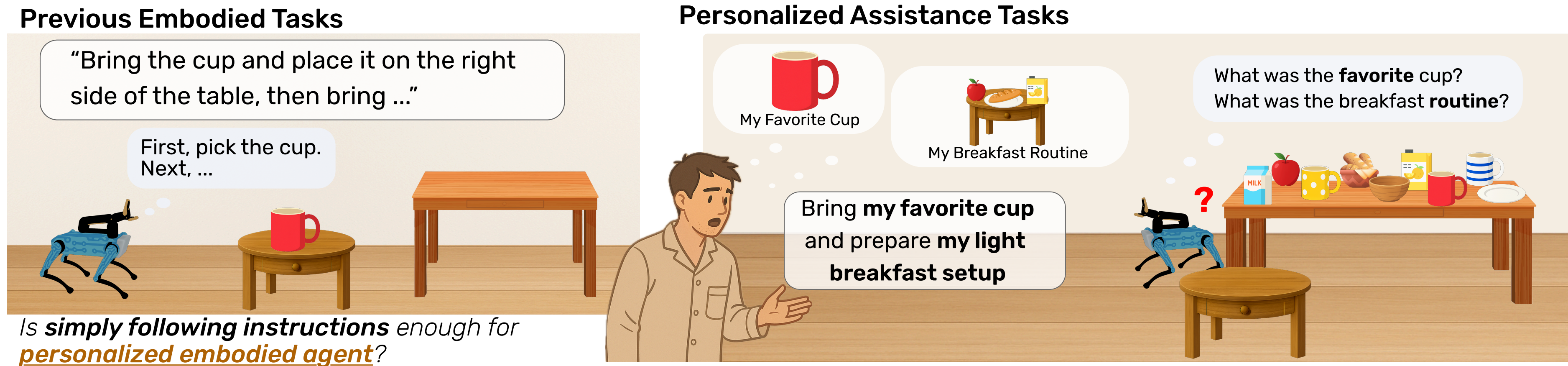


Investigating Challenges and Solutions Through the Lens of Memory Utilization

Taeyoon Kwon*, Dongwook Choi*, Hyojun Kim, Sunghwan Kim, Seungjun Moon, Beong-woo Kwak, Kuan-Hao Huang, Jinyoung Yeo

A Motivation

- Current LLM-powered embodied agents succeed at simple tasks but **struggle with real-world personalized assistance**.
- Personalized assistance require agents to **understand user-specific meanings** (e.g., favorite items, routines) from **previous interaction histories**.

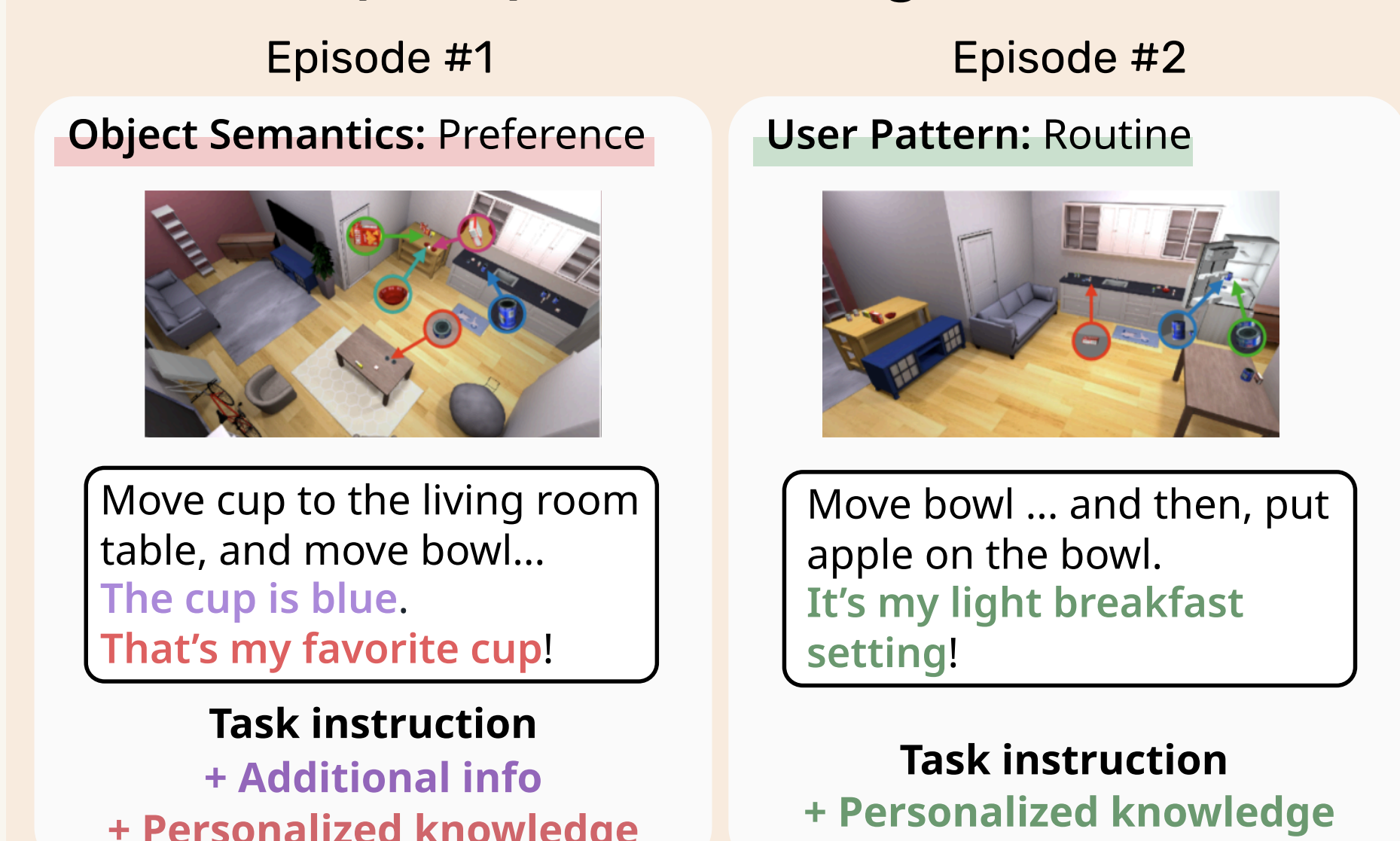


“We investigate the challenges faced by LLM-powered embodied agents in personalized assistance tasks through the lens of memory utilization”

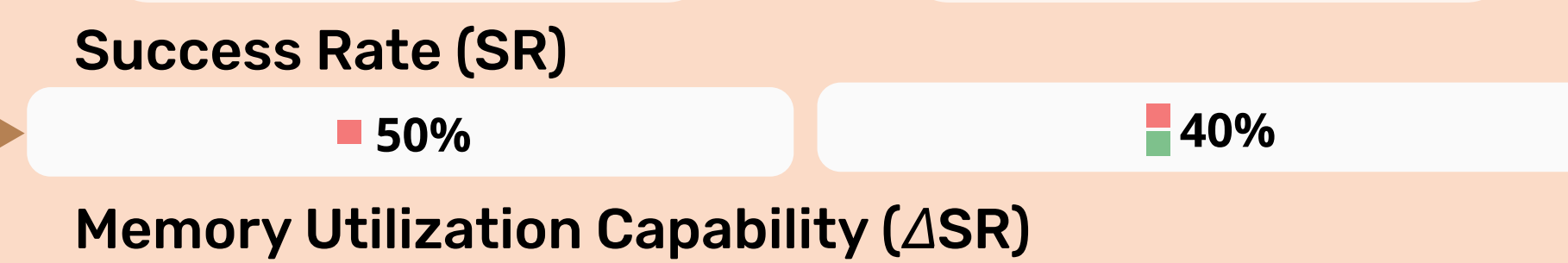
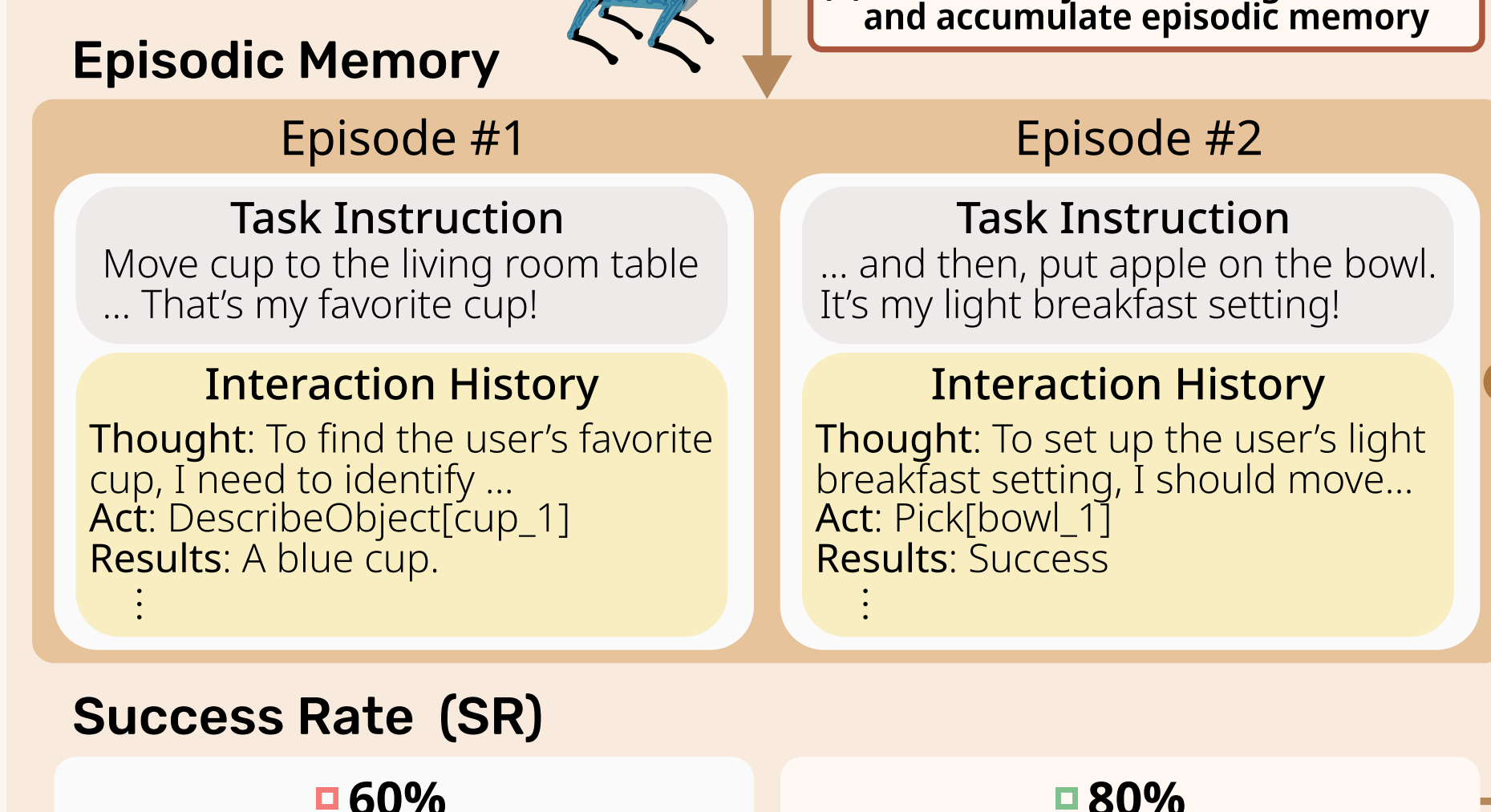
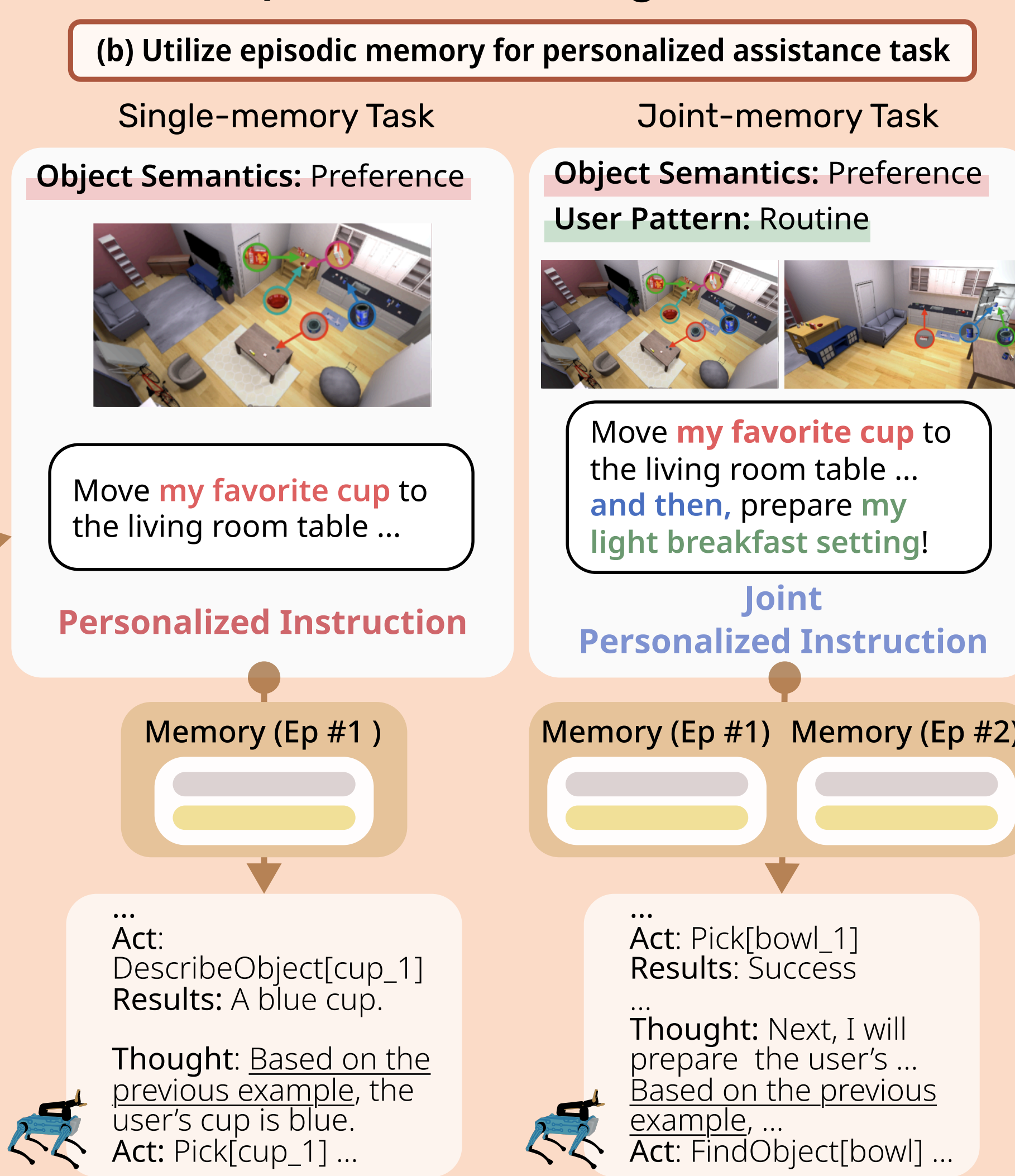
B Evaluation Setup: MEMENTO

- Dimension of personalized knowledge analysis
 - Object Semantics**: identifying objects based on personal meaning
 - User Patterns**: recalling sequences from behavioral routines
- Key concept: Share the same episodes while varying the instruction

1. Memory Acquisition Stage



2. Memory Utilization Stage



(c) Evaluate memory utilization capability by comparing the Success Rate (ΔSR)

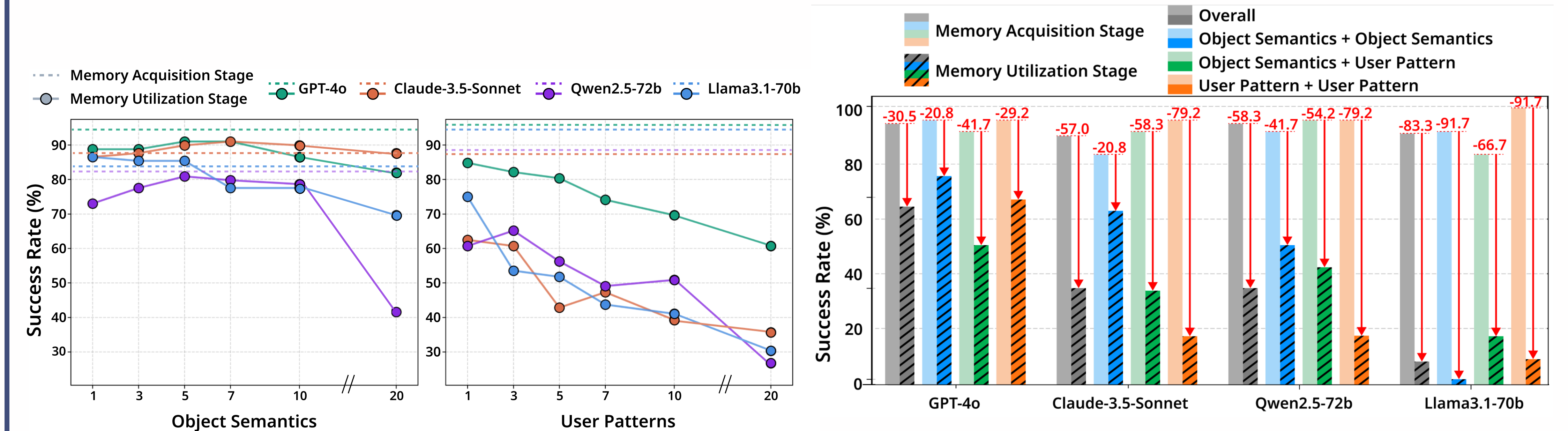
O Takeaways

- We propose **MEMENTO**, a two-stage evaluation framework to assess embodied agents' memory utilization capabilities
- Key findings**
 - While agents can effectively recall object semantics, they **struggle to comprehend user patterns**
 - Embodied agents suffer from **information overload** and **multi-memory coordination failures**
 - Episodic memory is an essential module** for personalized embodied agents
- We devise **user profile memory** a hierarchical knowledge graph that manages personalized knowledge independently



D Scenario-Based Analysis For Memory Utilization Bottlenecks

► **RQ2. What are the key factors that negatively affect embodied agents' memory utilization capabilities?**

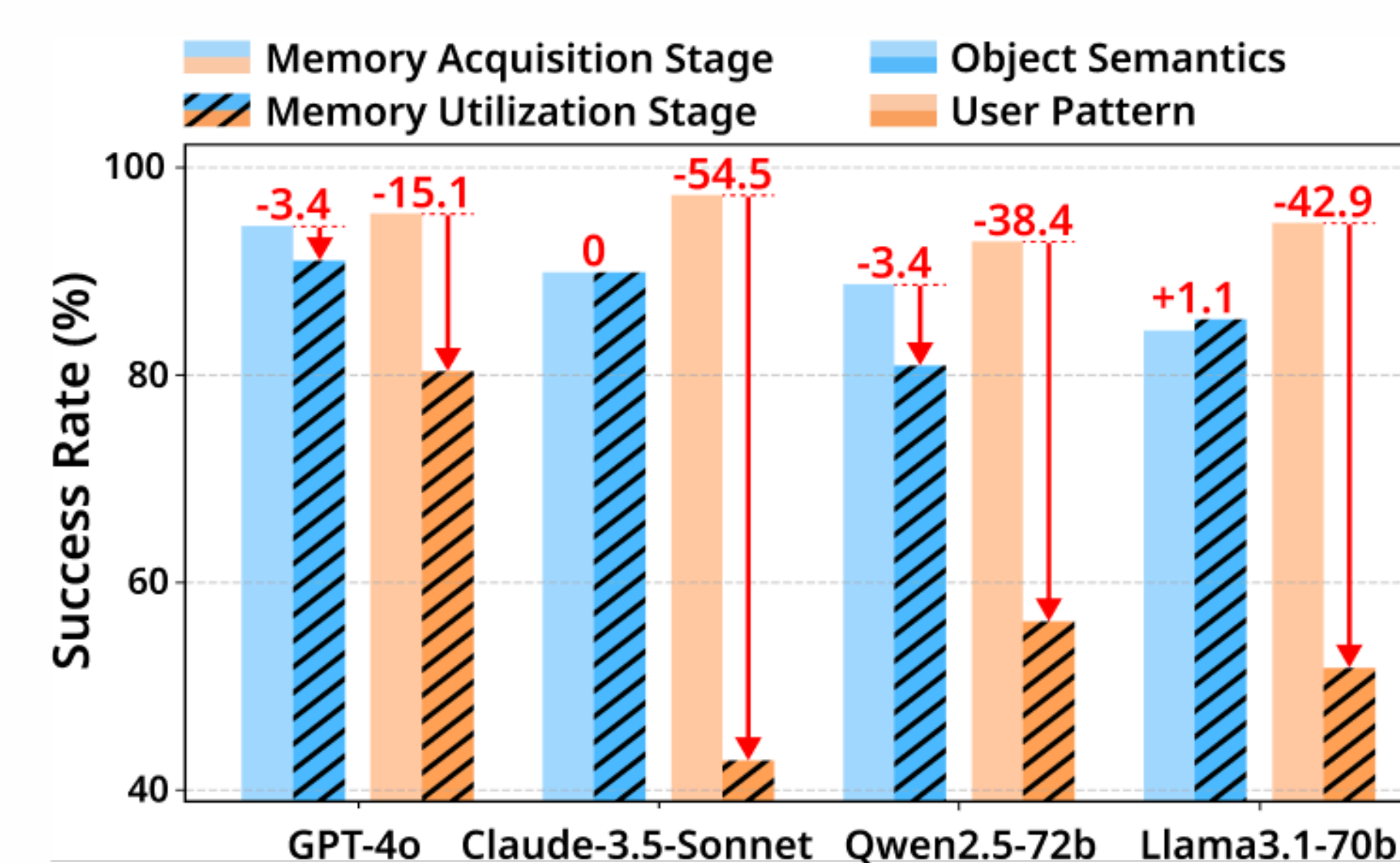


- Information overload** degrades memory utilization performance.
- Agents **fail to coordinate multiple memories**.

C Evaluating Memory Utilization Across Embodied Agents

► **RQ1. Can current LLM-powered embodied agents effectively utilize memory to perform personalized assistance tasks?**

Model	Stage	Task Type (Memory)	Planning Cycles ↓	Sim Steps ↓	Percent Complete ↑	Δ PC ↑	Success Rate ↑	Δ SR ↑
GPT-4o	Acquisition	-	16.5	2156.1	96.3	-	95.0	-
	Utilization	Single Joint	16.1 28.9	2450.8 3480.7	88.0 86.7	-8.3 -10.5	85.1 63.9	-9.9 -30.5
Claude-3.5-Sonnet	Acquisition	-	16.0	2104.1	96.2	-	94.0	-
	Utilization	Single Joint	15.3 27.8	2258.8 3198.8	69.3 64.6	-26.9 -30.1	63.7 33.3	-30.3 -57.0
Qwen-2.5-72b	Acquisition	-	17.5	2281.9	93.5	-	91.0	-
	Utilization	Single Joint	17.5 31.3	2691.2 4027.1	72.6 68.9	-20.9 -27.9	67.2 36.1	-23.8 -58.3
Llama-3.1-70b	Acquisition	-	17.7	2162.1	92.9	-	90.0	-
	Utilization	Single Joint	19.0 31.4	2566.6 3425.2	72.2 51.3	-20.7 -44.9	66.7 8.3	-23.3 -83.4



- LLM-powered embodied agents **struggle with understanding personalized knowledge**.
- Current embodied agents can effectively recall object semantics but **struggle to comprehend user patterns**.

E Exploring Memory Design for Personalized Embodied Agents

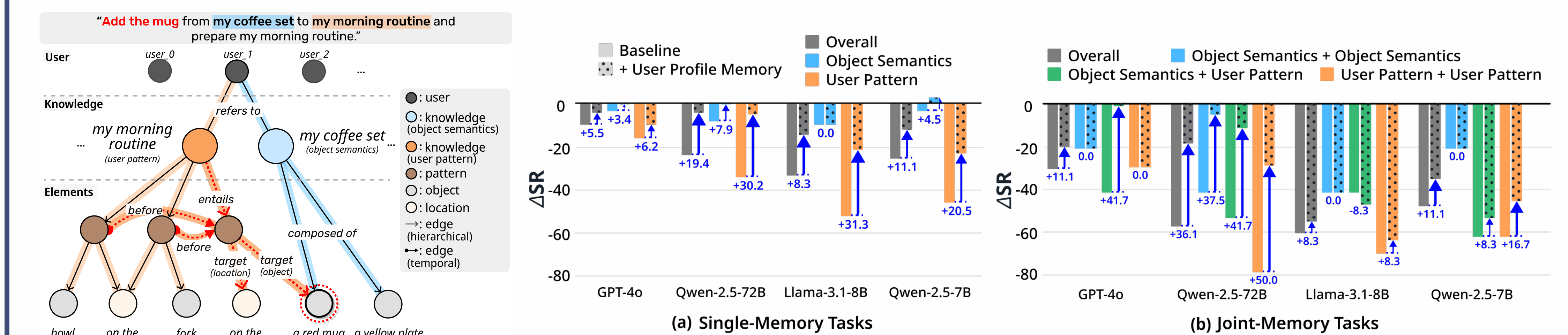
► **RQ3. How can we design memory architectures to better support personalized assistance tasks?**

- We compare three memory formats: **full episodic memory, summarized memory, and instruction-only** setup.
- Episodic memory provides both **personalized knowledge** and **in-context learning** benefits.

Model	Memory	PC (%)	SR (%)
GPT-4o	Full Memory	90.0	83.3
	Summarization	88.0	83.3
	Instruction-only	62.4	50.0
Qwen-2.5-72b	Full Memory	77.2	66.7
	Summarization	77.4	70.0
	Instruction-only	51.3	40.0
Llama-3.1-8b	Full Memory	72.8	63.3
	Summarization	49.4	43.3
	Instruction-only	40.0	30.0
Qwen-2.5-7b	Full Memory	50.1	43.3
	Summarization	43.9	36.7
	Instruction-only	35.6	23.3

User Profile Memory:

Hierarchical knowledge graph for personalized knowledge



- We construct **user profile memory** as a **three-level knowledge graph** with **hierarchical edges** for structure and **temporal edges** for sequential ordering.
- User profile memory** enhances embodied agents' memory utilization capability.