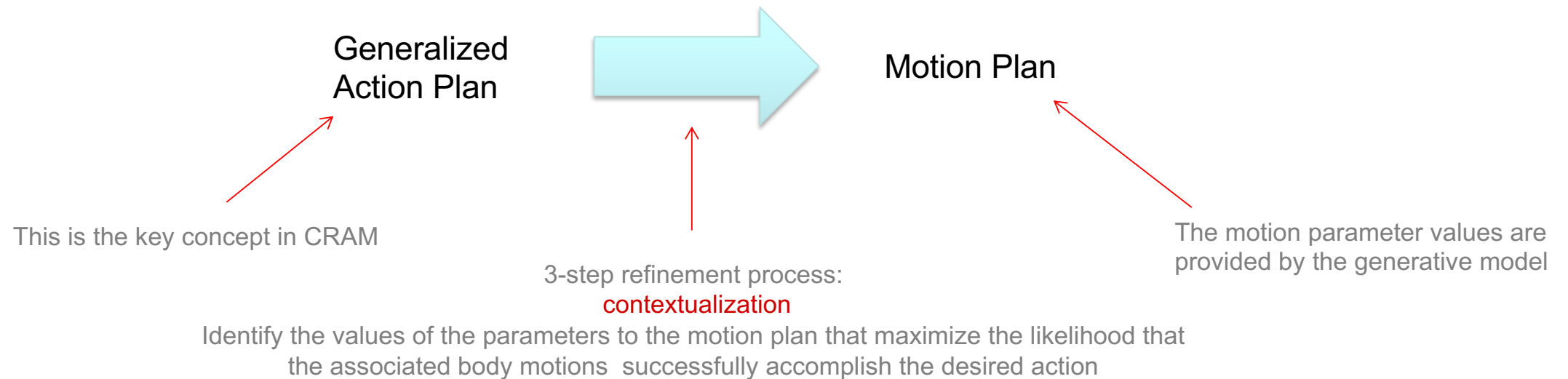
Design Principles

The control program is stated as a **generalized action plan**

- One plan for each category of **underdetermined action description**, e.g. fetch, place, pour, cut, ...
- The plan can be **executed**
- The plan can be **reasoned about** and transformed
 - **Self-programming**
 - Development and self-improvement through automatic generation of new plans

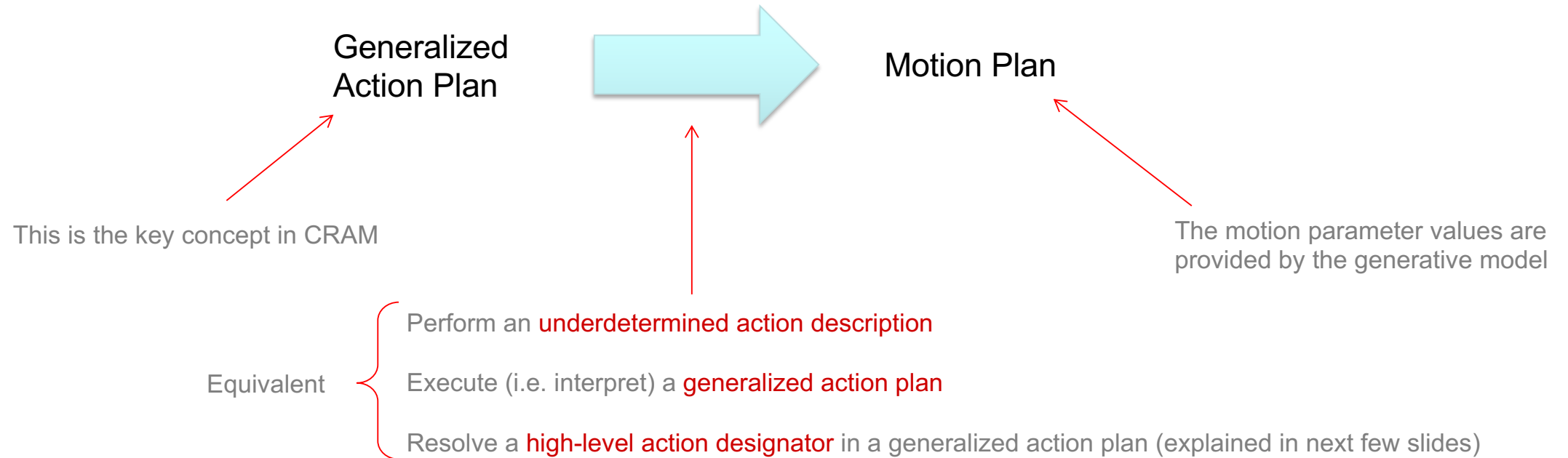
Design Principles

The control program is stated as a **generalized action plan**



Design Principles

The control program is stated as a **generalized action plan**



Control strategies in object manipulation tasks

J Randall Flanagan¹, Miles C Bowman¹ and Roland S Johansson²

The remarkable manipulative skill of the human hand is not the result of rapid sensorimotor processes, nor of fast or powerful effector mechanisms. Rather, the secret lies in the way manual tasks are organized and controlled by the nervous system. At the heart of this organization is prediction. Successful manipulation requires the ability both to predict the motor commands required to grasp, lift, and move objects and to predict the sensory events that arise as a consequence of these commands.

Addresses

¹ Department of Psychology and Centre for Neuroscience Studies, Queen's University, Kingston, ON, K7L 3N6, Canada

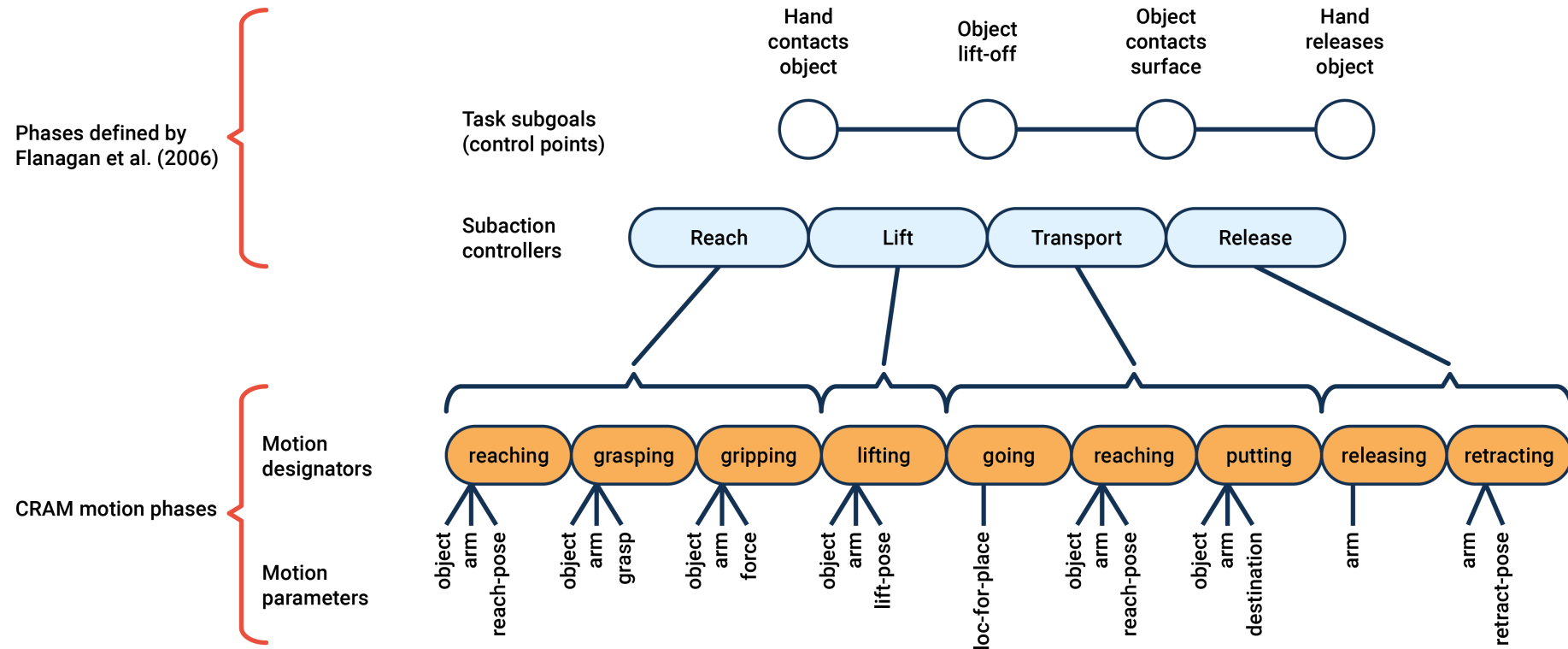
² Section for Physiology, Department of Integrative Medical Biology, Umeå University, SE-901 87 Umeå, Sweden

Corresponding author: Flanagan, J Randall

and another object or surface. Importantly, these contact events give rise to discrete and distinct sensory events, each characterized by a specific afferent neural signature. Because these sensory events provide information related to the functional goals of successive action phases, they have a crucial role in the sensory control of manipulations. In object manipulation, the brain not only forms action plans in terms of series of desired subgoals but also predicts the sensory events that signify subgoal attainment in conjunction with the generation of the motor commands. By comparing predicted sensory events with the actual sensory events, the motor system can monitor task progression and adjust subsequent motor commands if errors are detected. As discussed further below, such adjustments involve parametric adaptation of fingertip actions to the mechanical properties of objects, triggering

J Randall Flanagan, Miles C Bowman, and Roland S Johansson. Control strategies in object manipulation tasks. *Current opinion in neurobiology*, 16(6):650–659, 2006.

Motion Plan

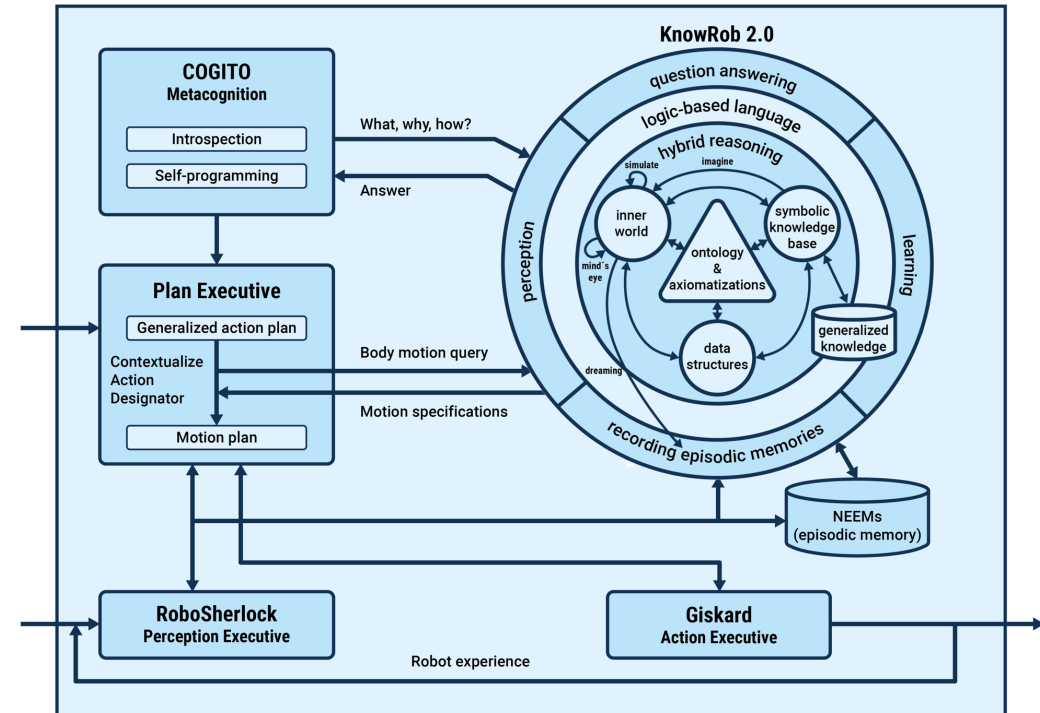


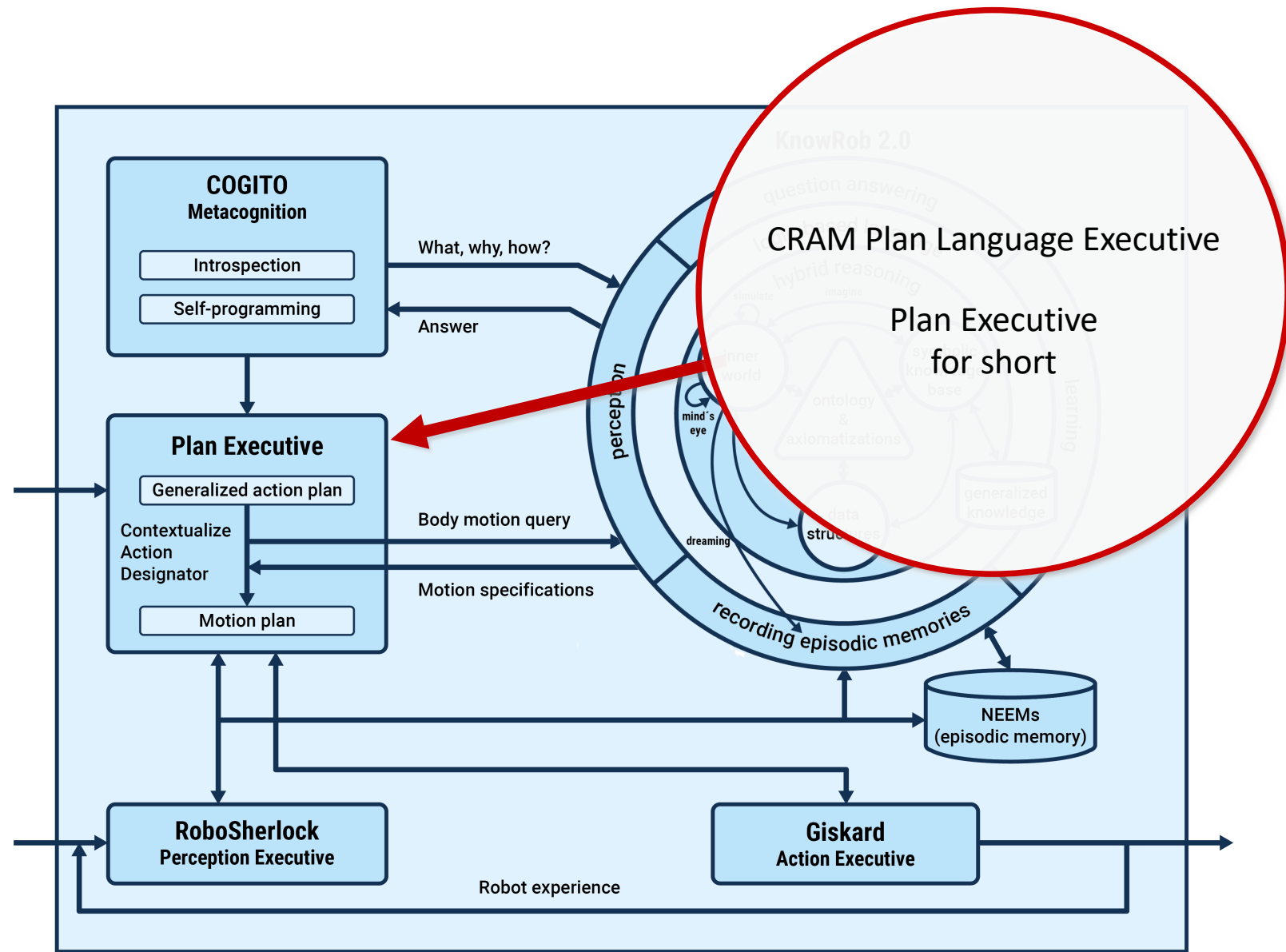
J Randall Flanagan, Miles C Bowman, and Roland S Johansson. Control strategies in object manipulation tasks. *Current opinion in neurobiology*, 16(6):650–659, 2006.

The CRAM 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





CRAM Plan Language (CPL) Executive

- Tasks are accomplished by executing plans written in the CRAM Plan Language (CPL)
- CPL is an extension of Lisp
- A CPL plan represents all key aspects of the plan as persistent **first-class objects** in a **first-order logic**
 - **Plans** themselves **can be reasoned about**, even at runtime
 - Particularly relevant for the meta-cognition system, COGITO


CRAM Plan Language (CPL) Executive

- Plans specify how the robot should respond to
 - Sensory events
 - Changes in belief states
 - Failures in plans
- All of which can be queried, inspected, and reasoned about

CRAM Plan Language (CPL) Executive

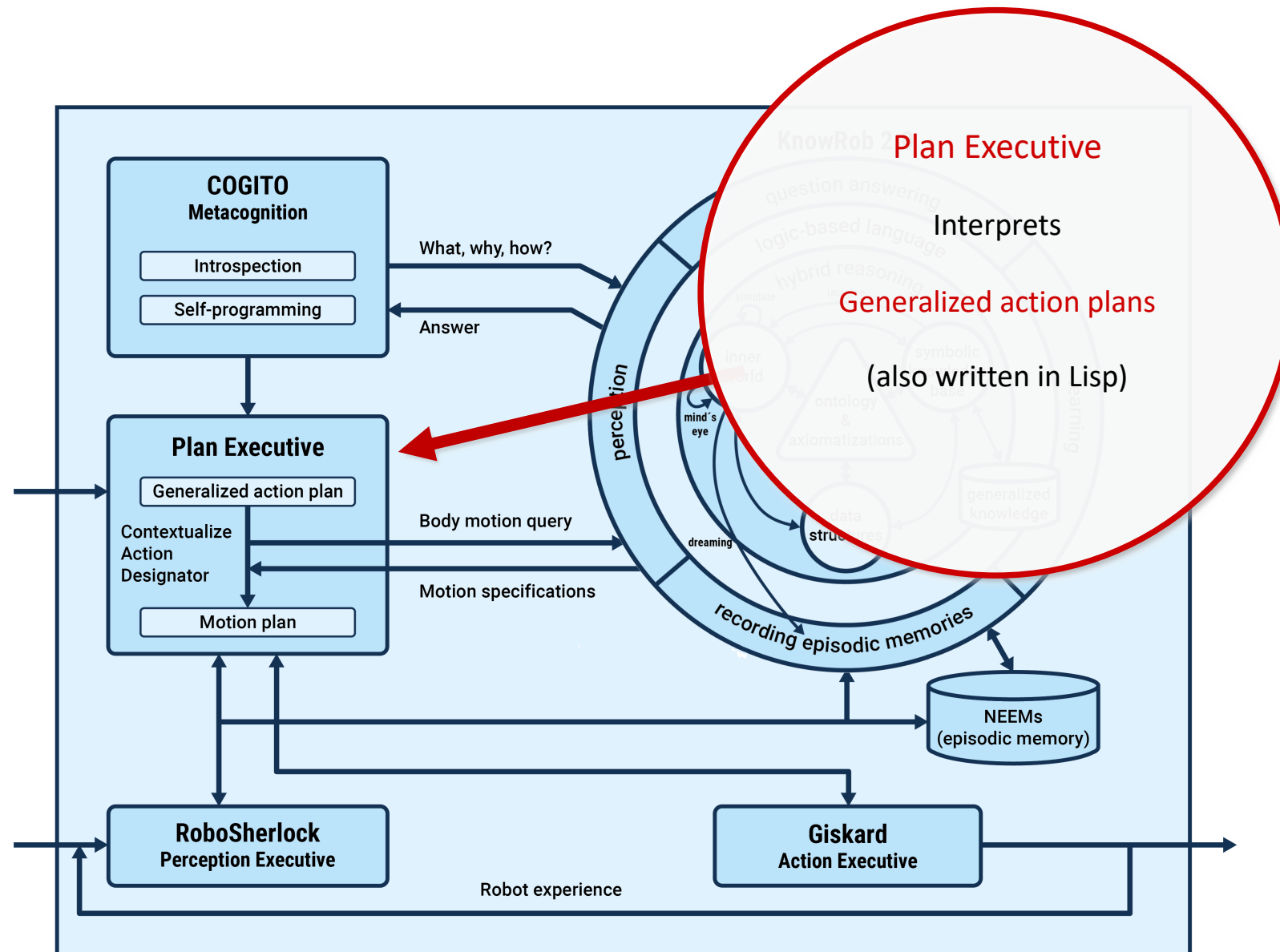
- A plan comprises set of **abstract plan designators** for

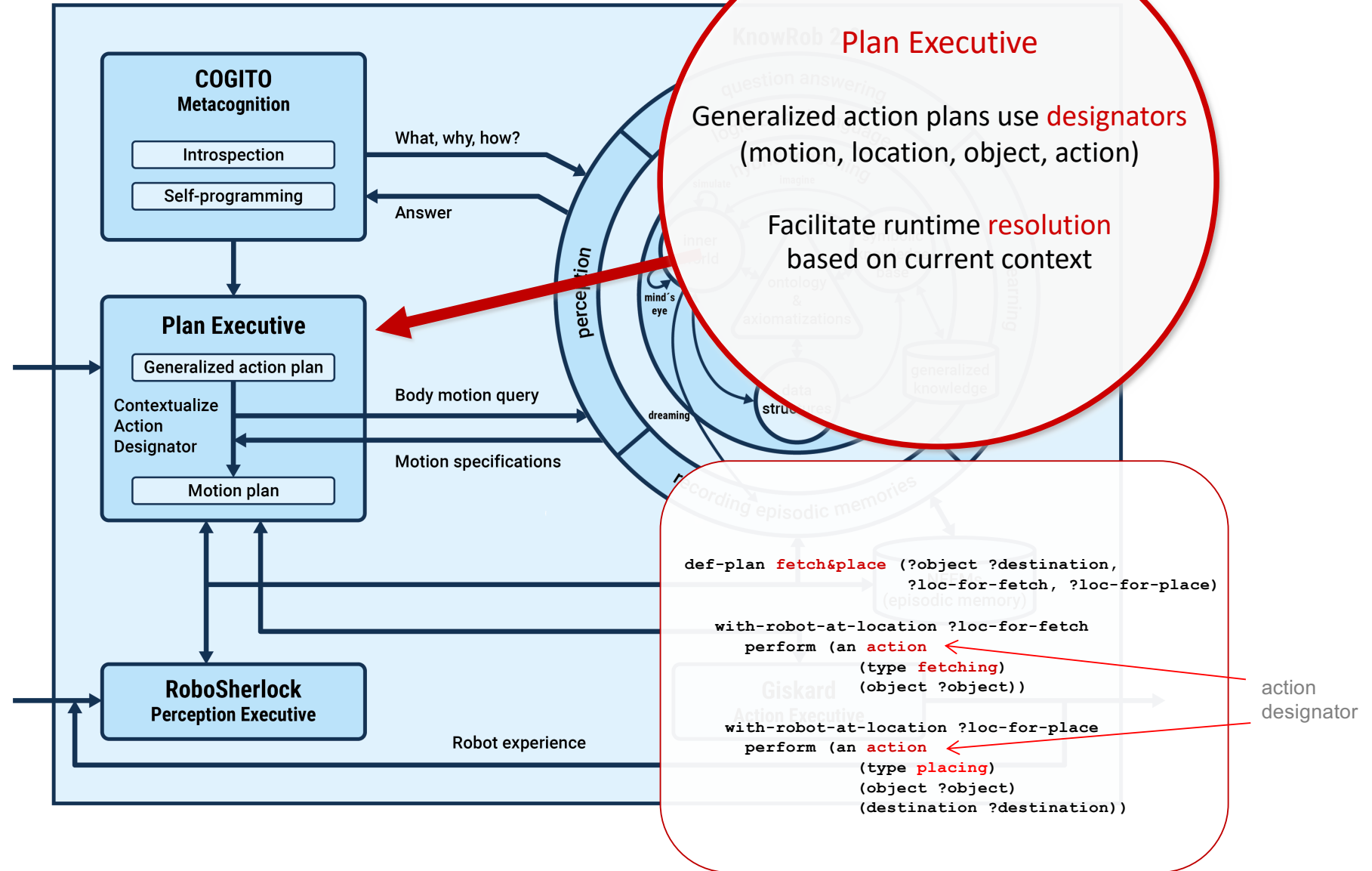
- actions
- objects
- locations
- motions (i.e. elementary movements)

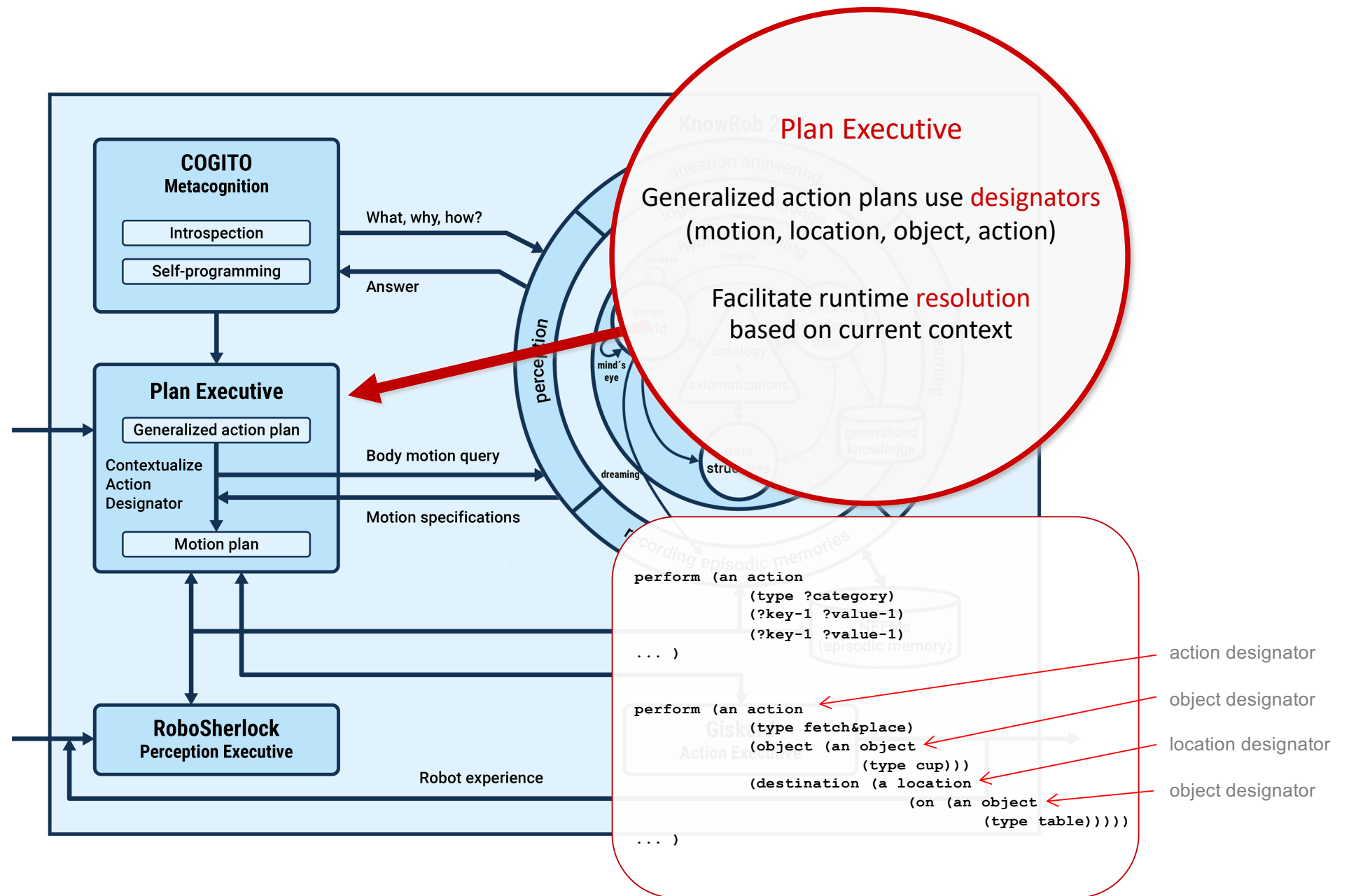


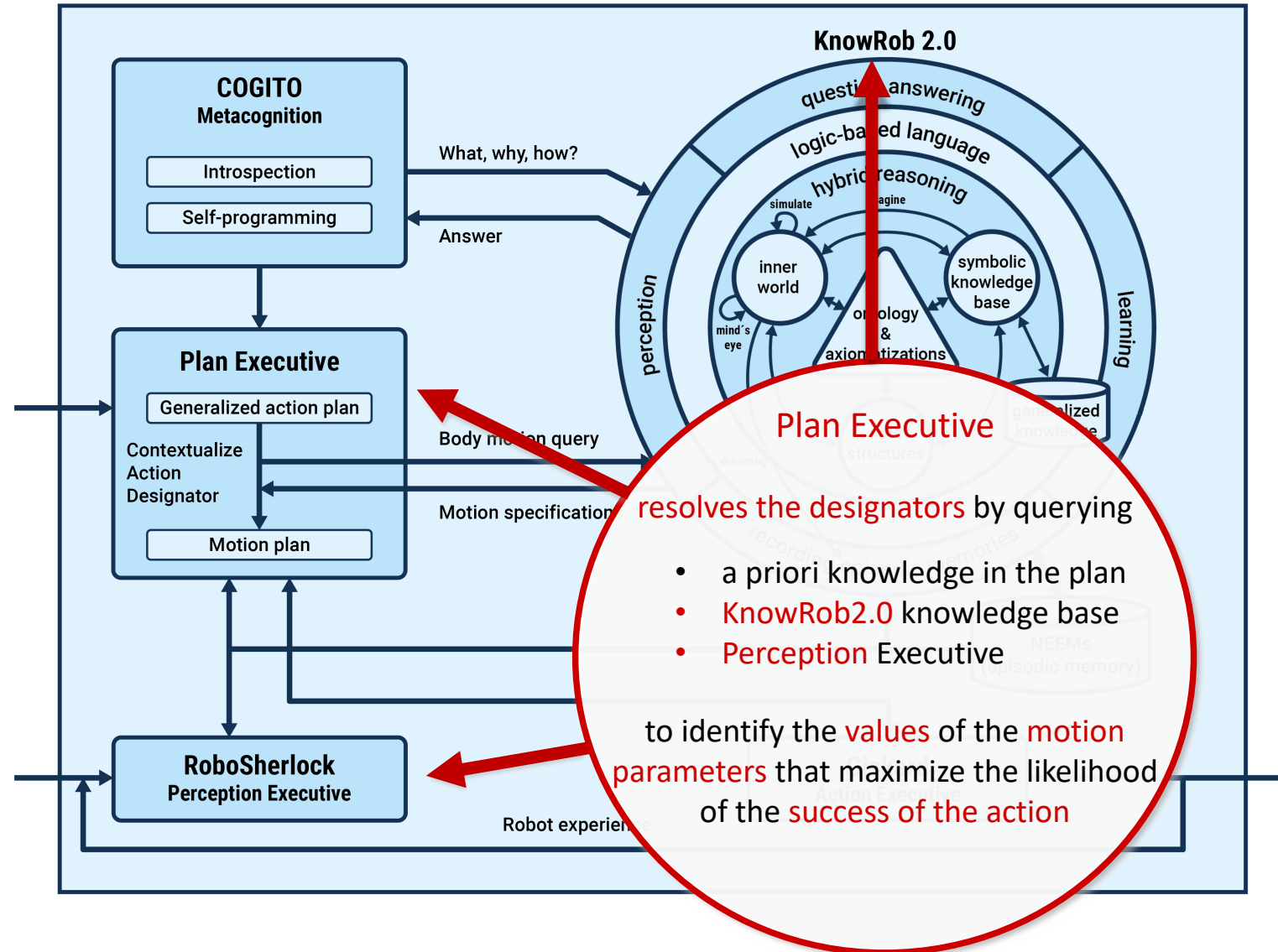
Designators are effectively
placeholders

require runtime **resolution** based on
the current **context** of the task action



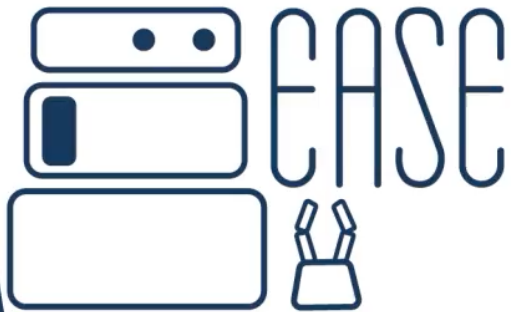


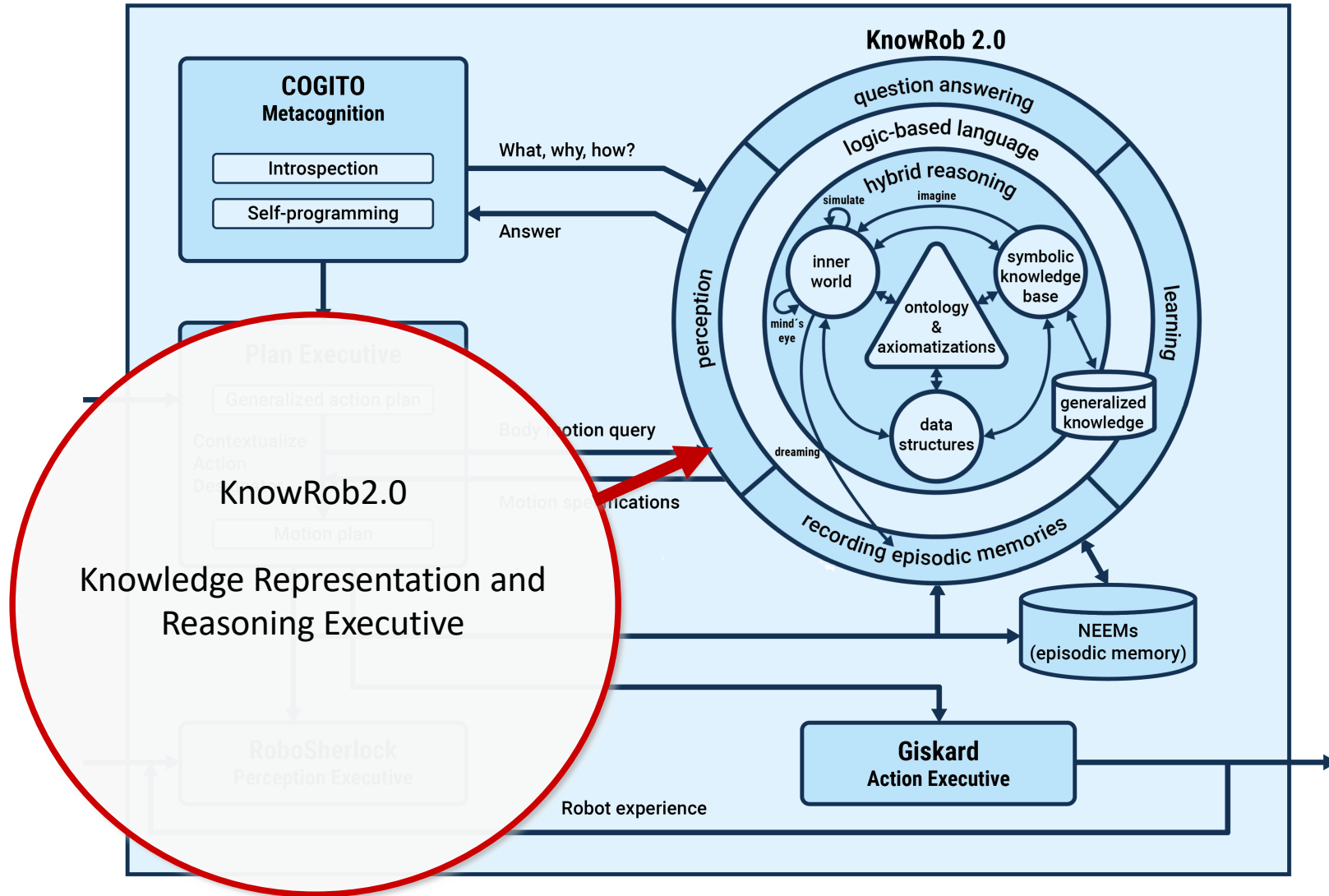


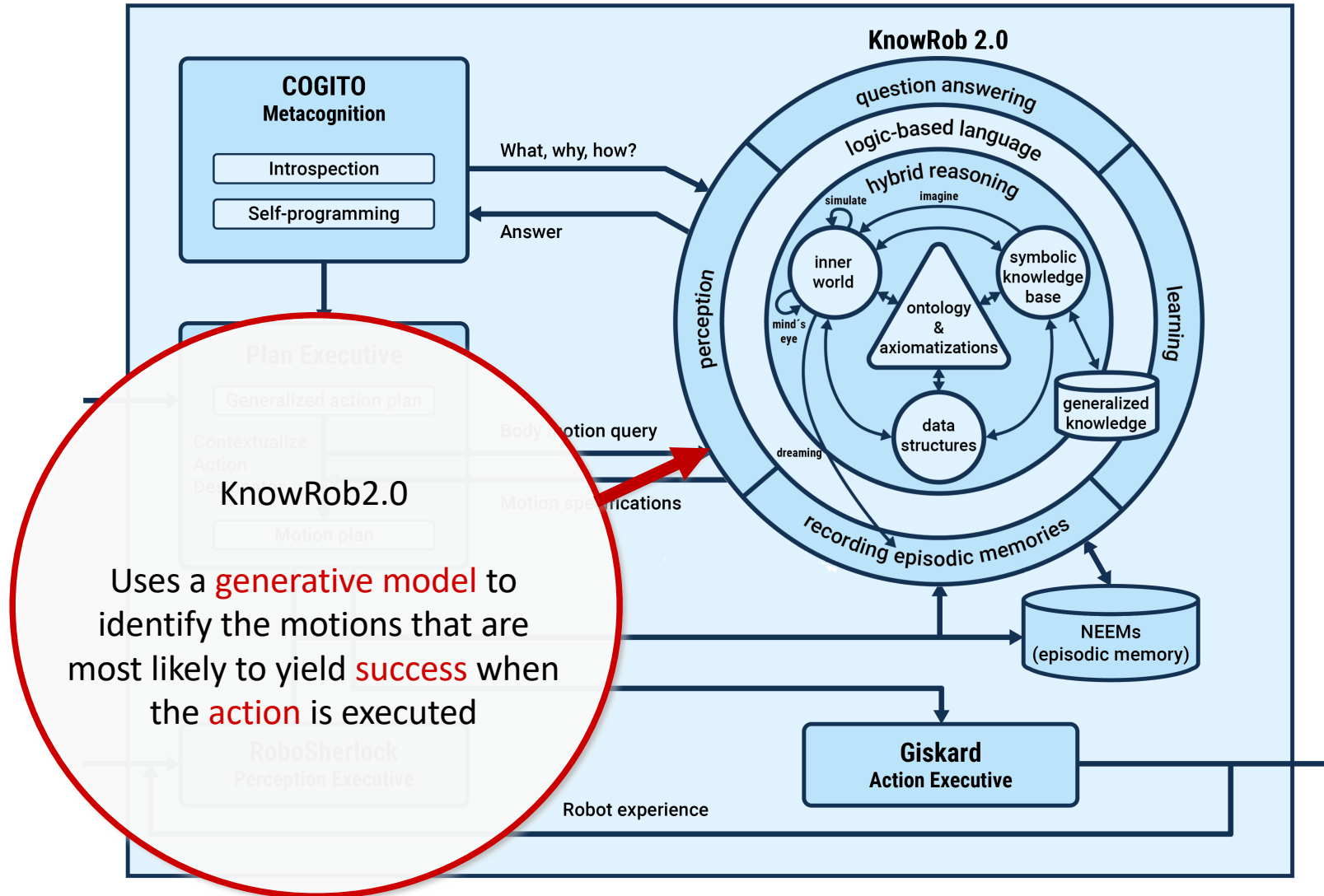


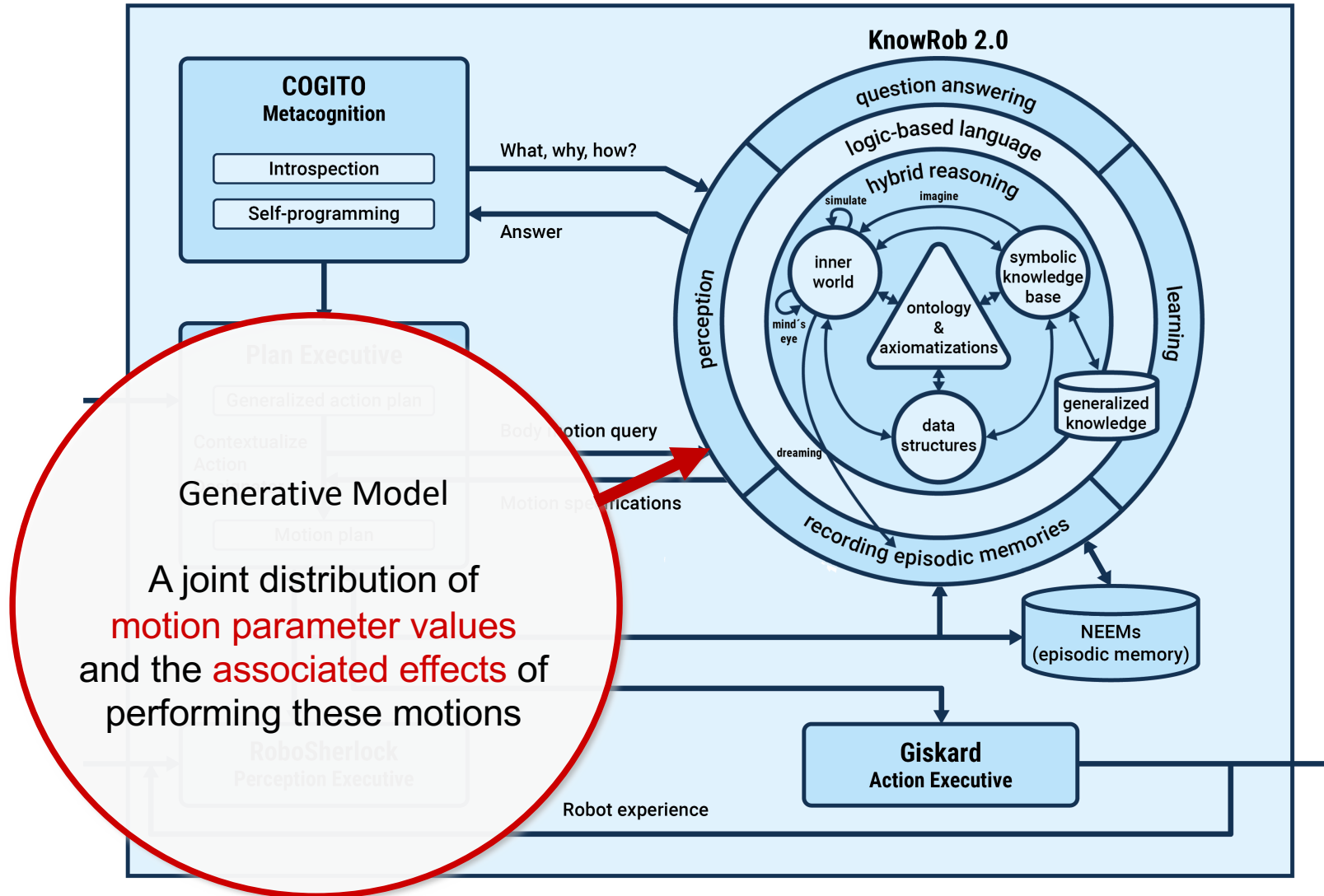
Programming Robotic Agents with Action Descriptions

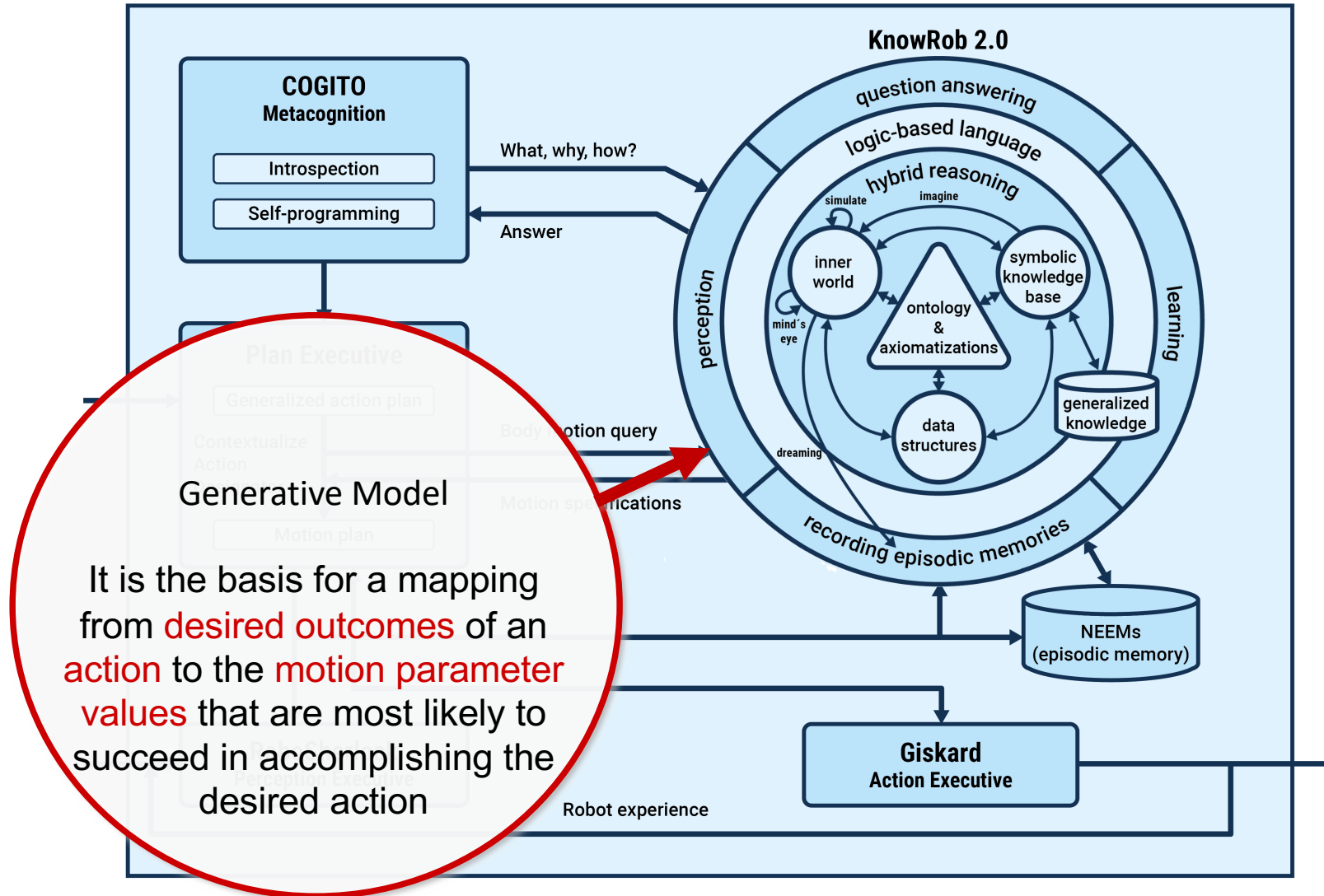
Gayane Kazhoyan and Michael Beetz











It is the basis for a mapping from **desired outcomes** of an **action** to the **motion parameter values** that are most likely to succeed in accomplishing the desired action

KnowRob2 Knowledge Base

- Knowledge representation and reasoning framework for robotic agents
 - Implemented in Prolog
- Exposed as a conventional first-order time interval logic knowledge-base
- However, many logic expressions are constructed on-demand from sensorimotor data computed in real-time

KnowRob2 Knowledge Base

- Provides the **background commonsense intuitive-physics knowledge** required by the CPL executive to implement its goal-directed under-determined task plans, e.g.
 - How to grasp an object (depending on the object's shape, weight, softness, and other properties)
 - How it has to be held while moving it (e.g. upright to avoid spilling its contents)
 - Where the object is normally located.

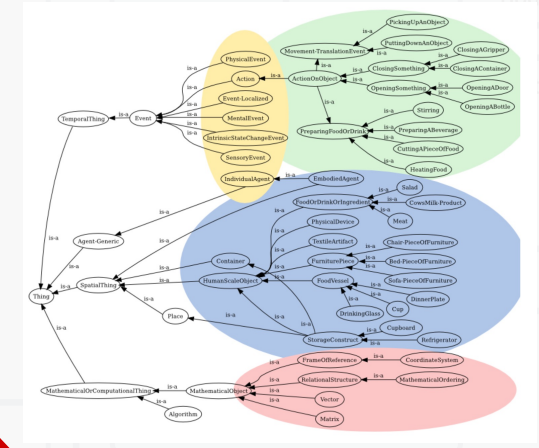
KnowRob2 Knowledge Base

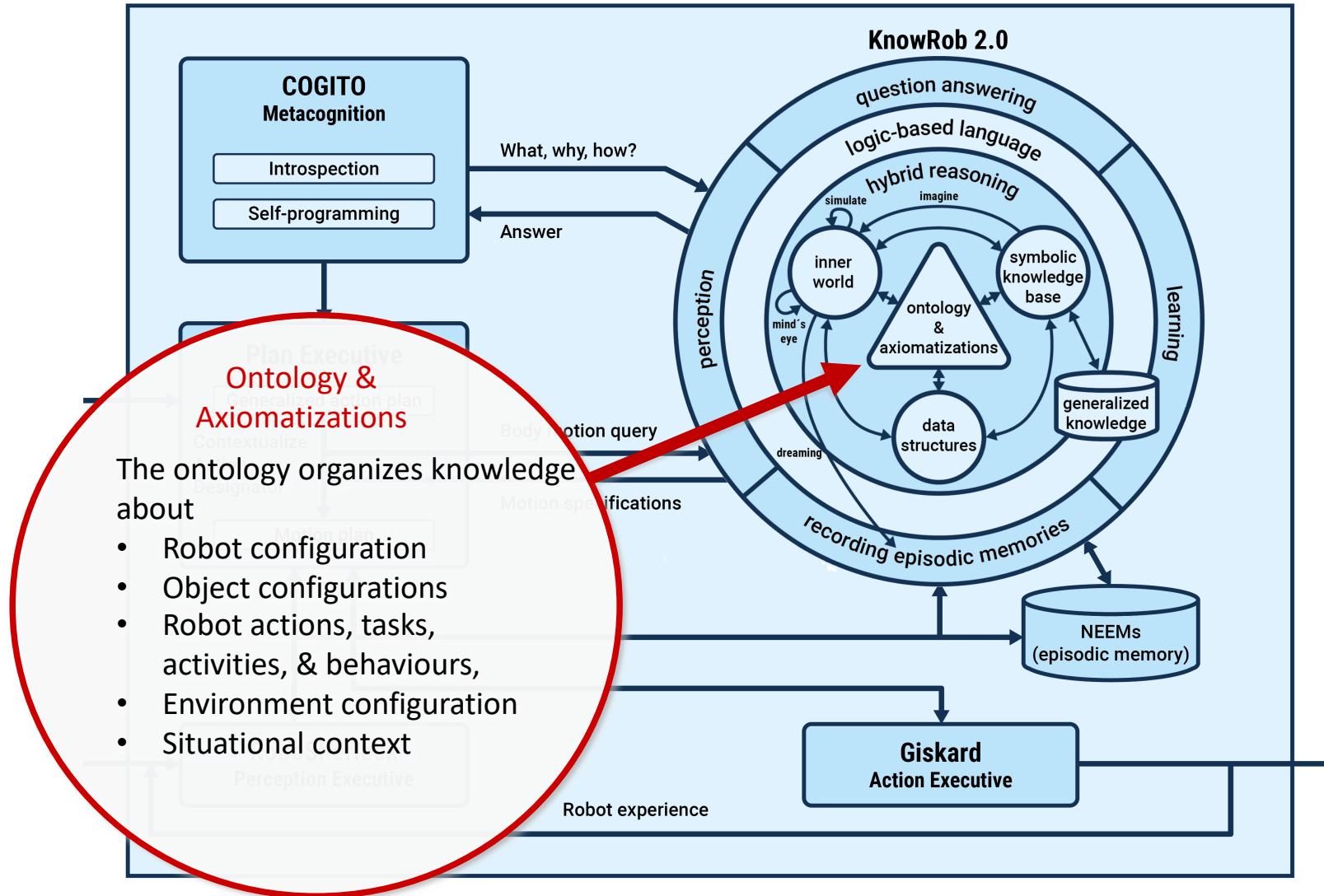
- Source of knowledge:
 - Some is specified **a priori**
 - Some is derived from **experience**
 - Some is the result of **simulated execution of candidate actions** using a high-fidelity virtual reality physics engine simulator
- All represented by a first-order time interval logic expression, **and reasoned about** as needed.

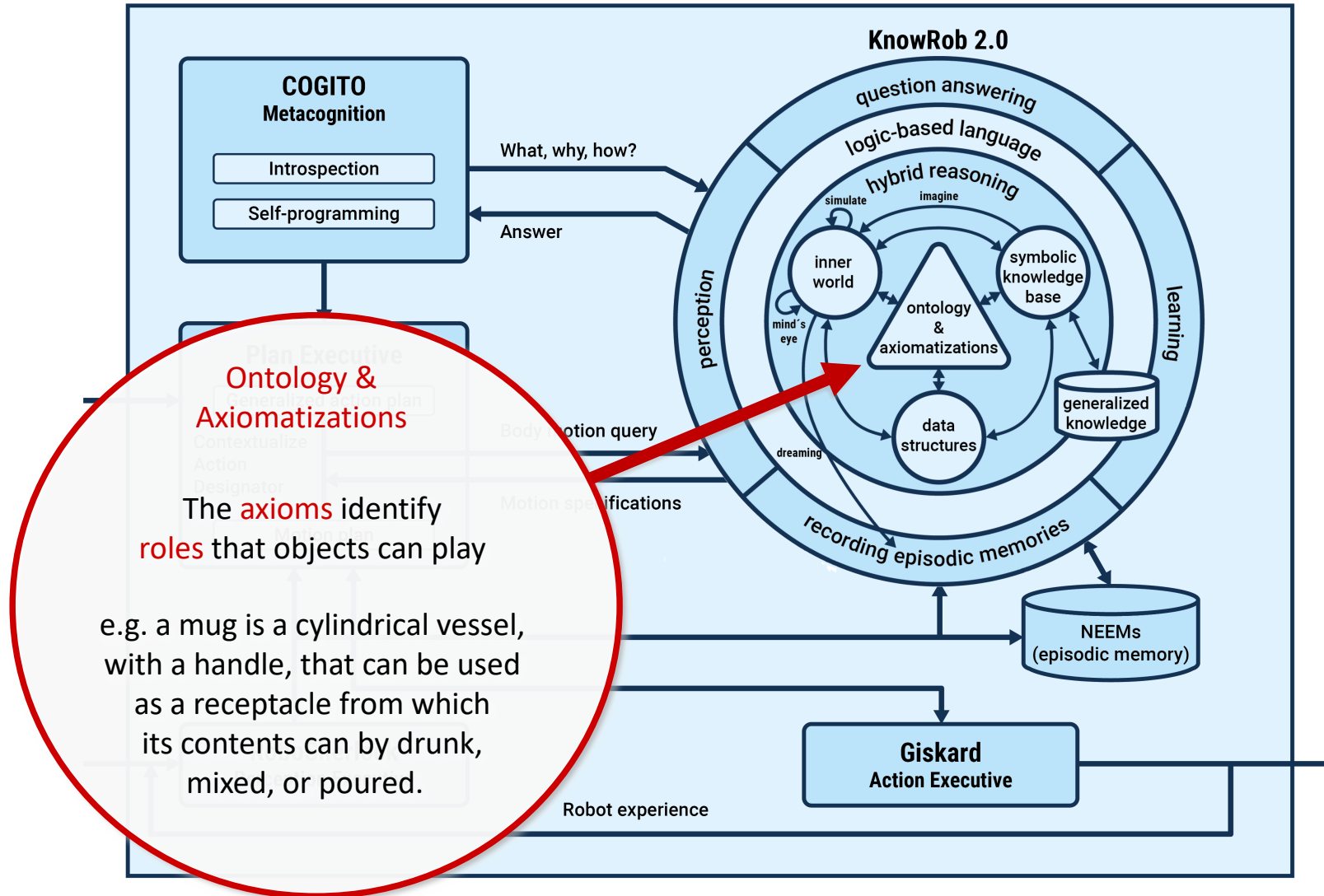
KnowRob2 Knowledge Base

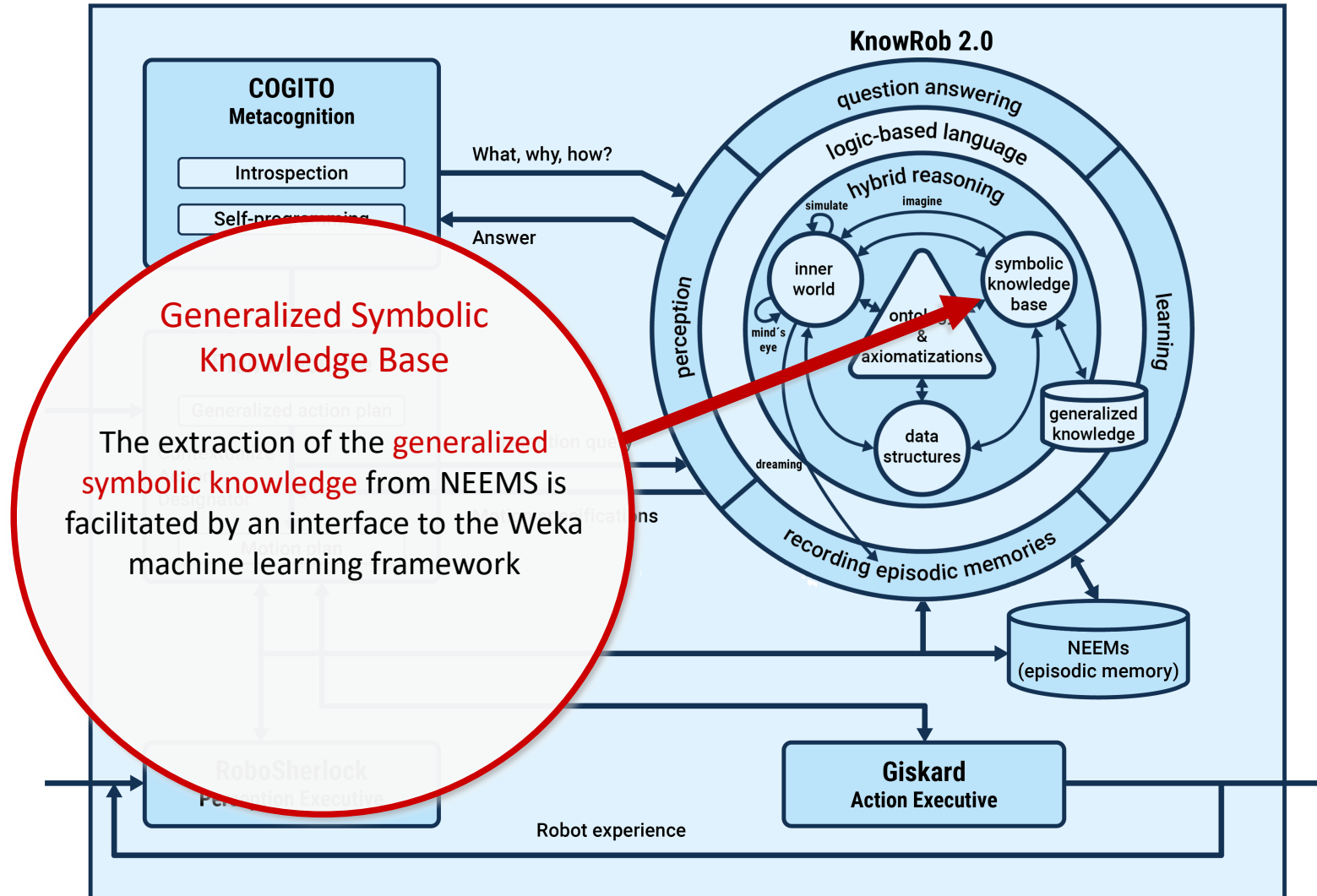
KnowRob2 comprises **five core elements**

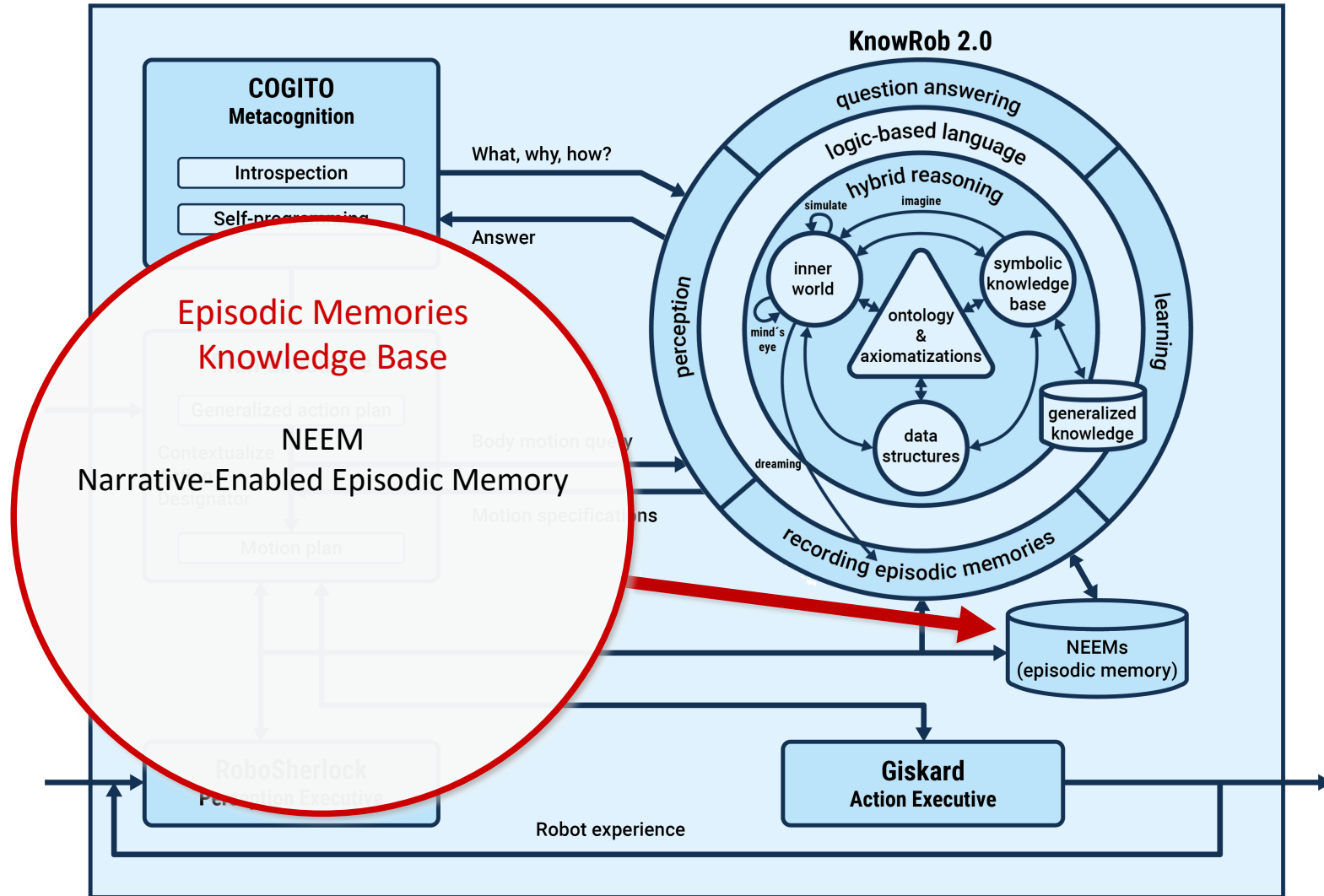
- embedded in a **hybrid** multi-formalism **reasoning** shell
- exposed through a **logic-based language** layer

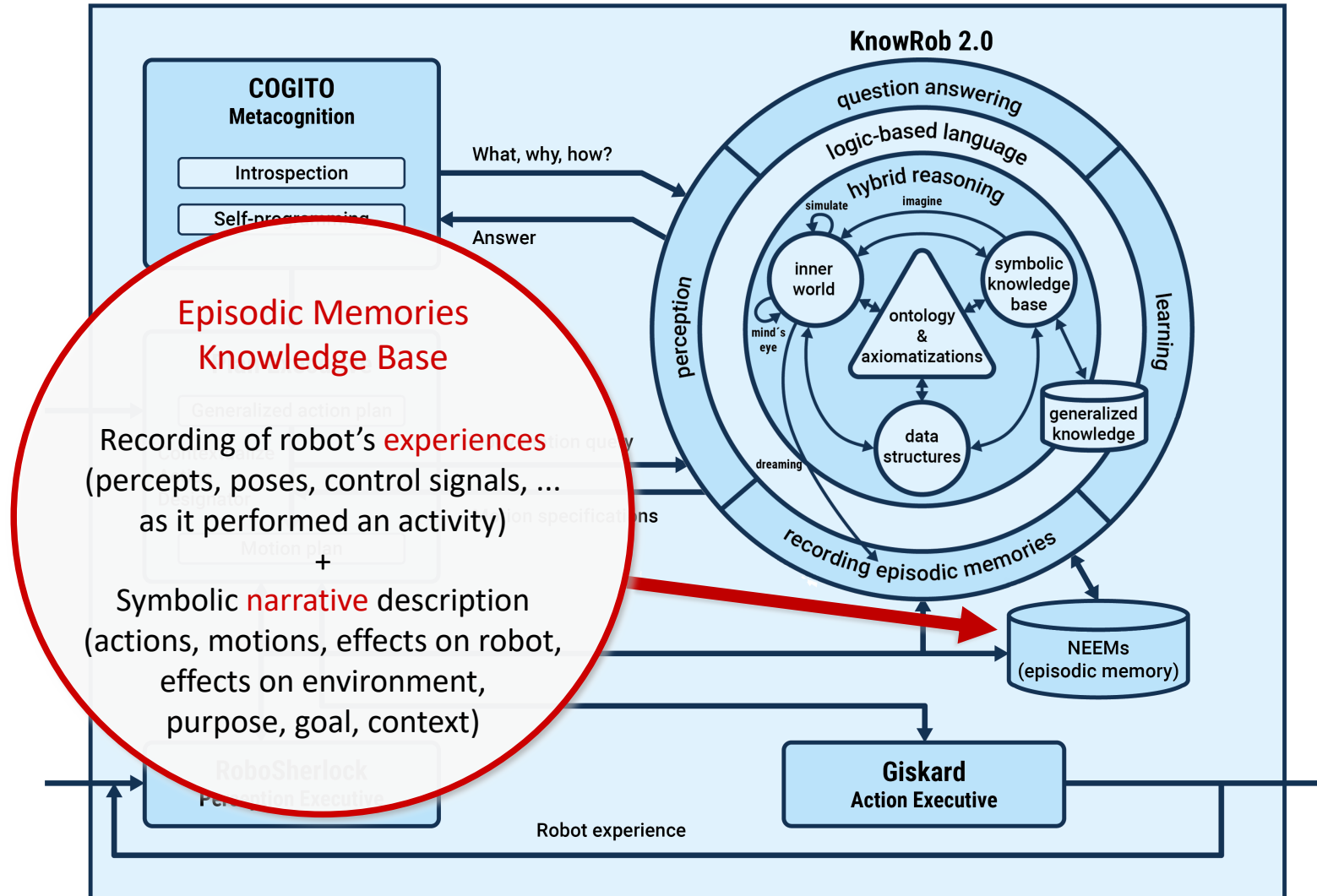


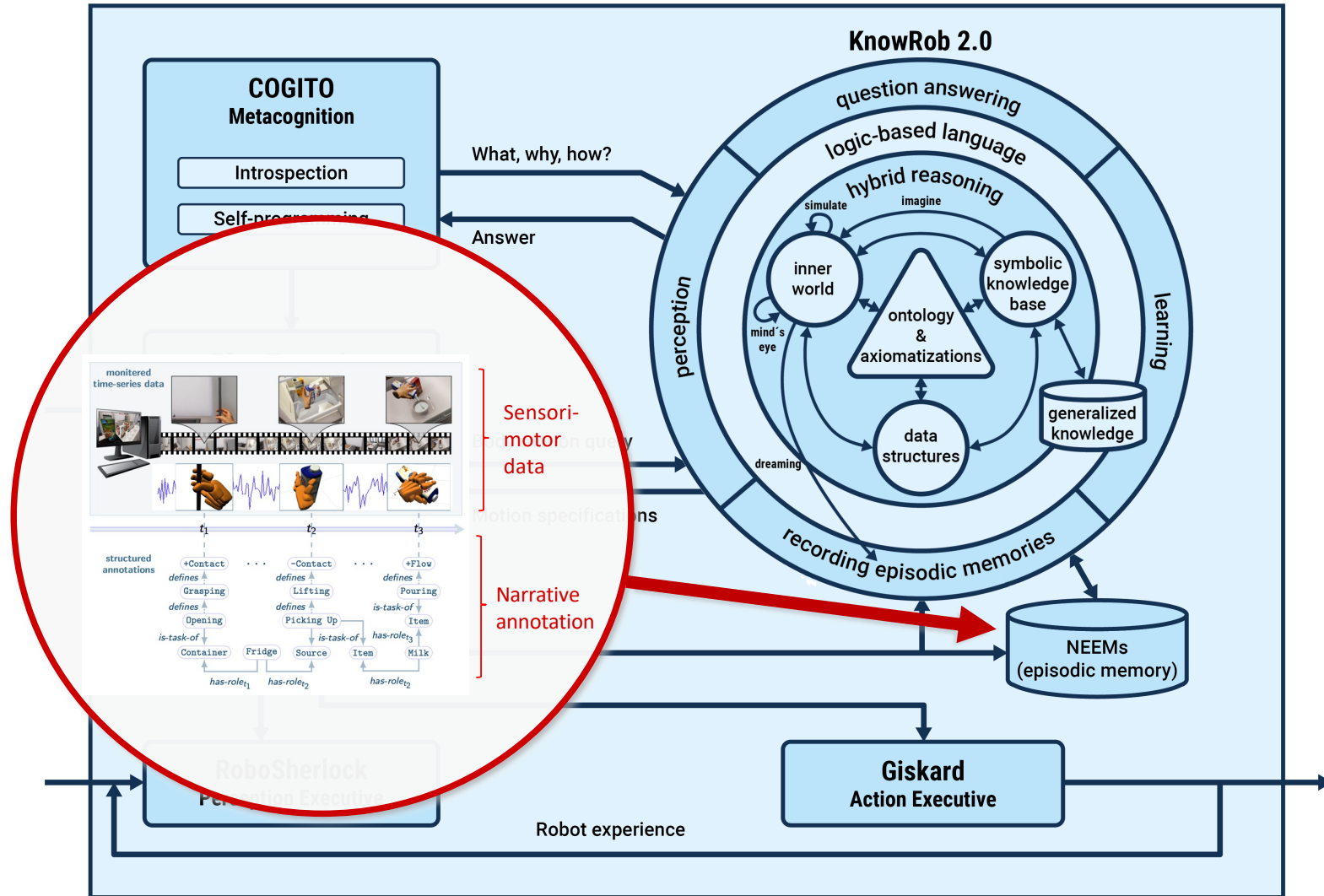


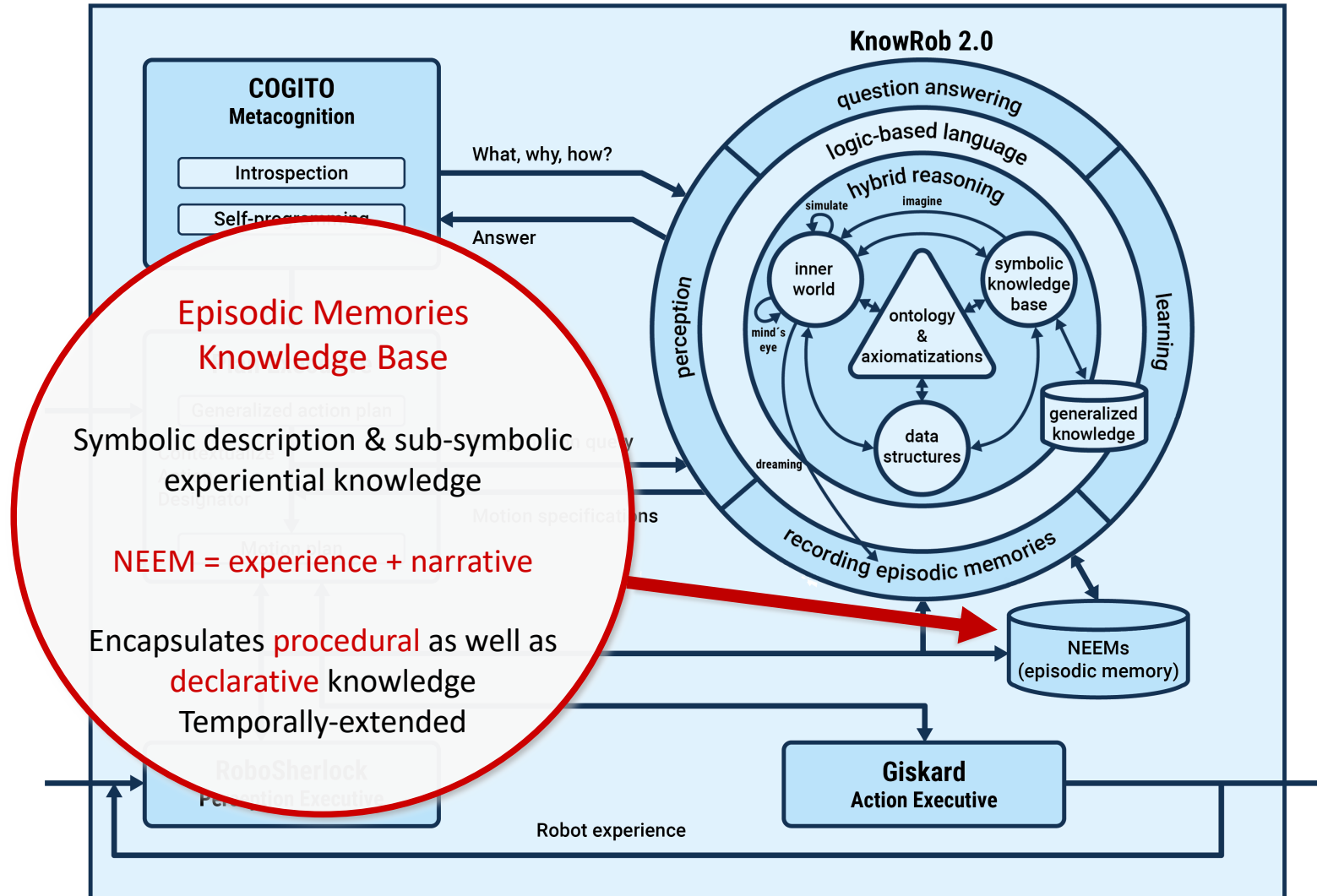


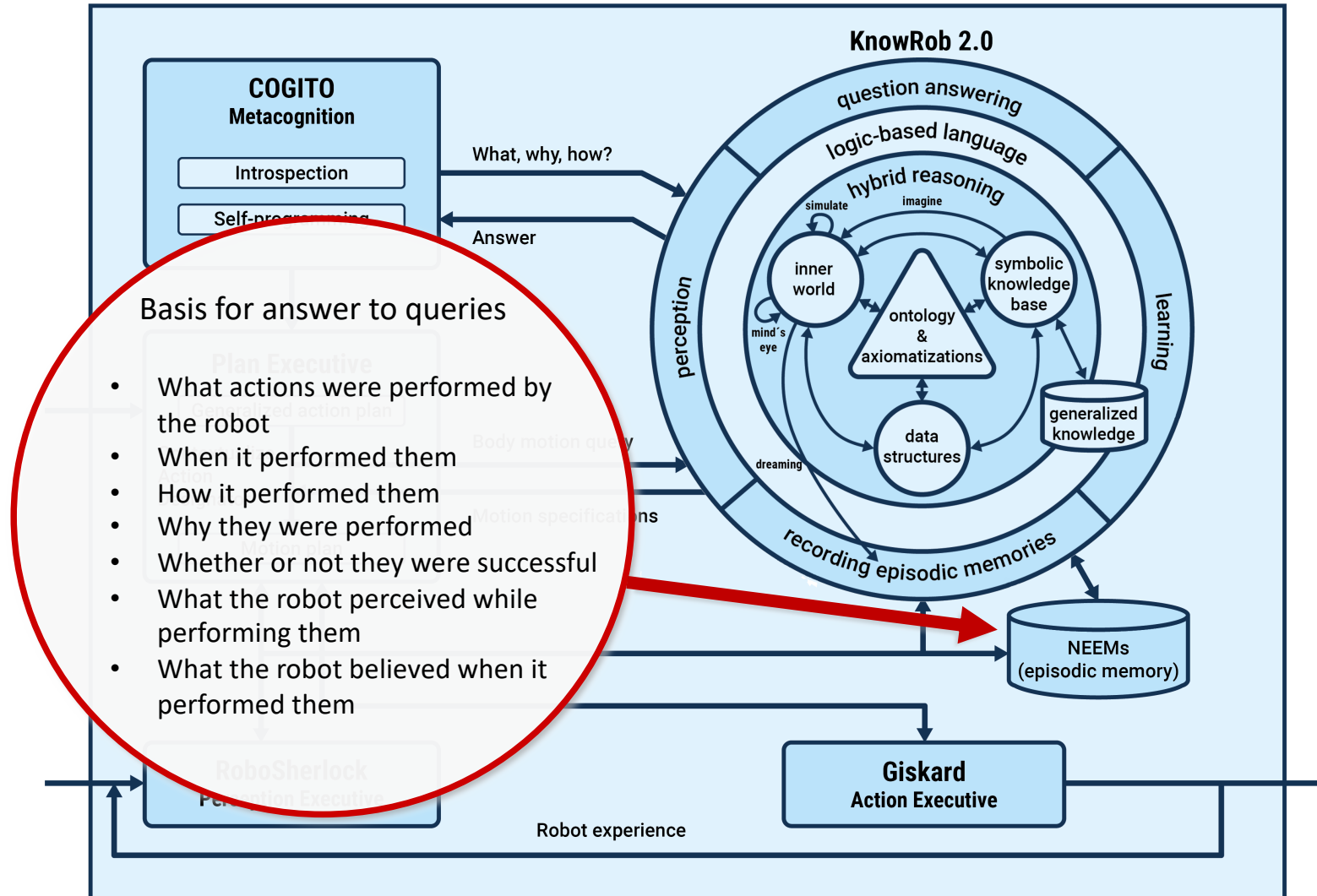


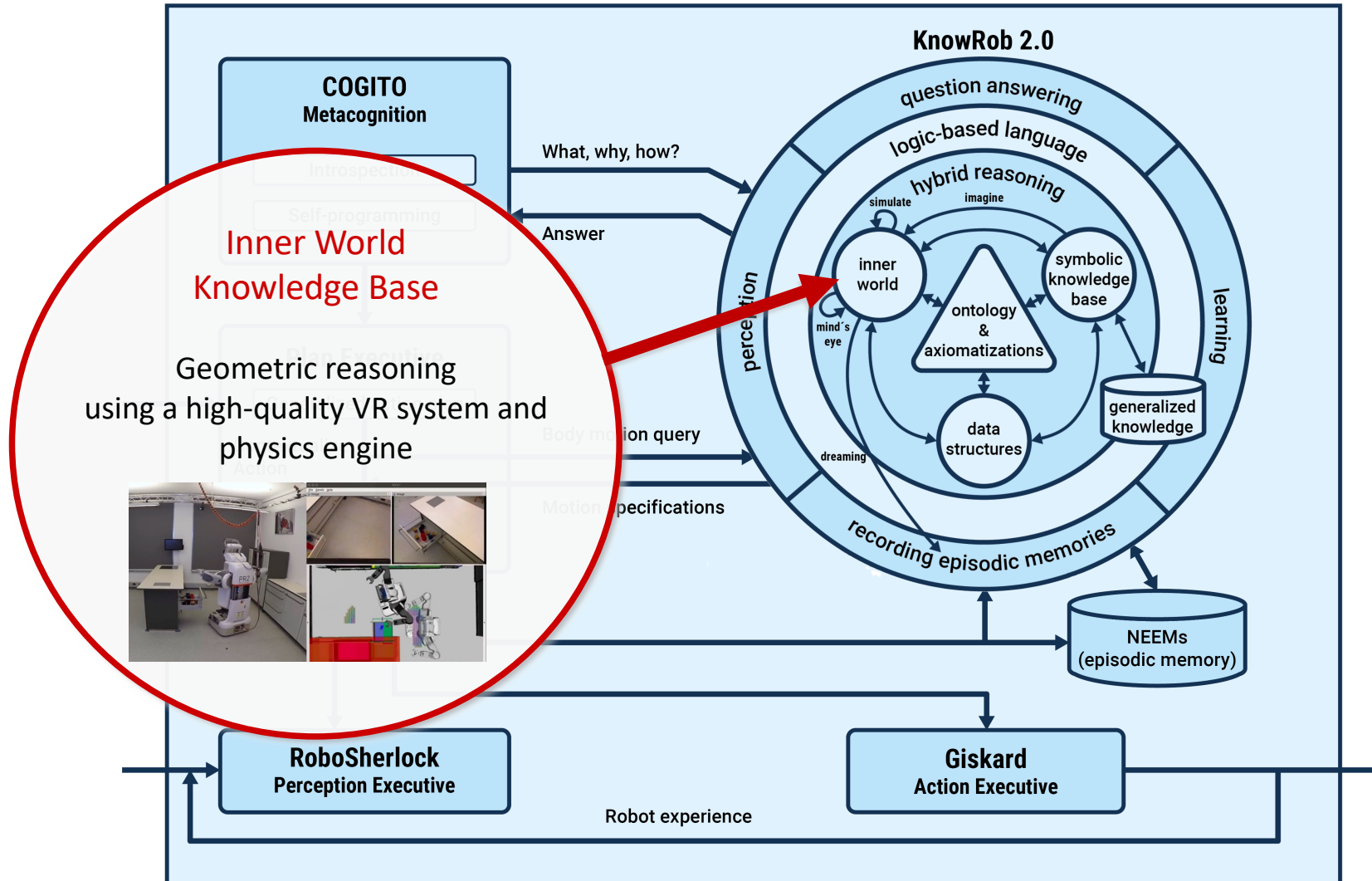


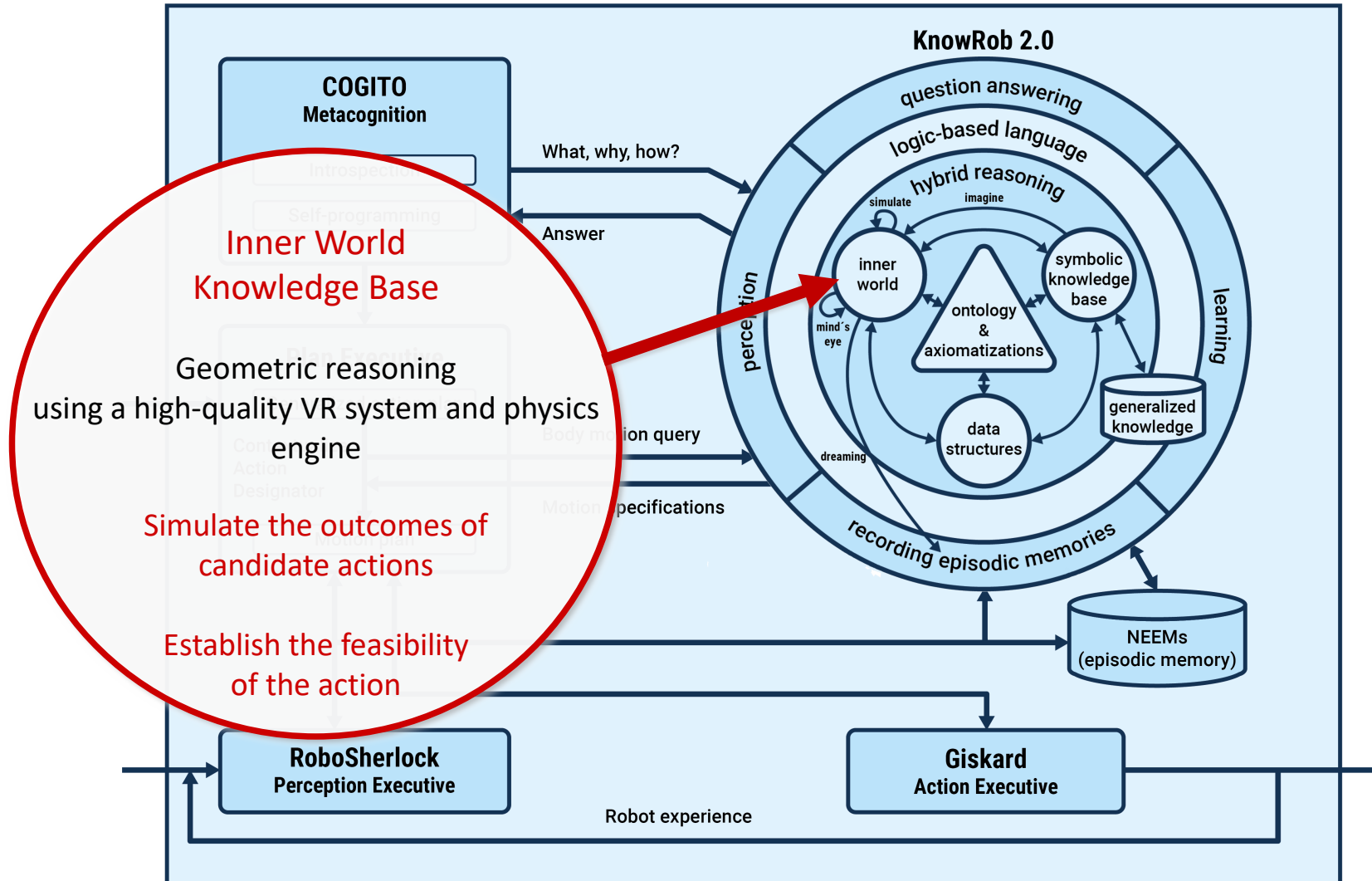


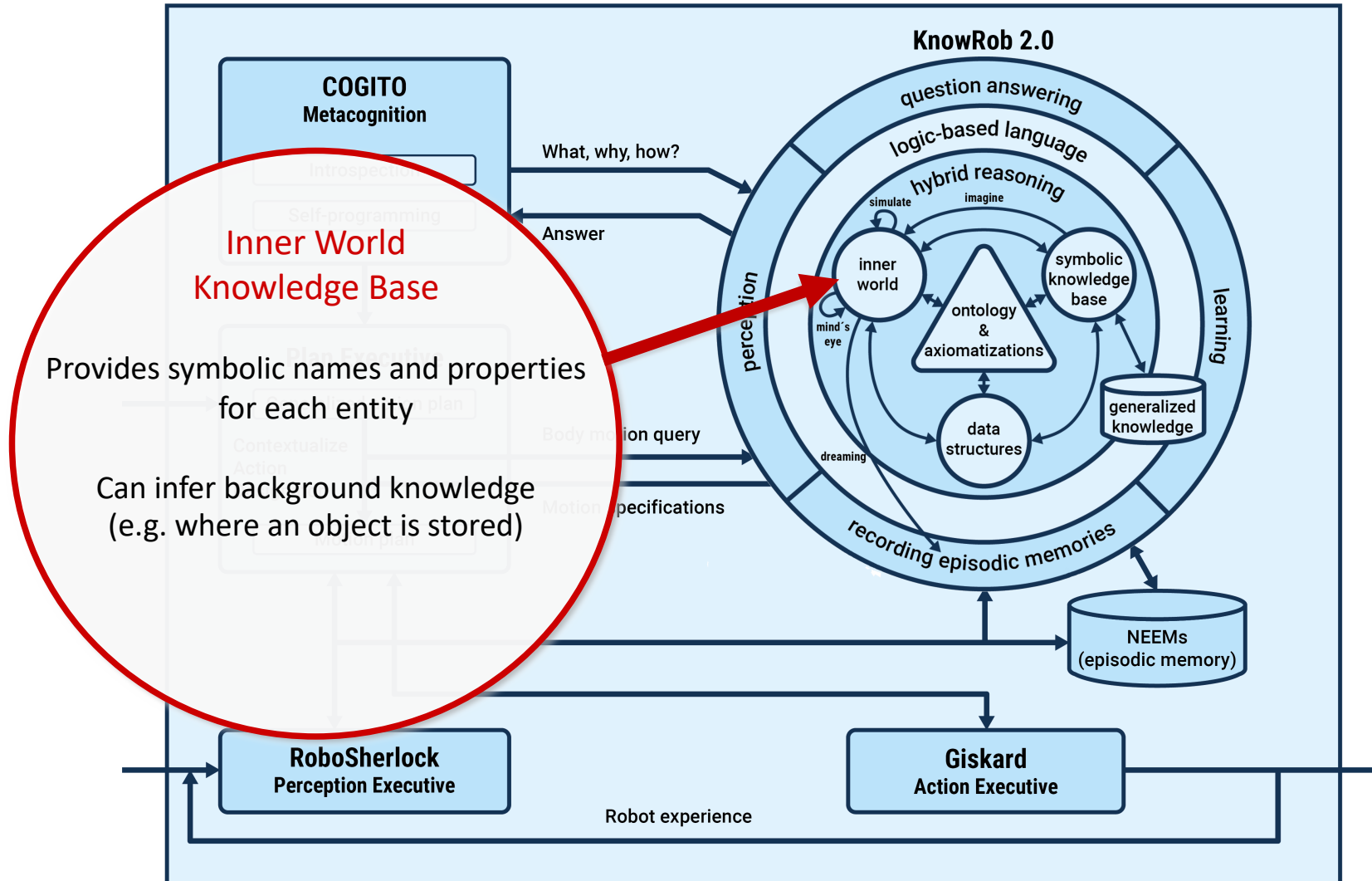


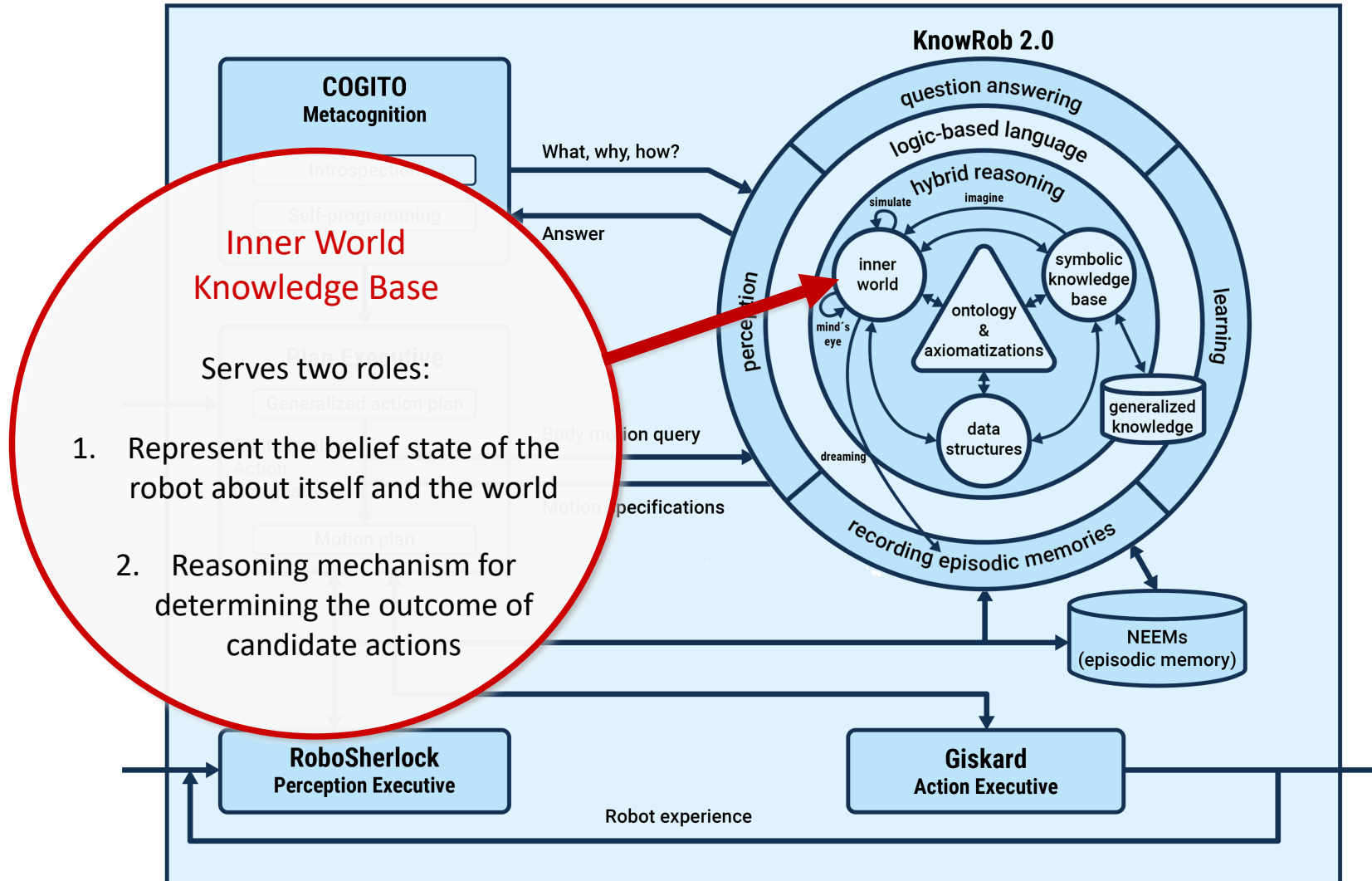


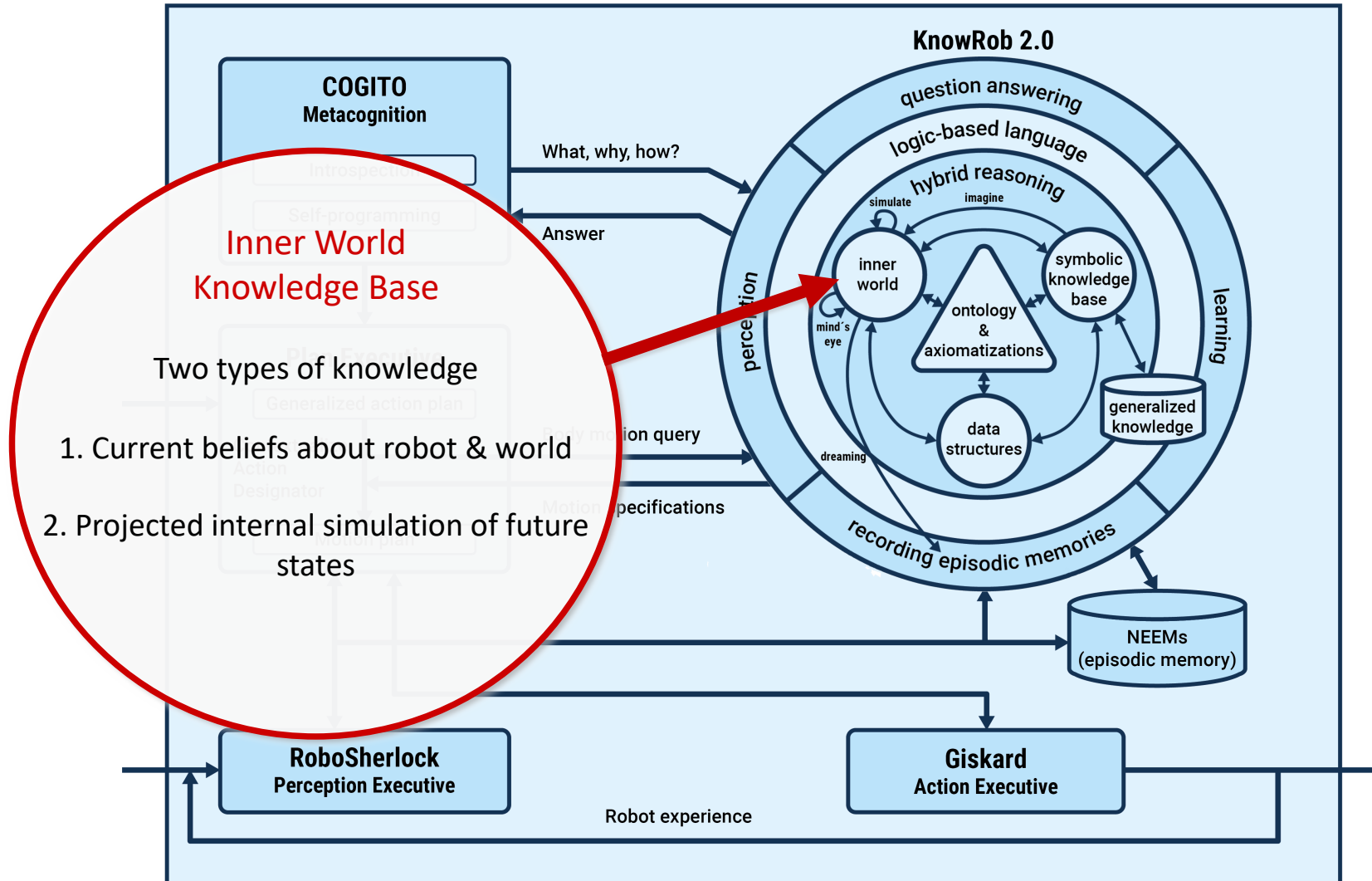


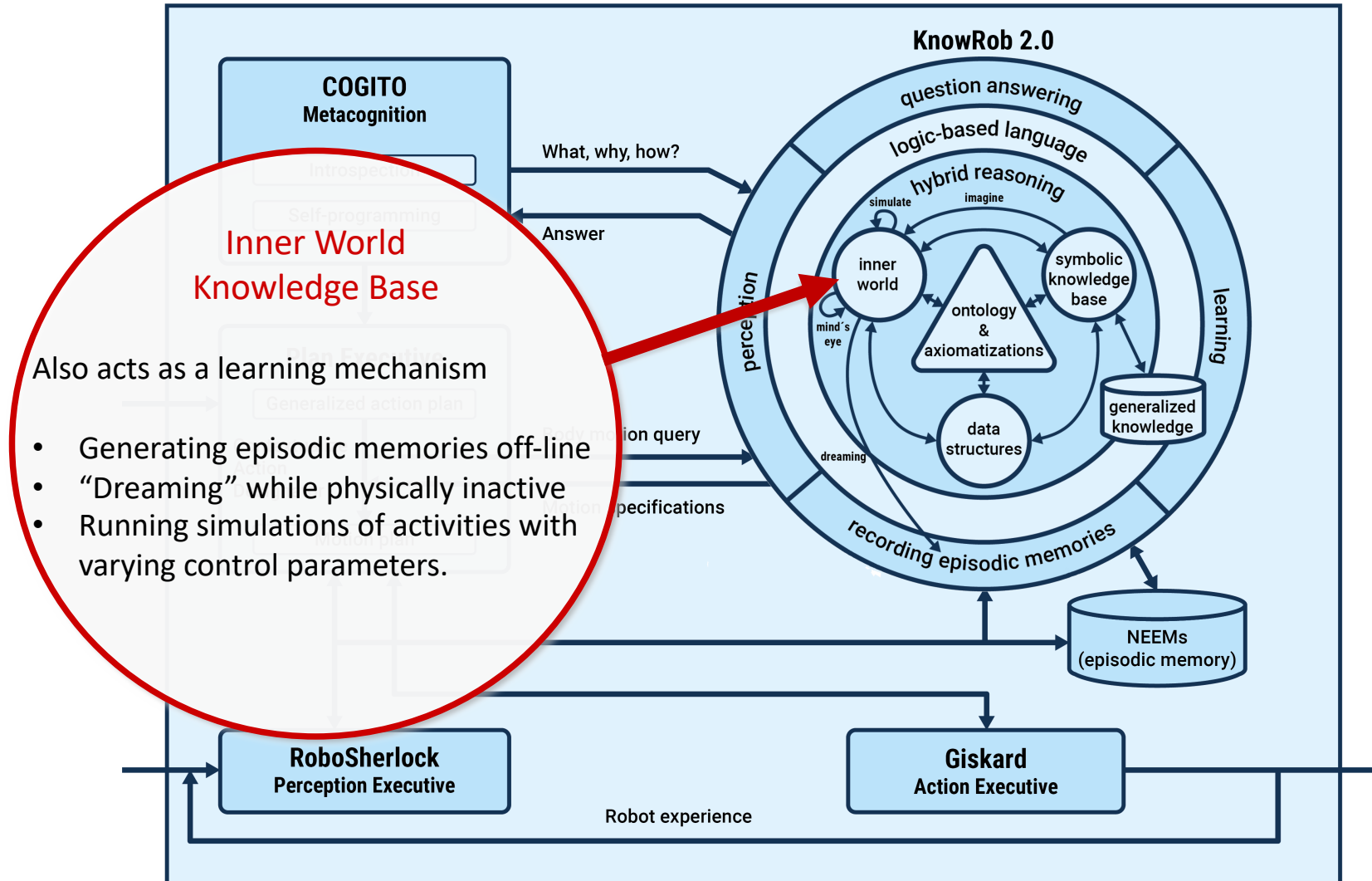


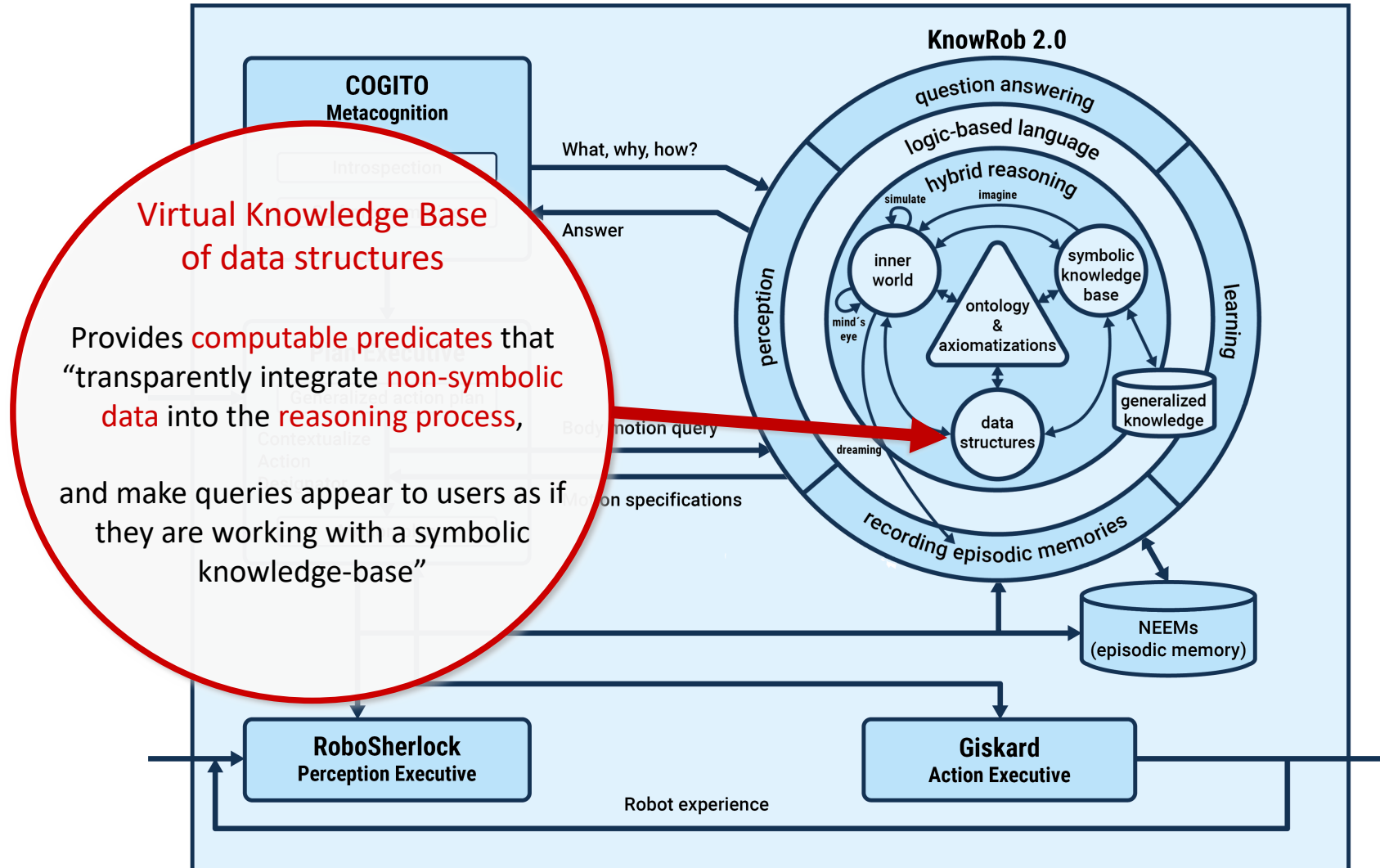


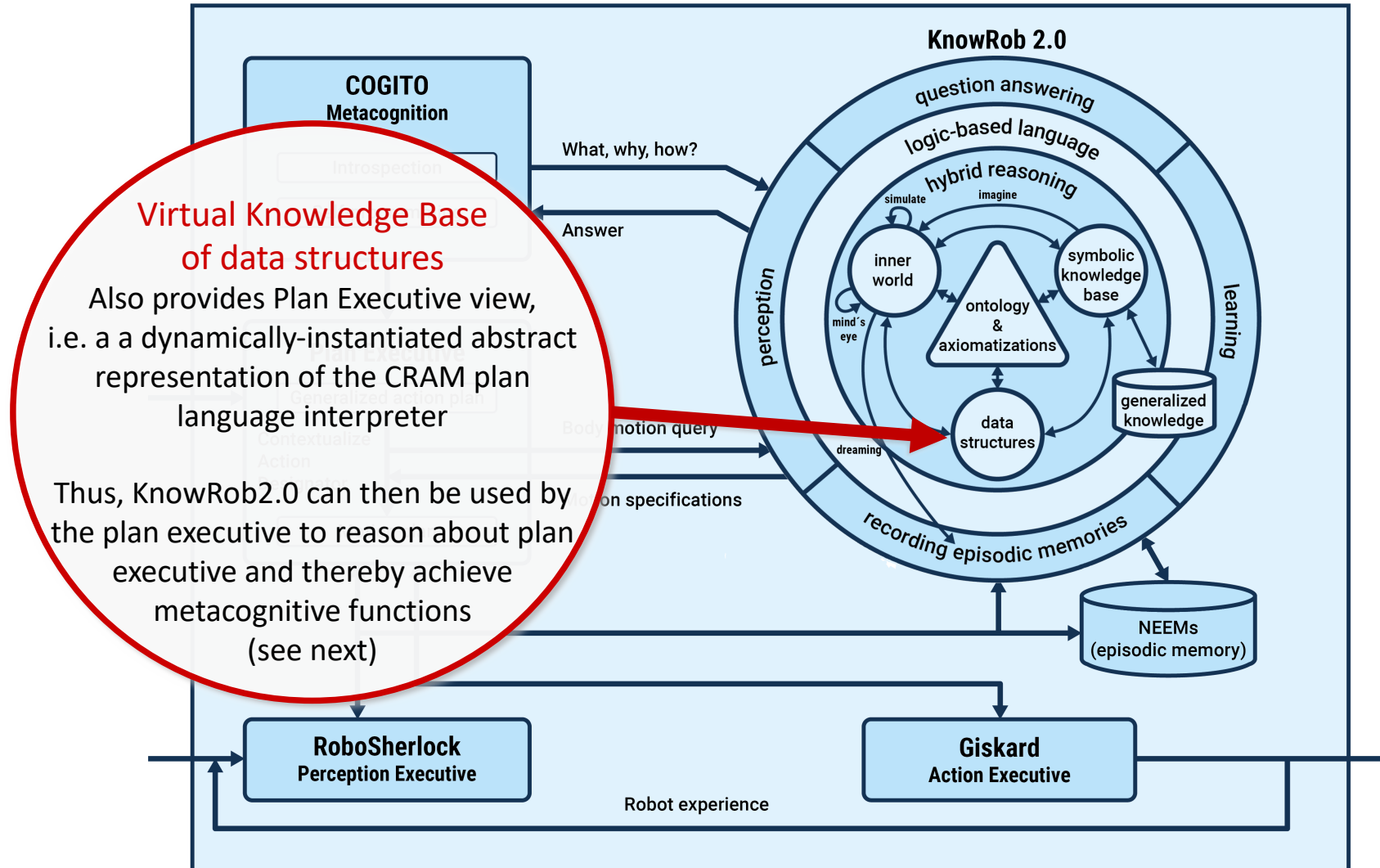


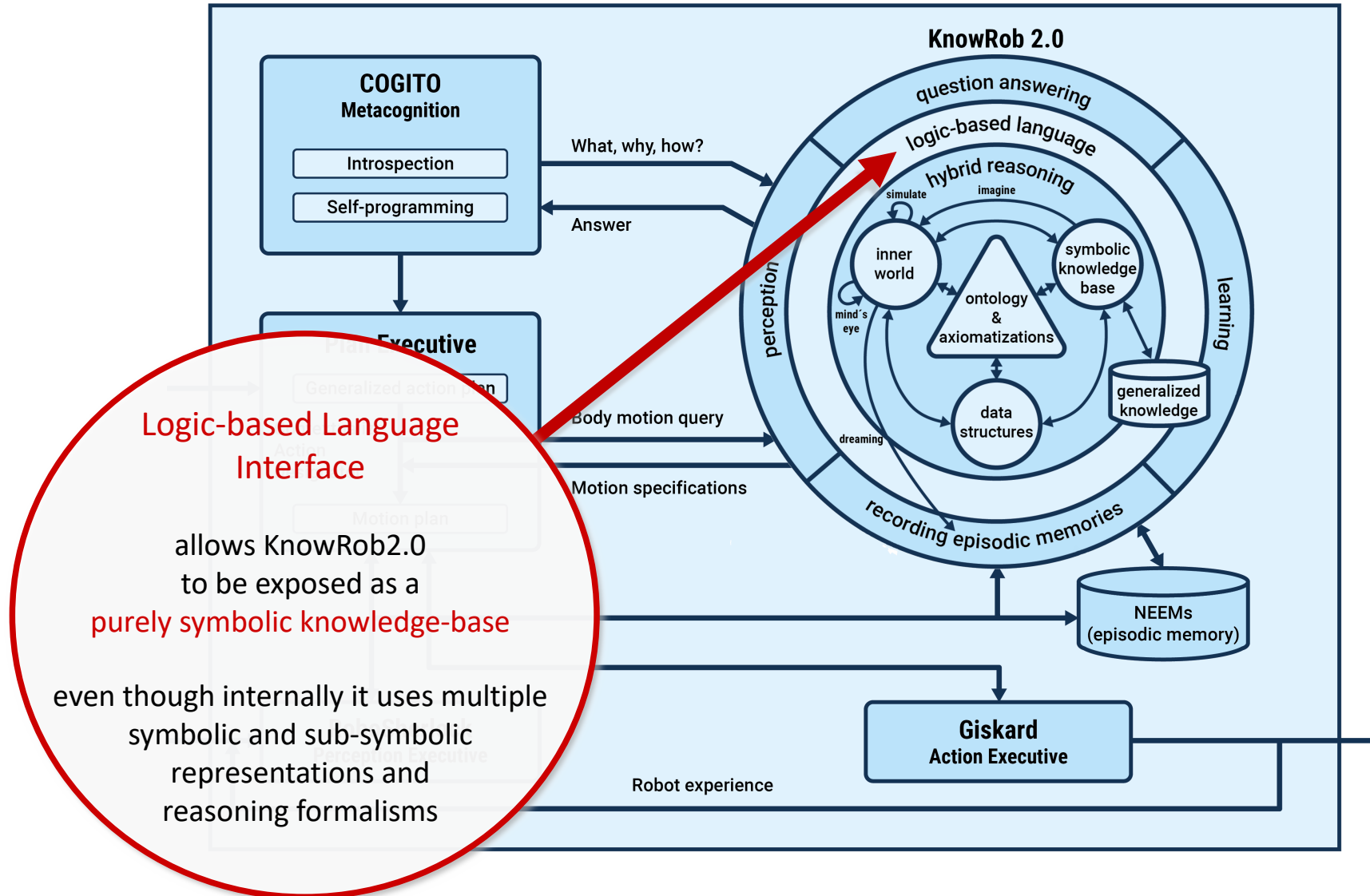


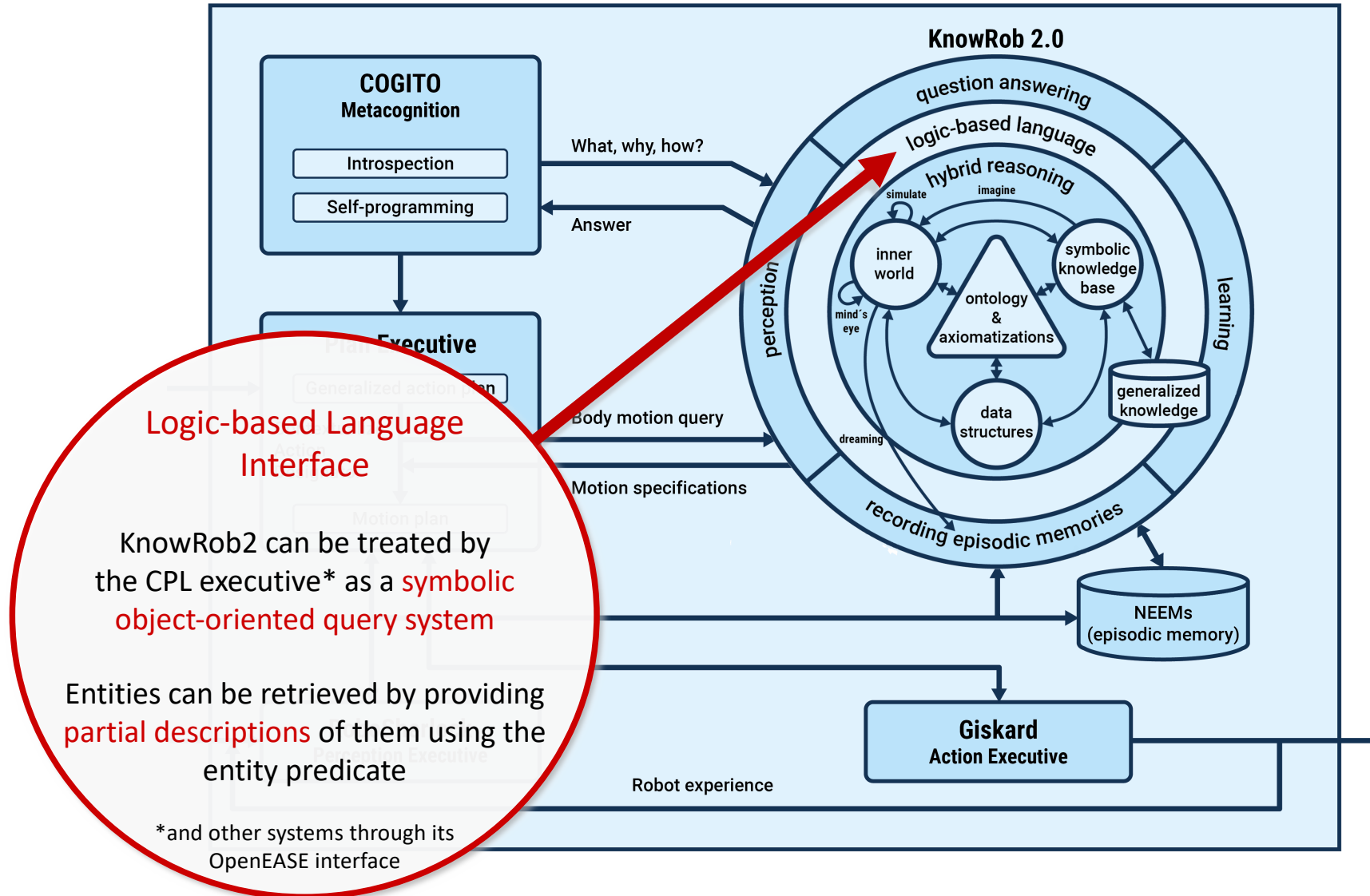


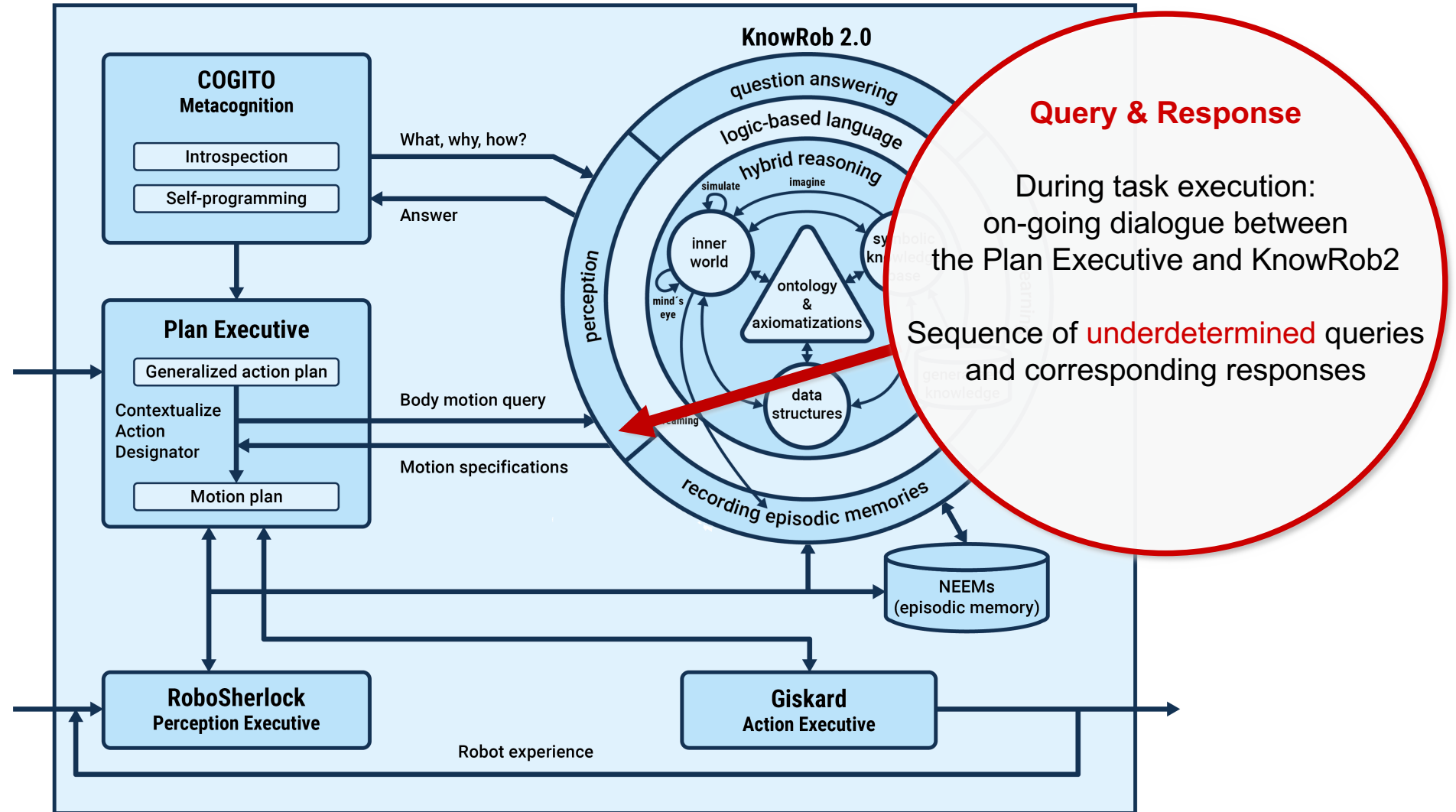


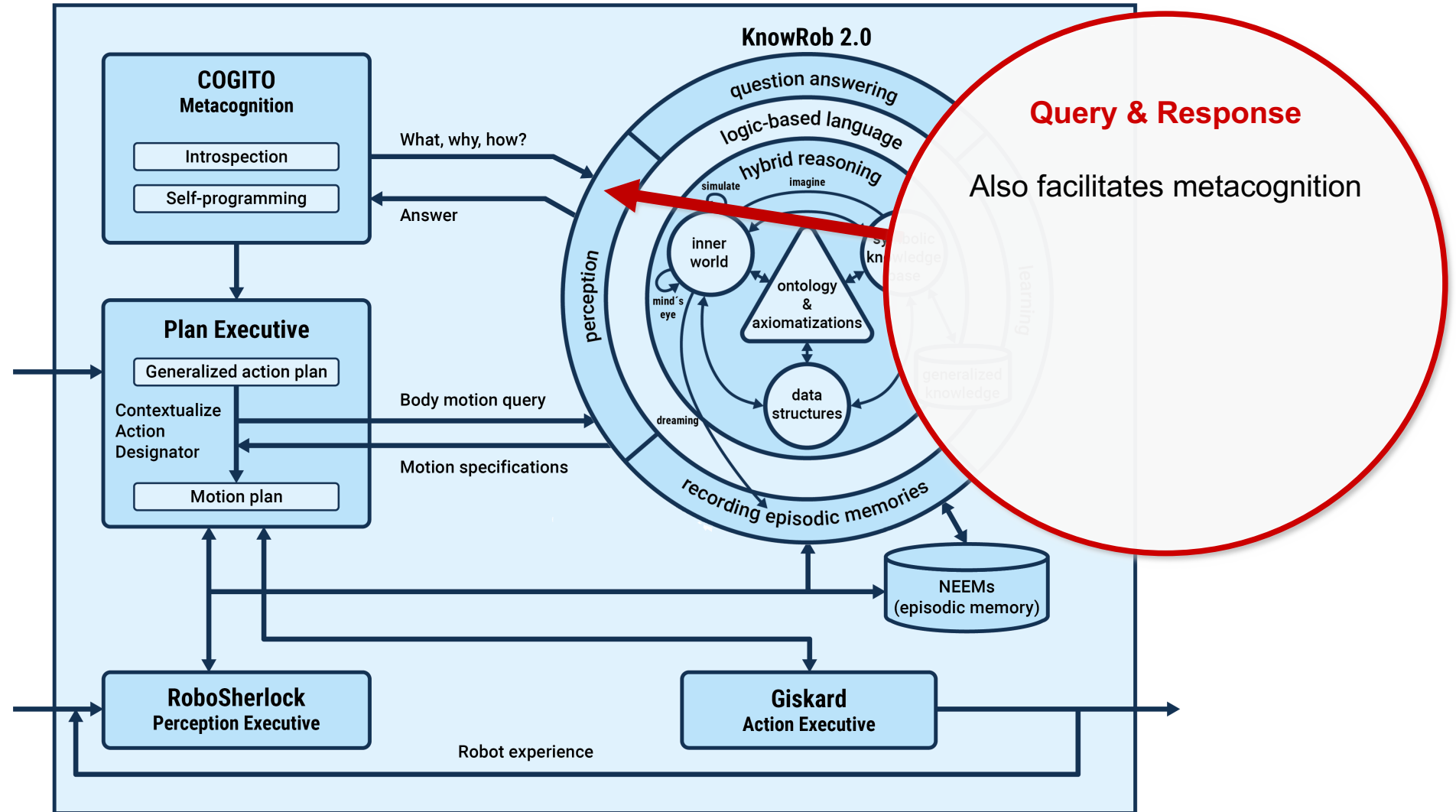


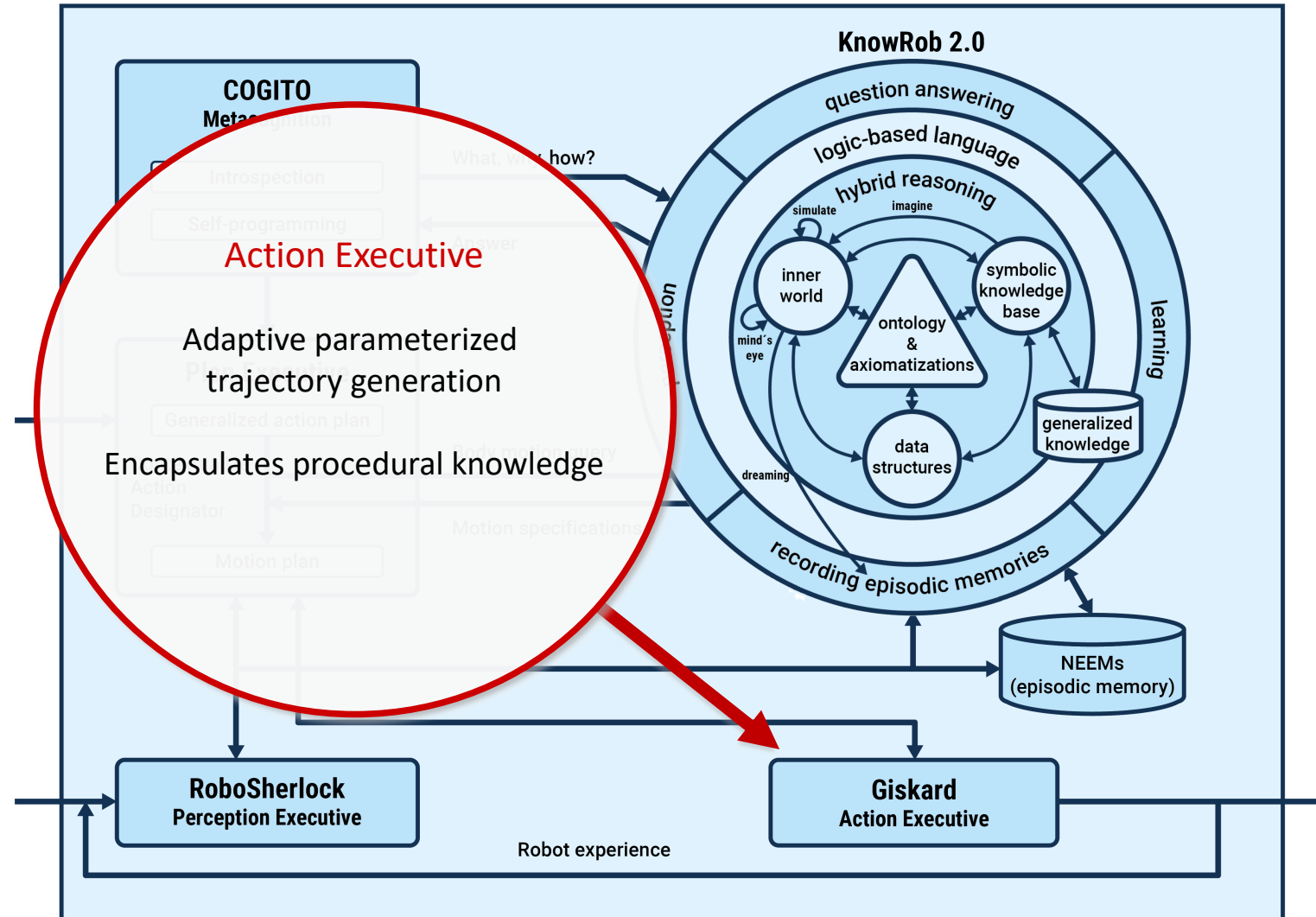


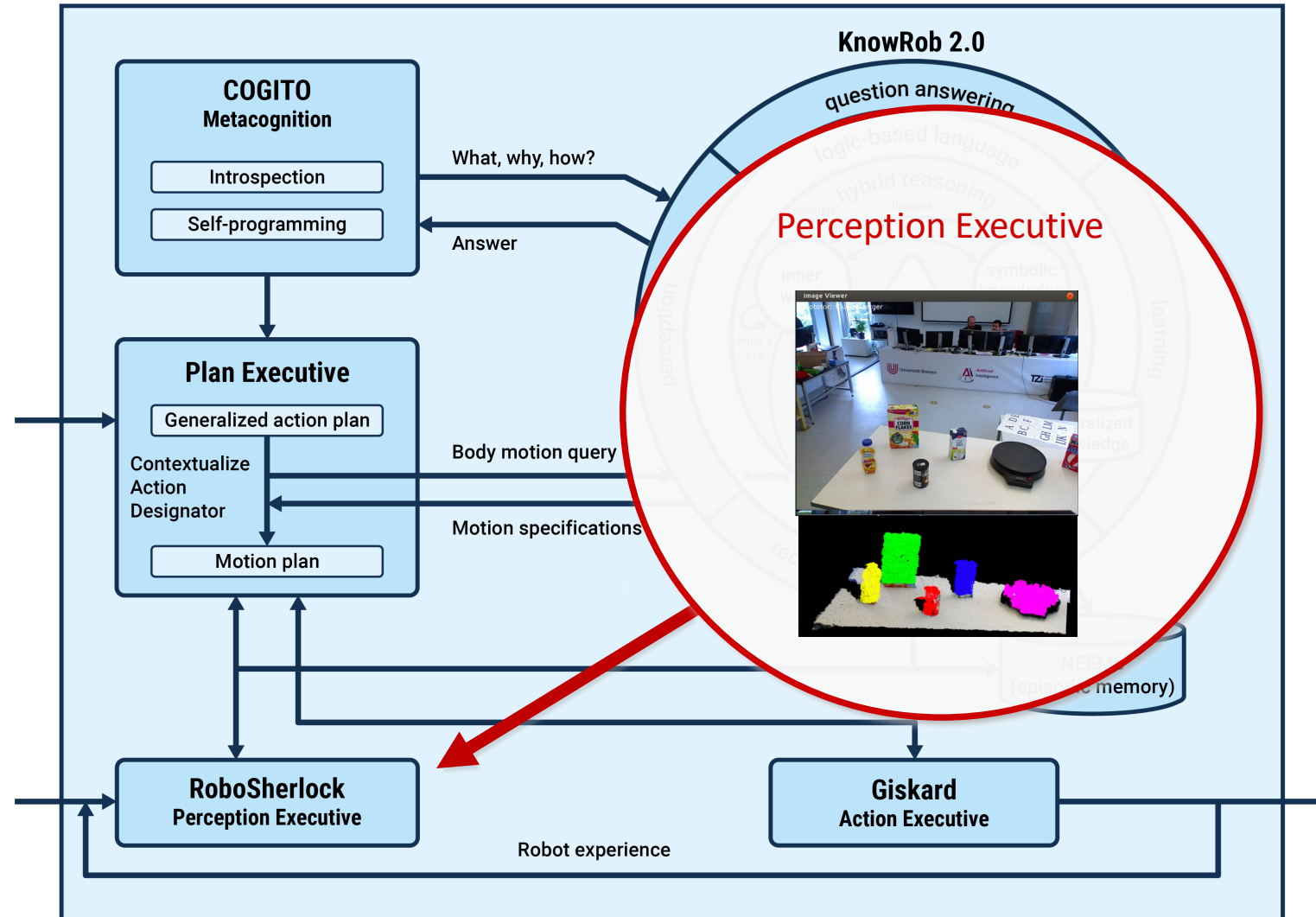


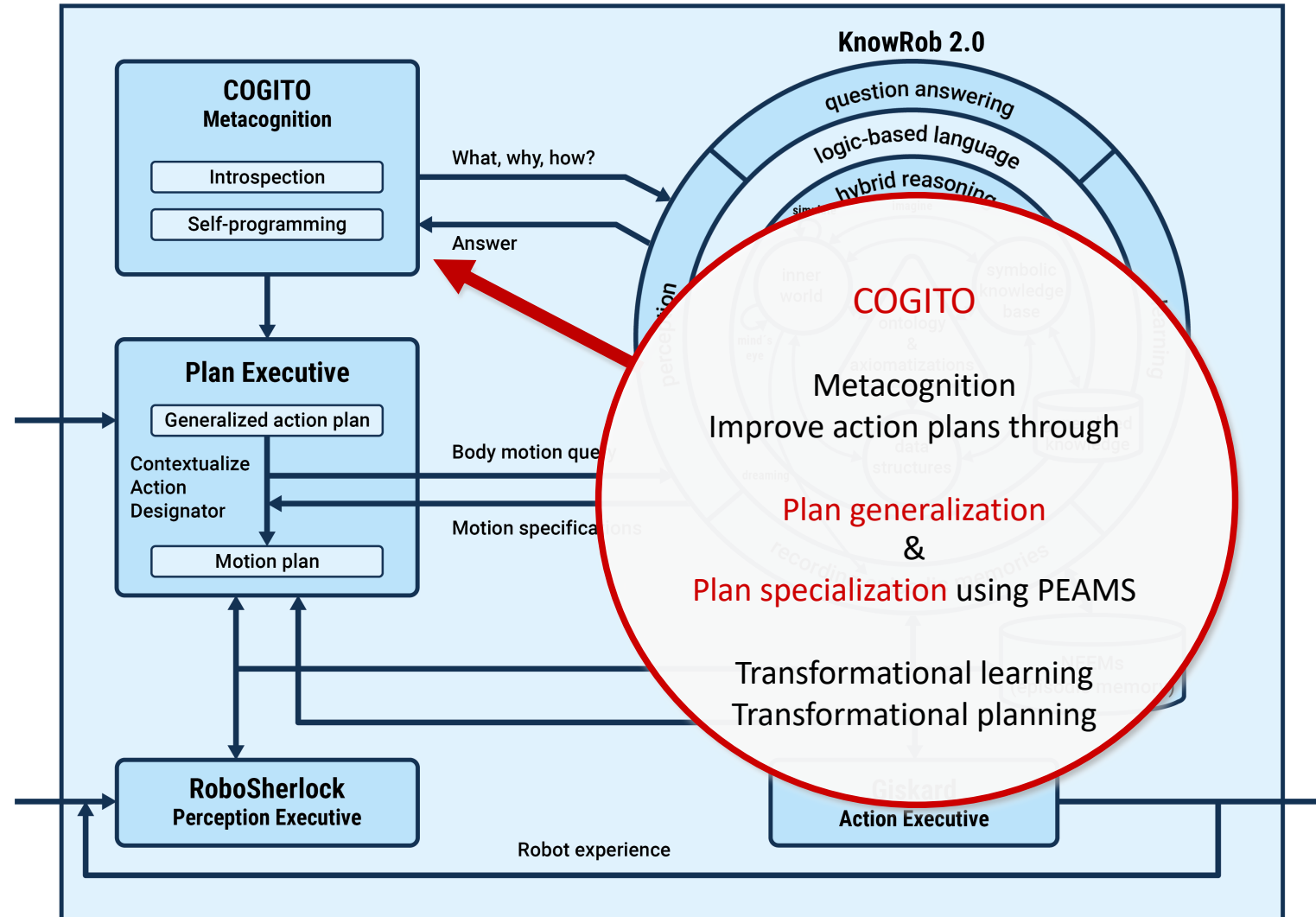


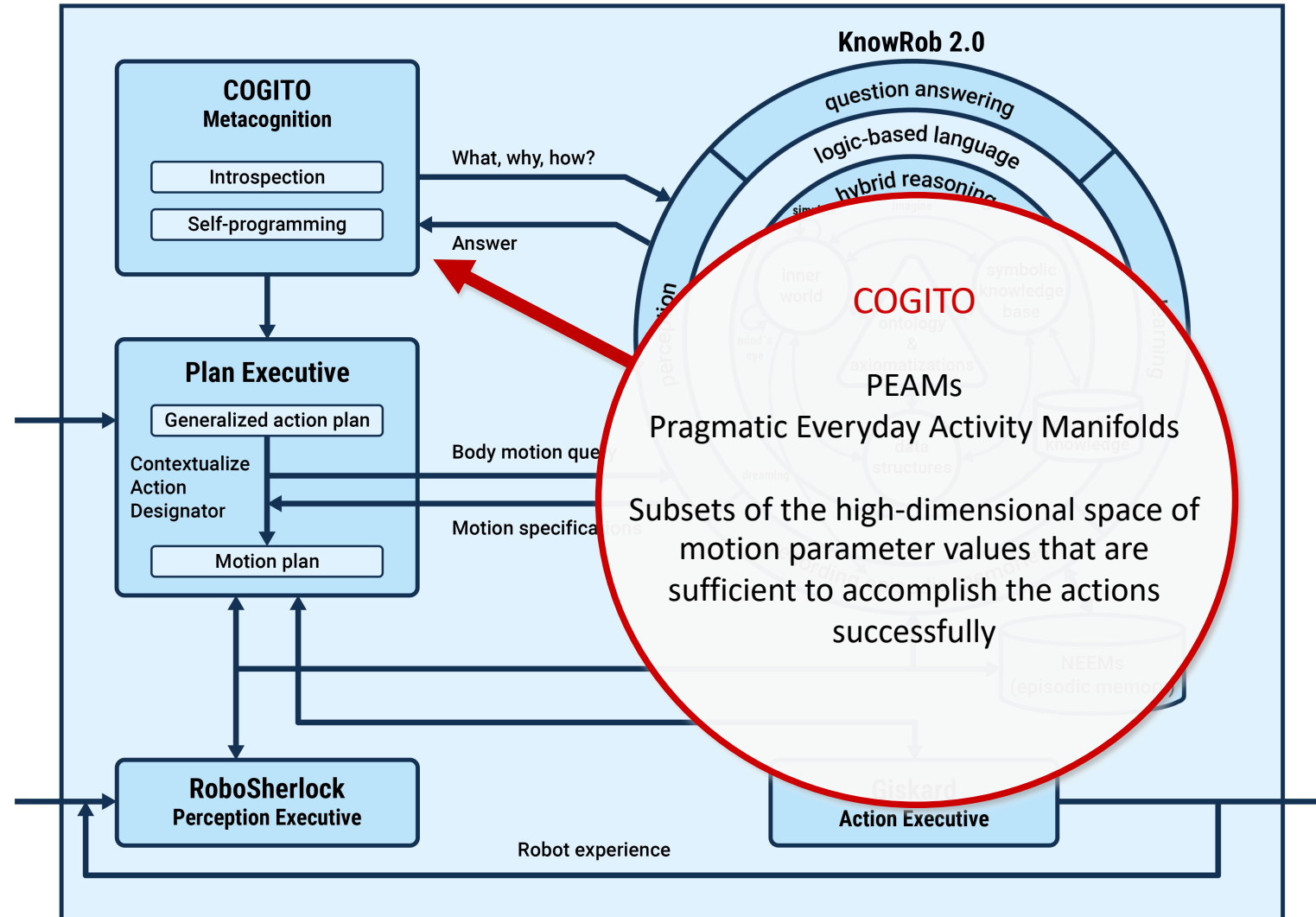










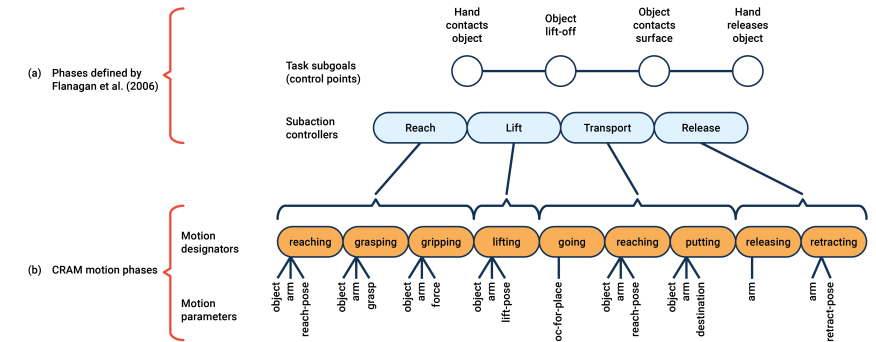


Walk through the execution of a **generalized action plan**

Generalized
Action Plan



Motion Plan

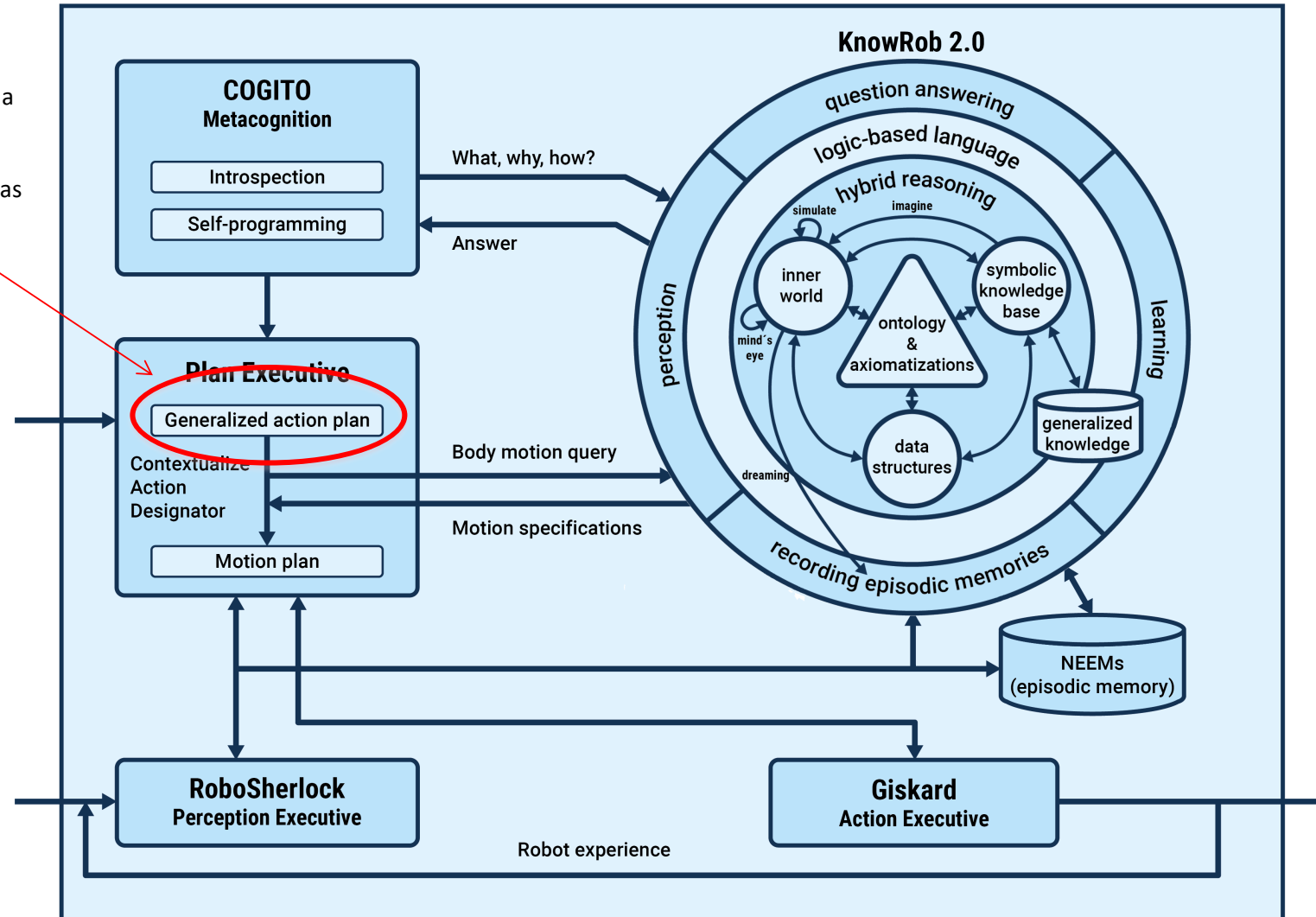


3-step refinement process:
contextualization

Identify the values of the parameters to the motion plan that maximize the likelihood that the associated body motions successfully accomplish the desired action

The motion parameter values are provided by the generative model

Recall:
The robot agent is equipped with a **generalized action plan** for each **action category**, which typically corresponds to action verbs such as **fetch**, **place**, **pour**, and **cut**.

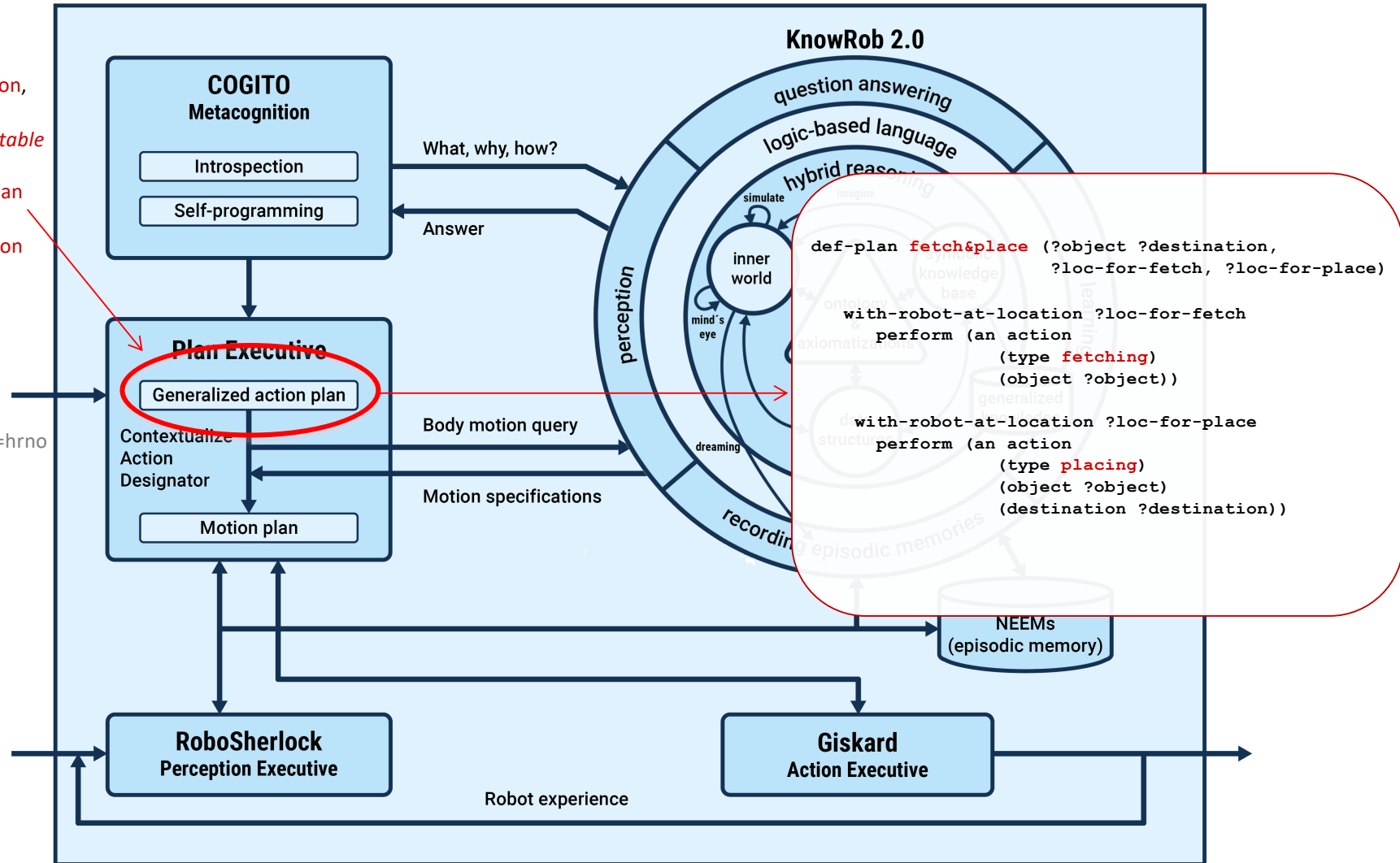


An action plan is invoked with a request to perform an **underdetermined action description**,

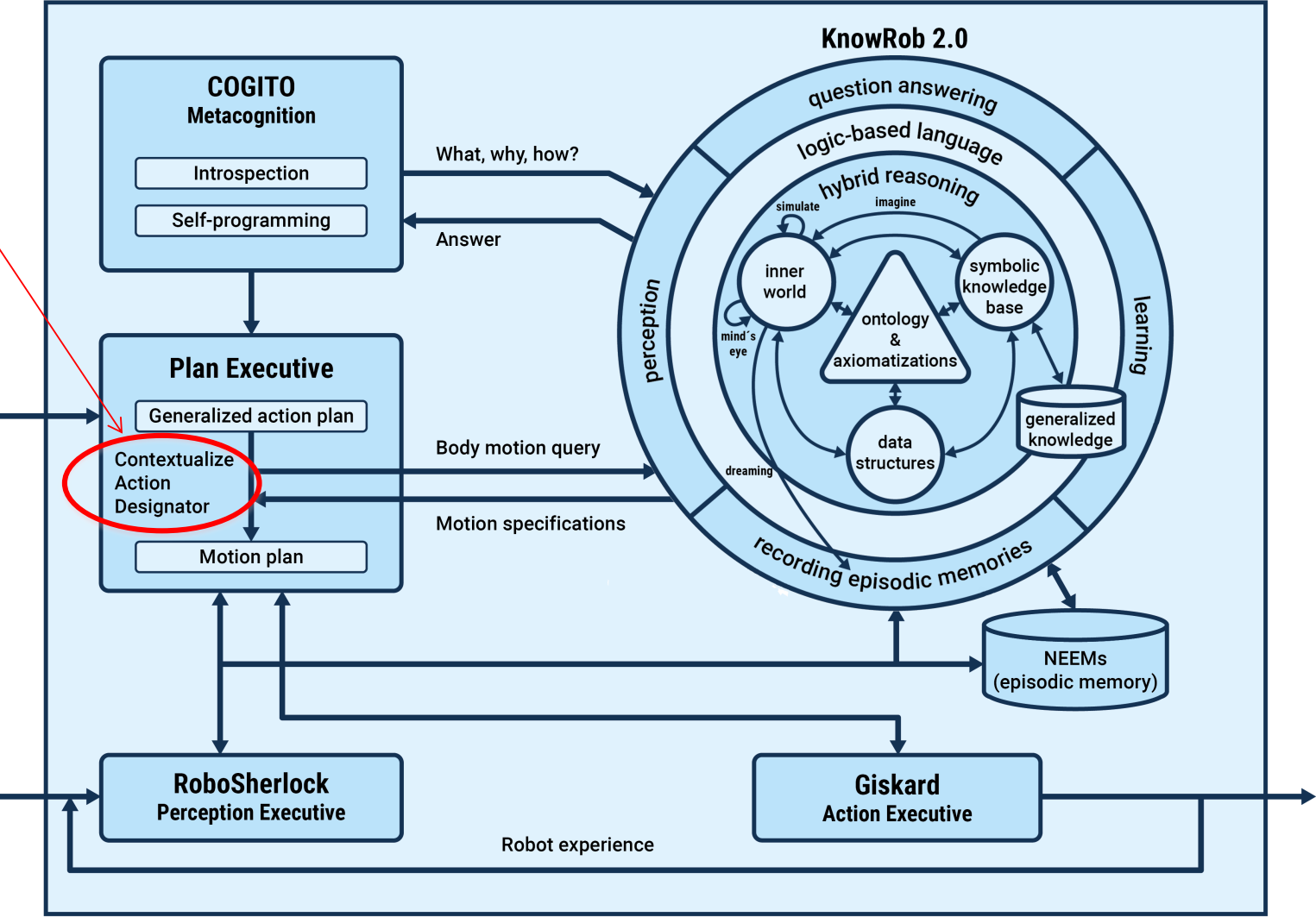
e.g. *fetch the cup and place it on the table*

by selecting the **generalized action plan** for the **action category** corresponding to the **action description**

(e.g. see the spoon challenge at <https://www.youtube.com/watch?v=hrnoY6J8ddE&feature=youtu.be>)

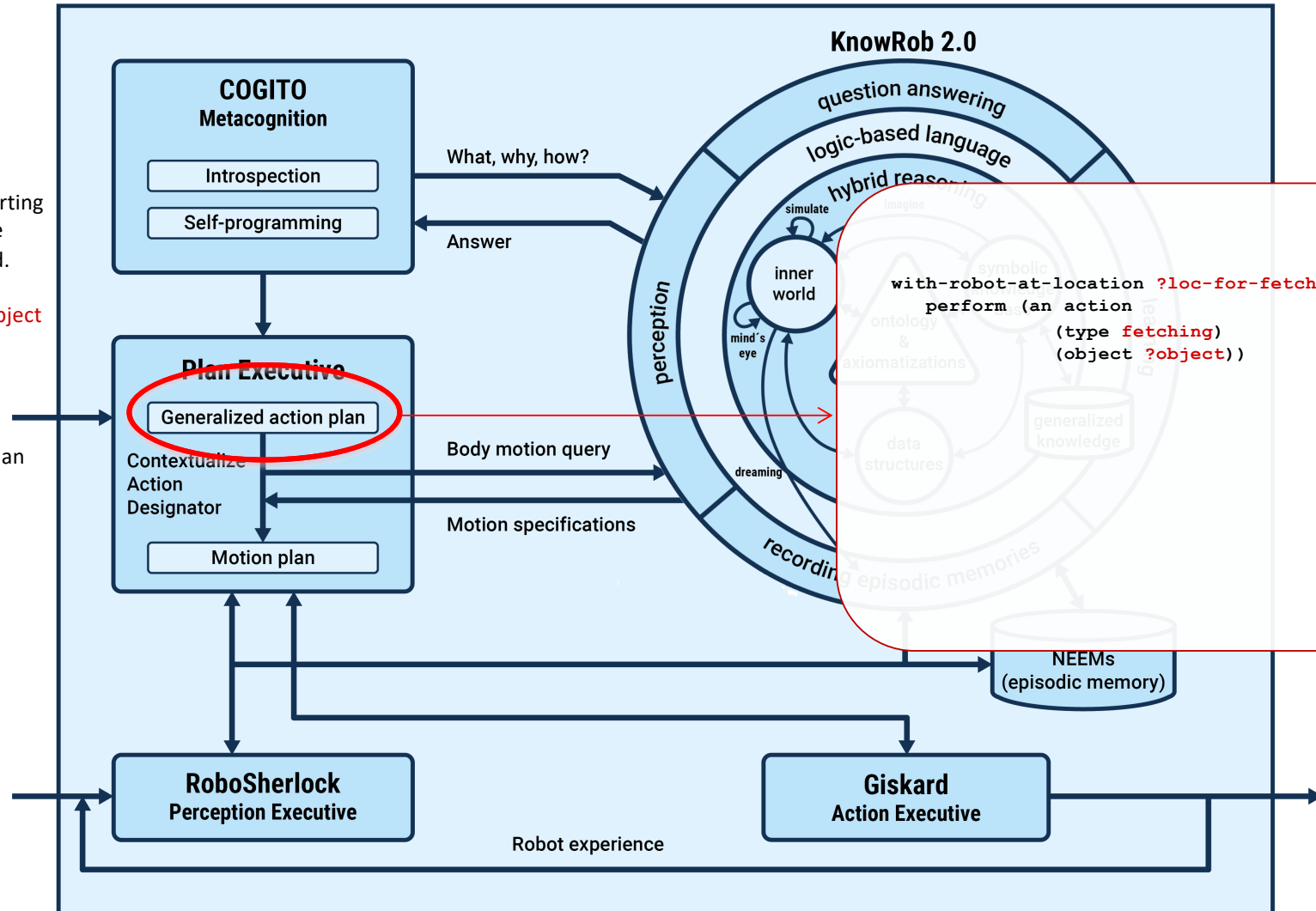


The Plan Executive interprets the generalized action plan in process referred to as **contextualization** in **three** steps:



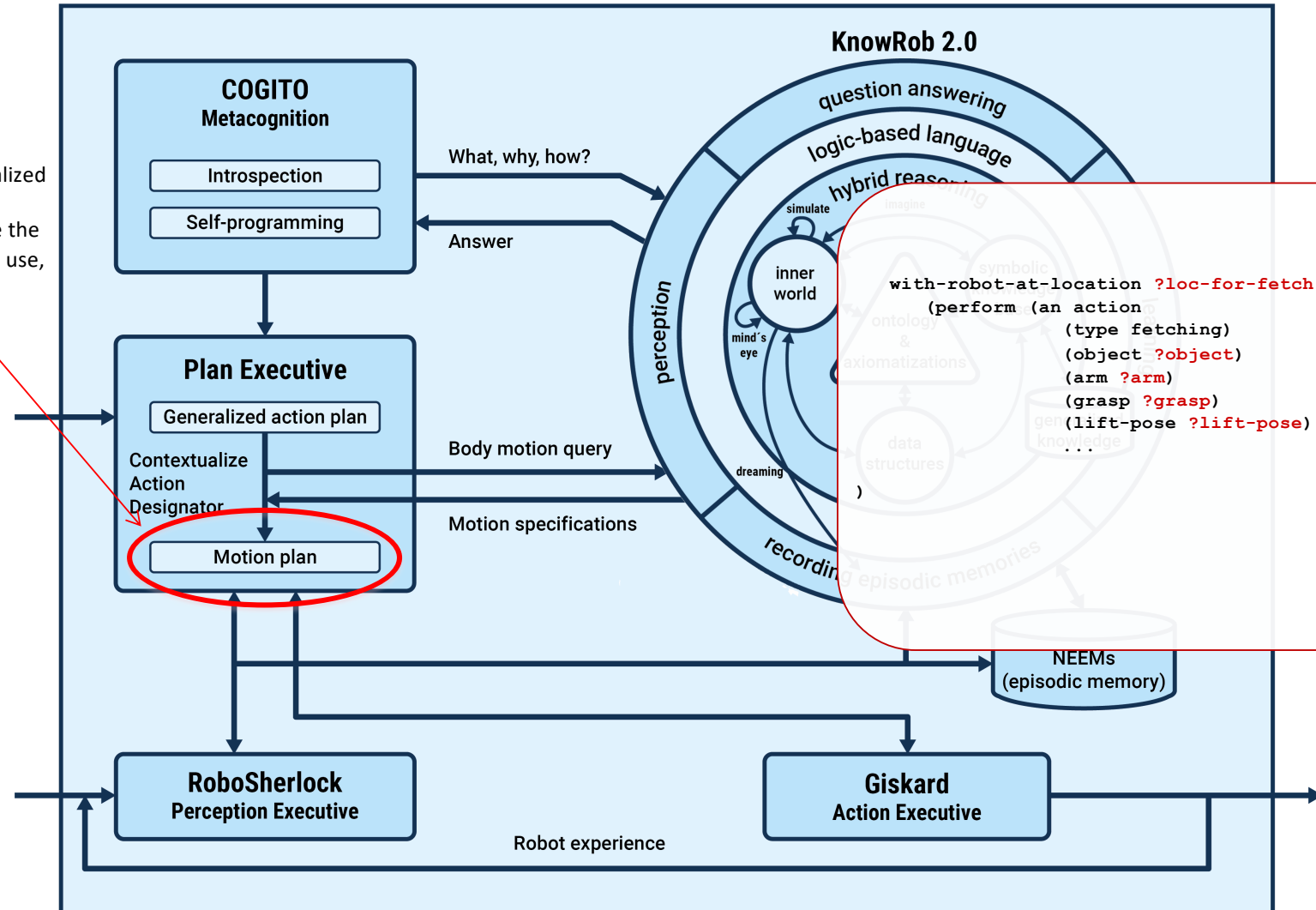
The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

1. **Instantiate** the selected generalized action plan by inserting the arguments required for the specific action to be performed.
- For example, the type of the **object** to be manipulated or the destination **location**.
- These arguments are typically **designators** of some kind, e.g., an **action**, **object**, or **location** designator.



The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

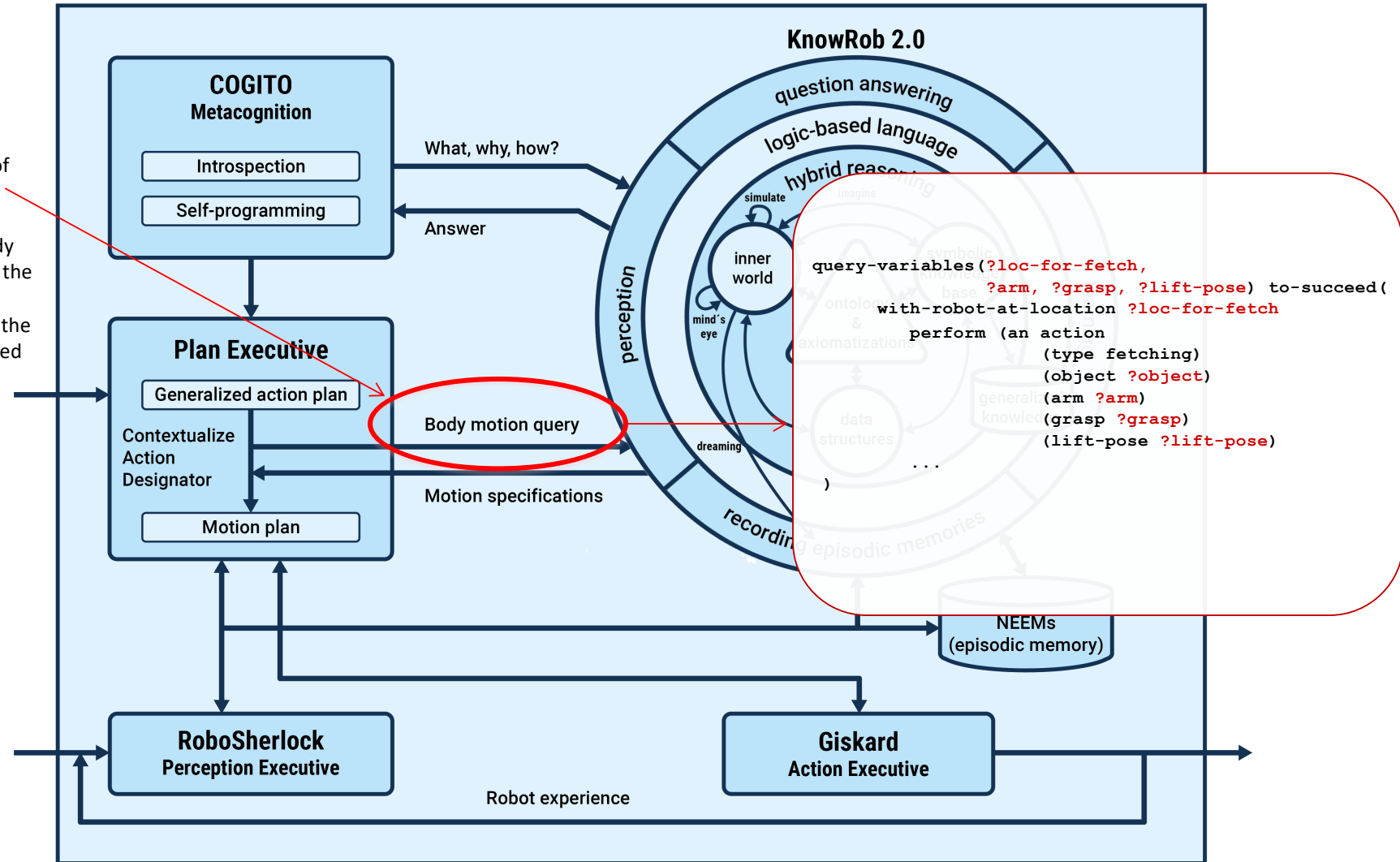
2. **Extend** the instantiated generalized action plan by **adding the parameters** needed to execute the **motion plan**, e.g. which arm to use, what grasp pose to use.



The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

3. Create a **query** for the values of these parameters

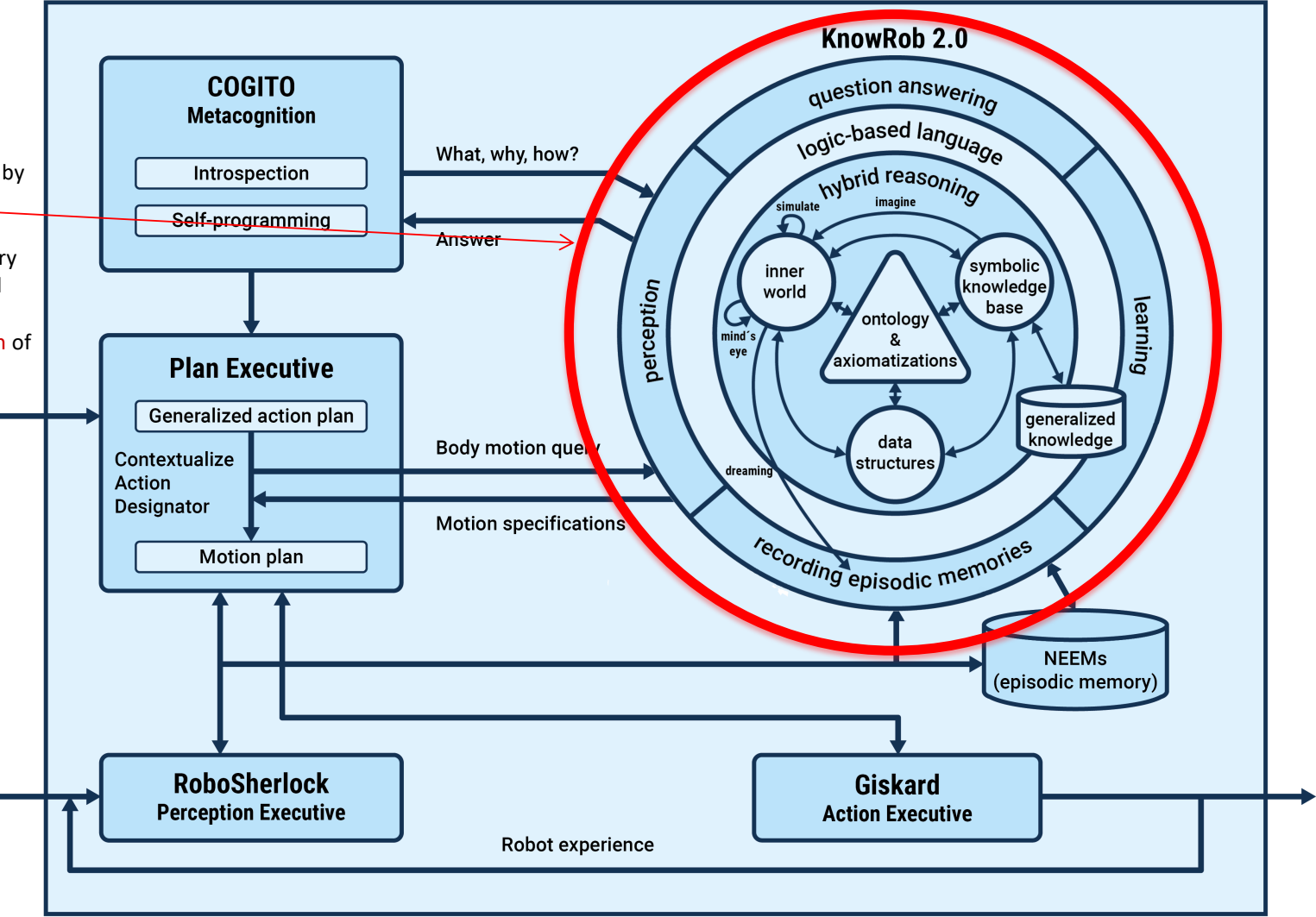
(that would produce robot body motions to achieve the goal of the underdetermined action description and, equivalently, the associated instantiated extended generalized action plan).



The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

The contextualization is accomplished by KnowRob2.0.

The motion parameter values necessary to carry out the action are determined using a **generative model** effectively sampling a **joint distribution** of the motion parameter values and the associated outcome.

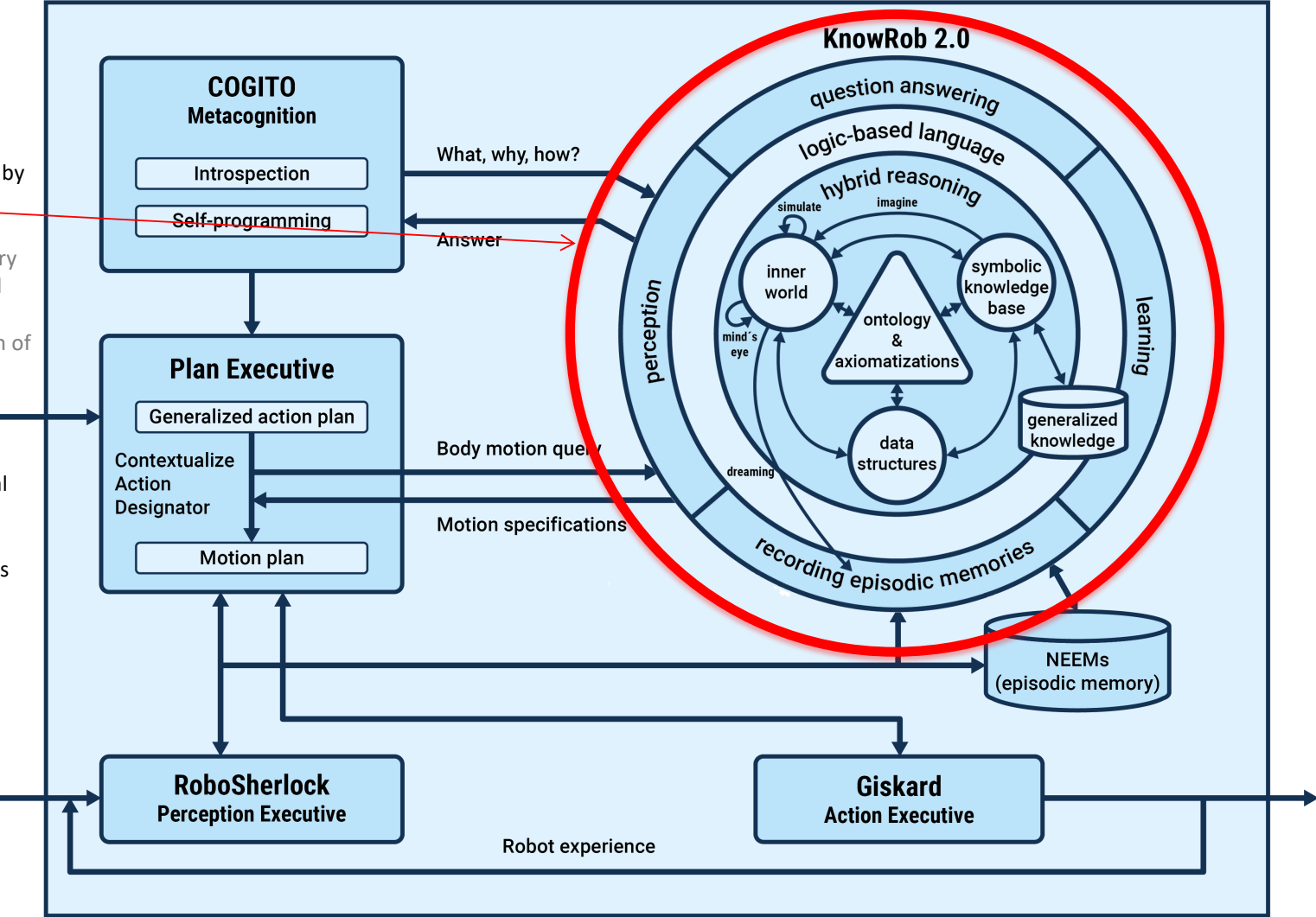


The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

The contextualization is accomplished by KnowRob2.0.

The motion parameter values necessary to carry out the action are determined using a generative model effectively sampling a joint distribution of the motion parameter values and the associated outcome.

It uses **knowledge and reasoning**, exploiting the constraints of contextual knowledge and current perceptual information, and **prospection**, to maximize the likelihood that the values identified are most likely to result in a successful action.



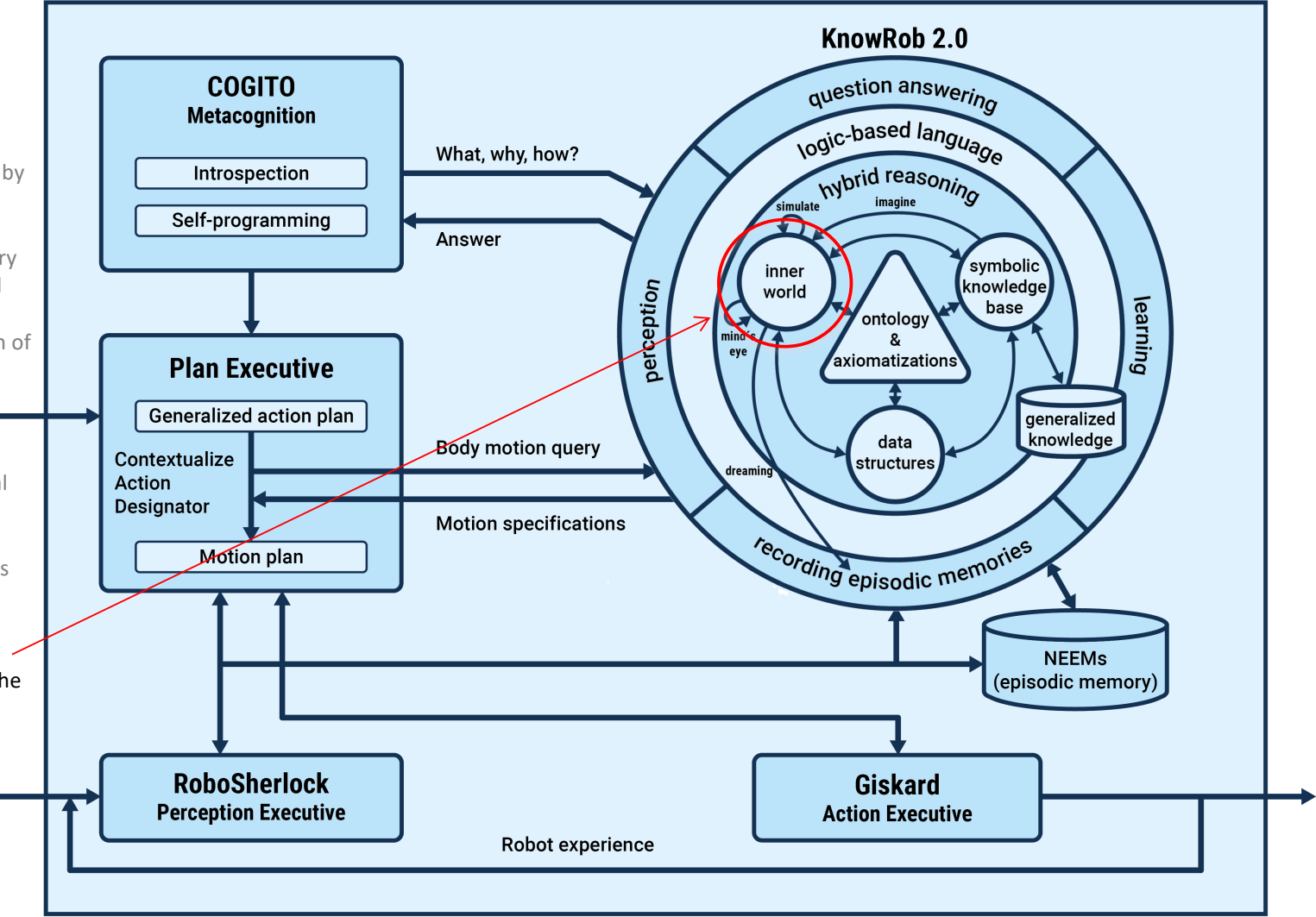
The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

The contextualization is accomplished by KnowRob2.0.

The motion parameter values necessary to carry out the action are determined using a generative model effectively sampling a joint distribution of the motion parameter values and the associated outcome.

It uses knowledge and reasoning, exploiting the constraints of contextual knowledge and current perceptual information, and prospection, to maximize the likelihood that the values identified are most likely to result in a successful action.

It accomplishes prospection by using the robot's **inner world** to simulate plan execution



The Plan Executive interprets the generalized action plan in process referred to as contextualization in three steps:

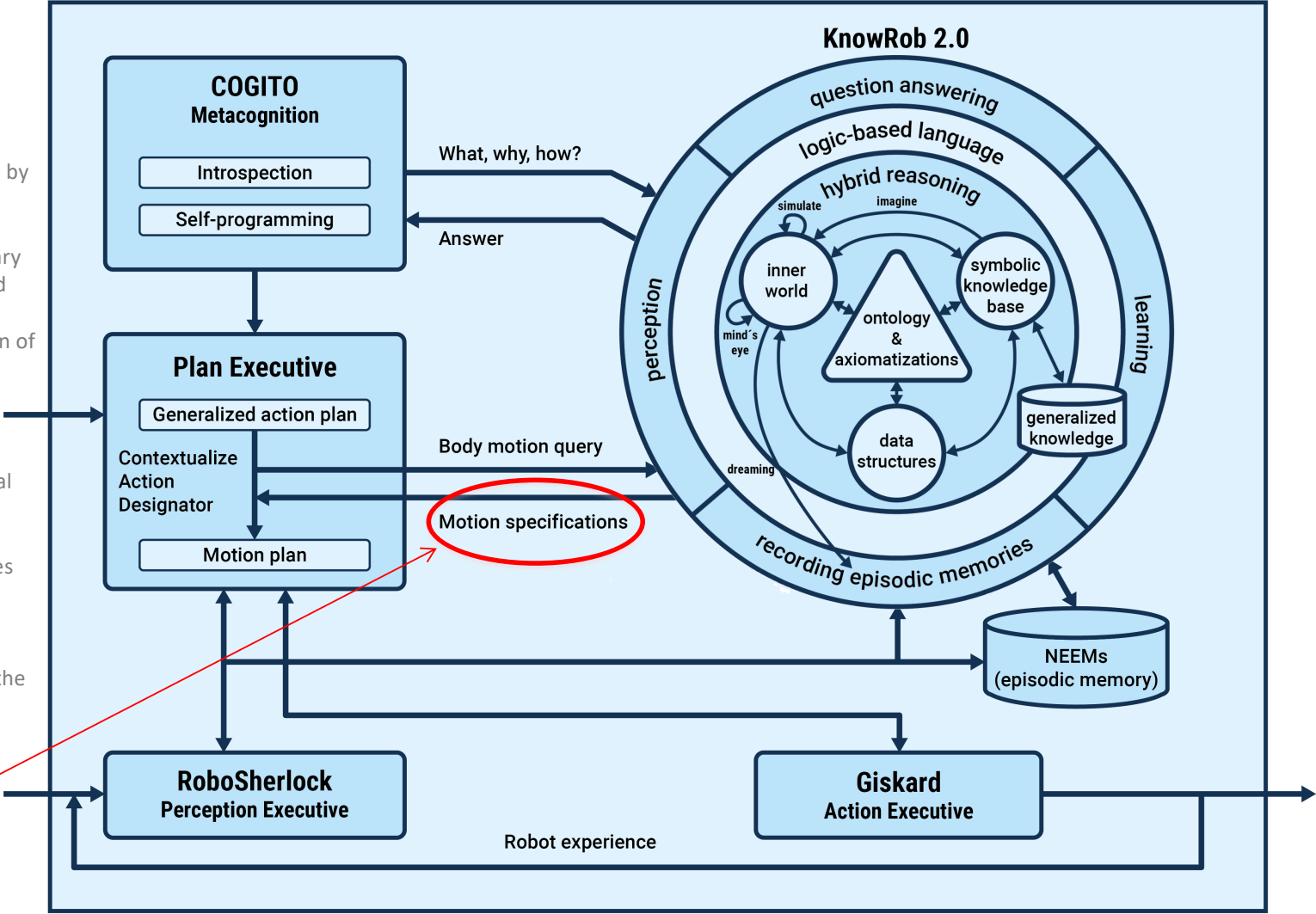
The contextualization is accomplished by KnowRob2.0.

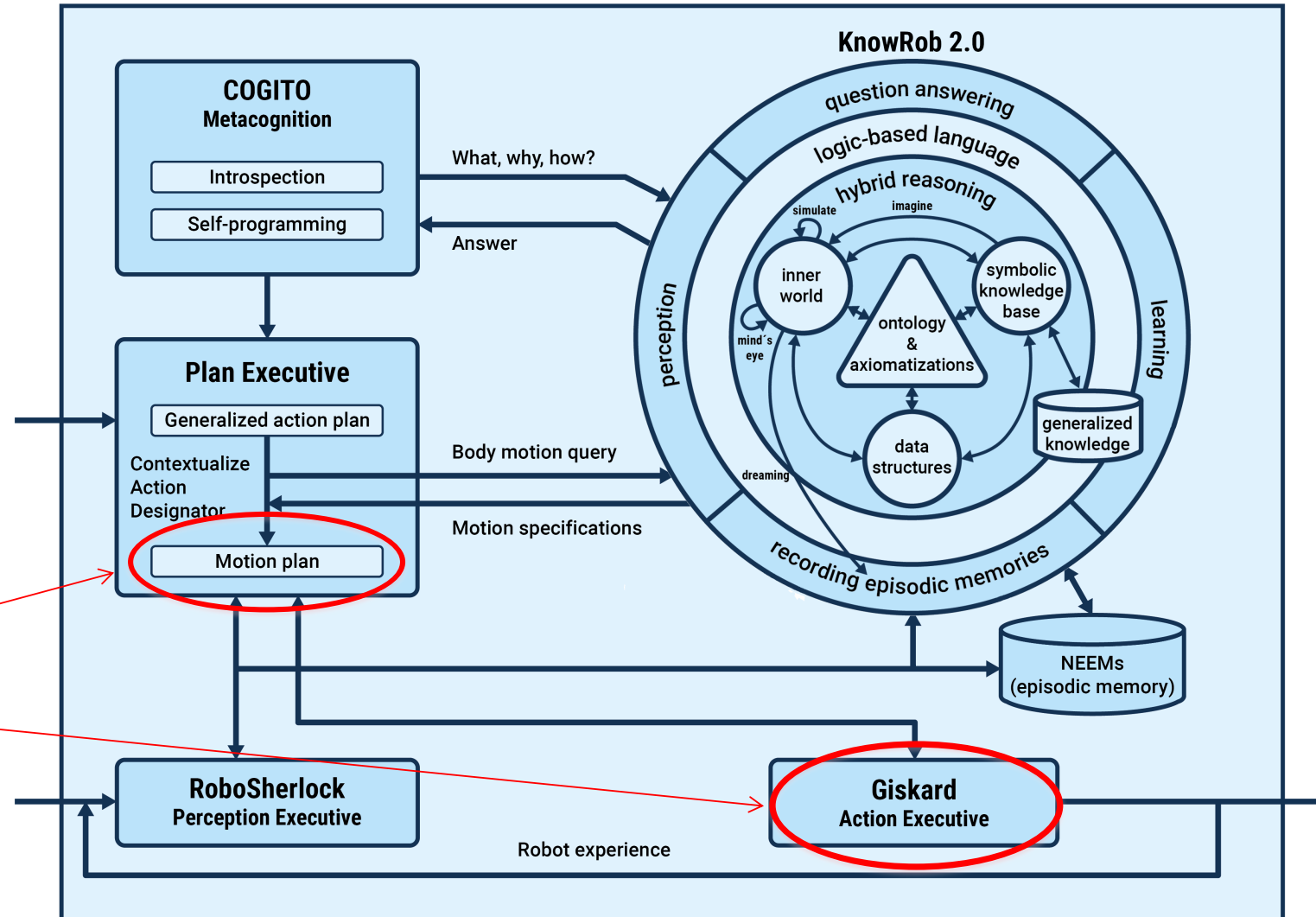
The motion parameter values necessary to carry out the action are determined using a generative model effectively sampling a joint distribution of the motion parameter values and the associated outcome.

It uses knowledge and reasoning, exploiting the constraints of contextual knowledge and current perceptual information, and prospection, to maximize the likelihood that the values identified are most likely to result in a successful action.

It accomplishes prospection by using the robot's inner world to simulate plan execution

The motion parameter values are returned to the Plan Executive





Reading

D. Vernon, "Cognitive Architectures", in Cognitive Robotics, A. Cangelosi and M. Asada (Eds.), MIT Press, Chapter 10, 2022,
Section 10.6.1.

Further Reading

D. Vernon, J. Albert, M. Beetz, S.-C. Chiou, H. Ritter, and W. X. Schneider, "Action Selection and Execution in Everyday Activities: A Cognitive Robotics & Situation Model Perspective", Topics in Cognitive Science, pp. 1-19, 2021.