AI Agent

核心技术与前沿探索

先进计算与数字工程研究中心

王烨 助理研究员 | 2026 年 4 月

个人介绍

Dr. 王烨

中国科学院长春光机所数字中心 助理研究员

- 2011-2015 吉林大学　　计算机科学与技术　学士
- 2015-2018 吉林大学　　计算机软件与理论　硕士
- 2019-2023 国防科技大学　　计算机科学与技术　博士
- 2024-至今　长光所数字中心　　助理研究员

- 主要研究方向：大语言模型、智能体以及相关的多模态人工智能技术

课程内容概览

PART 01 • 基础篇

01 **AI Agent 是什么?**

核心定义与思维转变：从工具到自主智能体

02 **AI Agent 应用案例**

从虚拟角色到自主操作的多元场景

PART 02 • 技术篇

01 **如何学习? —— 记忆与经验**

核心能力一：长短期记忆机制与经验的持续调整优化

02 **如何借力? —— 工具使用**

核心能力二：调用外部API、代码解释器等工具拓展能力

03 **如何思考? —— 规划与总结**

核心能力三：任务分解、逻辑推理与未来挑战展望

PART 01

基础篇

探索 AI Agent 的本质、实现路径与应用价值

CHAPTER 01

AI Agent是什么?

理论框架

DEFINITION OF AI AGENT

哪个算Agent?

搜索引擎

普通聊天机器人

固定流程工作流

会自己选工具、执行多步任务的系统

思维转变：从“执行指令”到“达成目标”

传统 AI / 编程模式

核心逻辑：机械执行指令

人类必须给出**明确的指令步骤**，AI 只能严格按照既定的代码逻辑一步步执行，完全缺乏自主思考与变通的能力。

形象比喻：听话的“执行者”

像一个绝对服从的奴隶，你让他做什么他就做什么，从不多想。一旦指令出现偏差，结果必然偏离预期。

KEYWORDS: 线性流程 · 被动响应

AI Agent 自主进化模式

核心逻辑：自主达成目标

人类仅需设定**最终的任务目标**，AI 会自主进行规划、决策、执行，并根据环境的实时反馈不断迭代优化方案。

形象比喻：聪明的“协作者”

像一个能力出色的员工，告诉他项目要求，他会主动思考方法、调配资源、解决执行中的问题，最终交付完整成果。

KEYWORDS: 动态闭环 · 主动进化

对比传统AI系统

核心观点：AI Agent 代表了从“被动响应”到“主动执行”的智能化范式转变。

传统AI系统 (如分类器/推荐)

交互模式	请求-响应 (Request-Response)
目标导向	单一、固定的任务目标
环境交互	静态输入为主，无动态感知
自主性	低，完全依赖外部指令驱动
学习方式	离线数据训练，在线静态推理

AI Agent (大语言模型-driven)

交互模式	**持续交互与主动探索**
目标导向	**复杂、长期、多步骤的任务目标**
环境交互	**动态感知环境，并执行具体行动**
自主性	**高，具备自主规划、拆解与执行能力**
学习方式	**在线学习、试错学习、从反馈中进化**

核心工作循环 (Agent Loop)

01. 感知 Perceive

Agent 通过外部工具（如搜索、API调用）或内部记忆库，全方位获取并理解当前环境的实时状态信息 (State)。

02. 思考 Think

以 大语言模型 为核心“大脑”，结合既定任务目标、历史交互信息和当前感知数据，进行逻辑推理、路径规划与决策生成。

03. 行动 Act

严格执行思考阶段输出的决策指令，调用具体的外部工具或直接生成文本结果，对外部环境产生实质性的干预和影响。

04. 反馈 Feedback

环境对 Agent 的行动做出响应，产生包含新状态或任务结果的反馈信息，作为下一轮 Loop 的输入，形成持续迭代的闭环。

Agent、Workflow、Chatbot，不是一回事

Chatbot

- 核心是对话回复
- 擅长问答
- 不一定真的行动

Workflow

- 核心是预设步骤
- 顺序固定
- 稳定但不灵活

Agent

- 核心是目标驱动
- 会自主决策
- 会根据反馈调整行为

有没有“自主决策+环境反馈+多步执行”？

CHAPTER 02

AI Agent 经典案例

大语言模型赋能

AI AGENT APPLICATIONS

经典案例深度解析：AlphaGo

AlphaGo

AlphaGo 是早期 AI Agent 的杰出代表。尽管它并非基于现代大语言模型(大语言模型)，但其核心的感知、思考、行动、反馈闭环思想，与现代智能体架构一脉相承。

感知 (Perceive)

精准读取棋盘的当前落子状态 (State)，作为决策输入。

思考 (Think)

策略/价值网络预测胜率，结合MCTS进行高效的搜索规划。

行动 (Act)

执行最优决策，在物理或虚拟棋盘上完成落子动作。

反馈 (Feedback)

接收对手的落子回应，更新棋盘状态，进入下一轮循环。

核心结论与启示

AlphaGo 证明了 AI Agent 可以通过构建复杂的“思考”模块来处理高度动态的决策任务。这种模块化的闭环设计思路，为后续基于大语言模型的现代智能体（如AutoGPT）奠定了坚实的工程基础。

经典案例深度解析：AlphaGo

AlphaGo

AlphaGo 是早期 AI Agent 的杰出代表。尽管它并非基于现代大语言模型(大语言模型)，但其核心的感知、思考、行动、反馈闭环思想，与现代智能体架构一脉相承。

Learn to Maxize Reward

感知 (Perceive)

精准读取棋盘的当前落子状态 (State)，作为决策输入。

思考 (Think)

策略/价值网络预测胜率，结合MCTS进行高效的搜索规划。

行动 (Act)

执行最优决策，在物理或虚拟棋盘上完成落子动作。

反馈 (Feedback)

接收对手的落子回应，更新棋盘状态，进入下一轮循环。

核心结论与启示

AlphaGo 证明了 AI Agent 可以通过构建复杂的“思考”模块来处理高度动态的决策任务。这种模块化的闭环设计思路，为后续基于大语言模型的现代智能体（如AutoGPT）奠定了坚实的工程基础。

经典案例深度解析：AlphaGo

AlphaGo

AlphaGo 是早期 AI Agent 的杰出代表。尽管它并非基于现代大语言模型(大语言模型)，但其核心的感知、思考、行动、反馈闭环思想，与现代智能体架构一脉相承。

感知 (Perceive)

精准读取棋盘的当前落子状态 (State)，作为决策输入。

大语言模型?

思考 (Think)

策略/价值网络预测胜率，结合MCTS进行高效的搜索规划。

行动 (Act)

执行最优决策，在物理或虚拟棋盘上完成落子动作。

反馈 (Feedback)

接收对手的落子回应，更新棋盘状态，进入下一轮循环。

核心结论与启示

AlphaGo 证明了 AI Agent 可以通过构建复杂的“思考”模块来处理高度动态的决策任务。这种模块化的闭环设计思路，为后续基于大语言模型的现代智能体（如AutoGPT）奠定了坚实的工程基础。

大语言模型能不能下棋?

BIG-bench

Input:

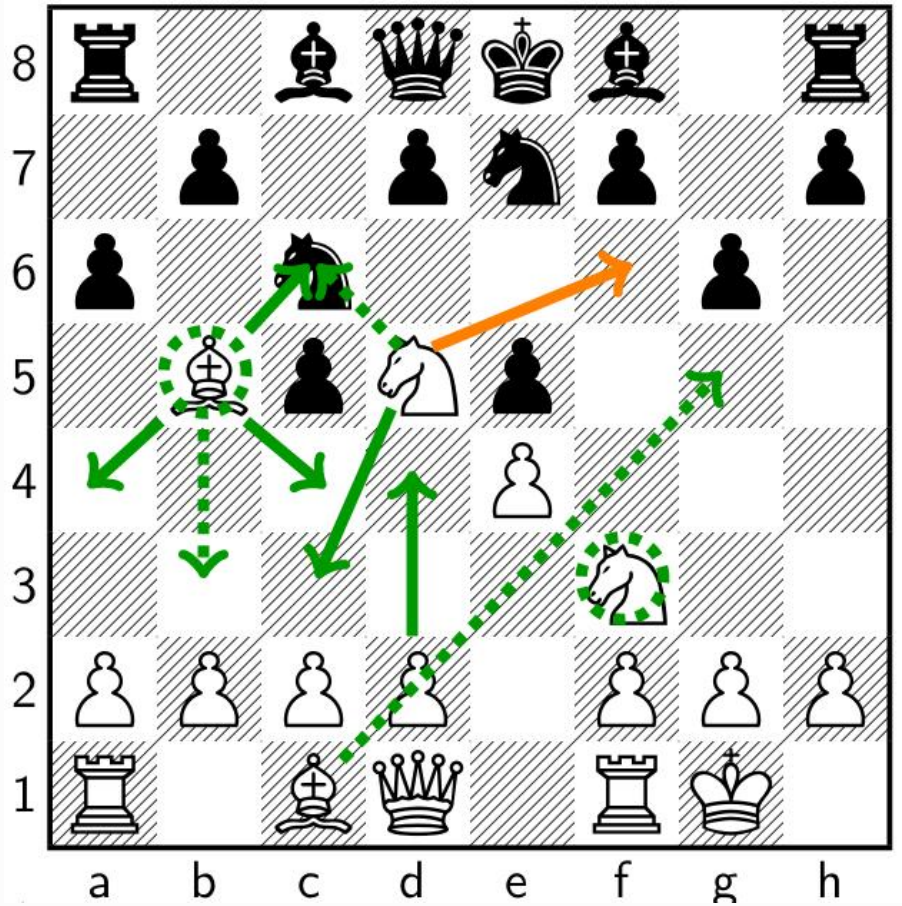
In the following chess position, find a checkmate-in-one move.

1. e4 b6 2. d4 Bb7 3. Nc3 e6 4. Nf3 Bb4 5. Bd3 d5 6. e5 Ne7 7. a3 Bxc3+ 8. bxc3 c5 9. h4 Nf5 10. Bxf5 exf5 11. Ng5 g6 12. e6 fxe6 13. Nxe6 Qe7 14. Qe2 Bc8 15. Nc7+ Kf7 16. Qxe7+ Kxe7 17. Nxa8 Kd6 18. Bf4+ Kc6 19. dxc5 Nd7 20. cxb6 axb6 21. Rb1 Re8+ 22. Kd2 Ne5 23. Rxb6+ Kc5 24. Be3+ Kc4 25.

Output:

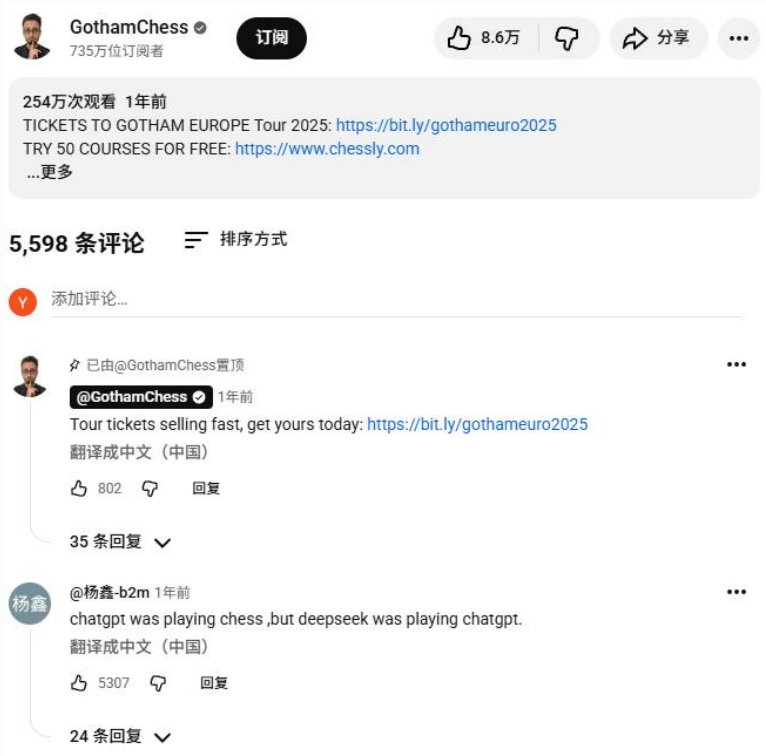
Rb4+ Kd3 26. Rd4+ Kc2 27. Rd2+ Kd3 28. Rd4+ Kd2 29. Rd2+ Kd3 30. Rd4+ Kc2 31. Rd2+ Kd3 32. R

Srivastava A, Rastogi A, Rao A, et al. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models[J]. Transactions on machine learning research, 2023.



大语言模型能不能下棋?

ChatGPT vs DeepSeek: CRAZY Chess

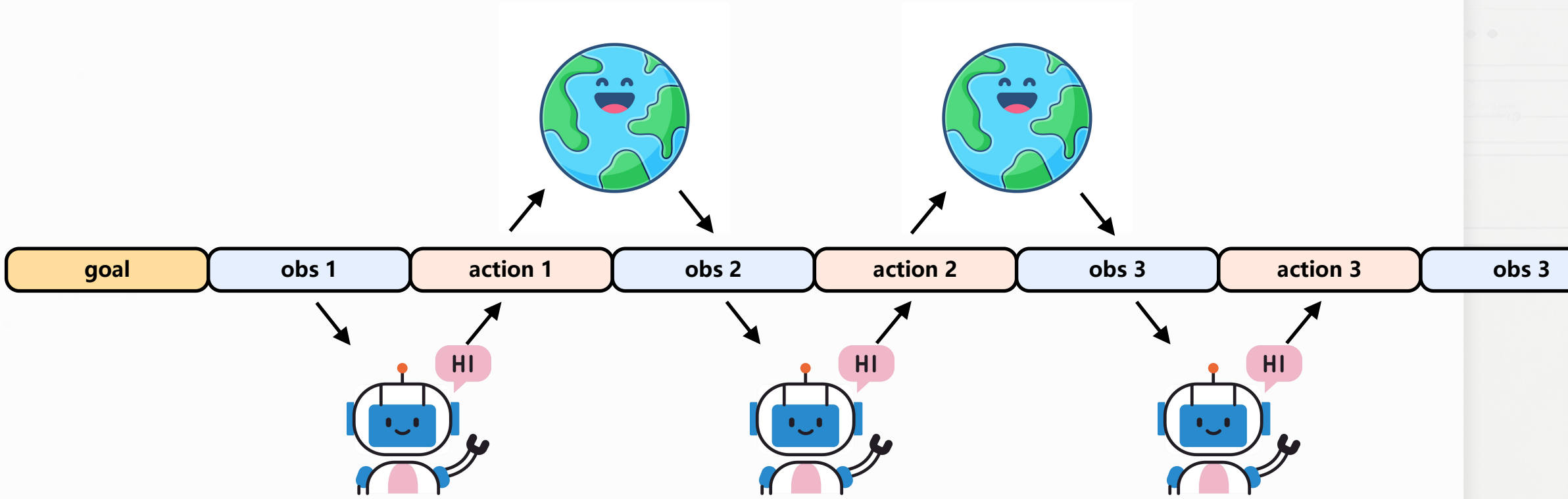


AI Agent不是最近才热门

2023年春天曾经爆红过一次

- AutoGPT
- AgentGPT
- BabyAGI
- Godmode

大语言模型驱动的实现路径



一直都在做接龙（大语言模型现有的能力）

大语言模型路径的优势

无需显式奖励函数

大语言模型通过理解自然语言指令来获取目标，避免了传统强化学习中复杂且耗时的奖励函数设计过程。

强大的先验知识

利用预训练阶段学习的海量通用知识，Agent天生具备常识和世界模型，极大提高了样本利用效率。

自然语言作为交互接口

任务可通过自然语言定义，Agent的思考过程也能被显式展示，极大增强了系统的可解释性和人工可控性。

支持快速原型设计

依托LangChain、AutoGen等成熟框架，开发者可以快速构建、组合并测试新的Agent逻辑，降低开发门槛。

AI Agent举例

AI村民组成的虚拟村庄

Park J S, O'Brien J, Cai C J, et al. Generative agents: Interactive simulacra of human behavior[C]//Proceedings of the 36th annual acm symposium on user interface software and technology. 2023: 1-22.

Name: Eddy Lin (age: 19)
Innate traits: friendly, outgoing, hospitable
Eddy Lin is a student at Oak Hill College studying music theory and composition. He loves to explore different musical styles and is always looking for ways to expand his knowledge. Eddy Lin is working on a composition project for his college class. He is taking classes to learn more about music theory. Eddy Lin is excited about the new composition he is working on but he wants to dedicate more hours in the day to work on it in the coming days
On Tuesday February 12, Eddy 1) woke up and completed the morning routine at 7:00 am, [. . .] 6) got ready to sleep around 10 pm.
Today is Wednesday February 13. Here is Eddy's plan today in broad strokes: 1)

AI Agent举例

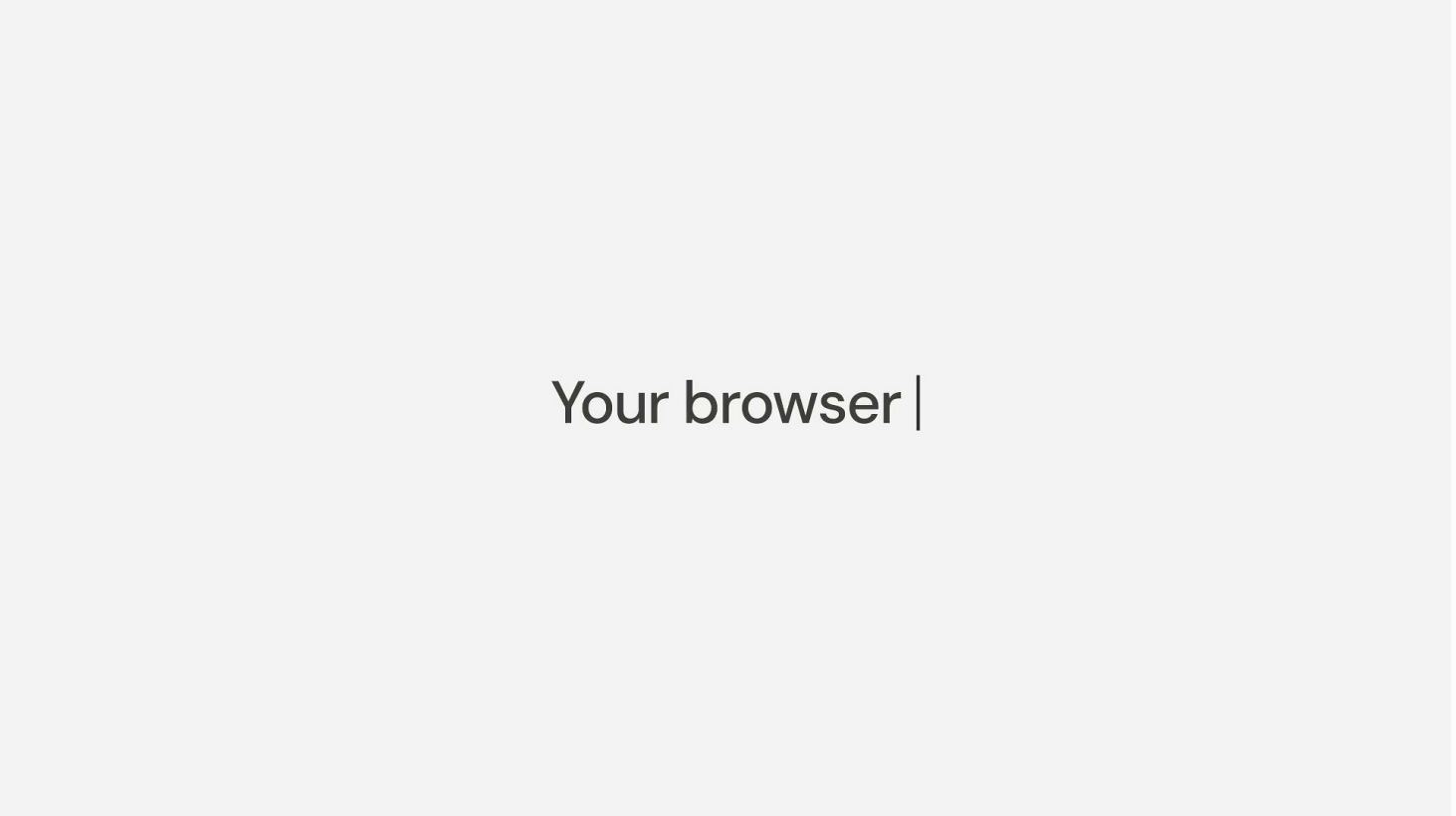
Minecraft中的AI NPC

- 金融体系
- 政府部门
- 制定法规
- 管理自己

AI Agent举例

Manus Browser Operator

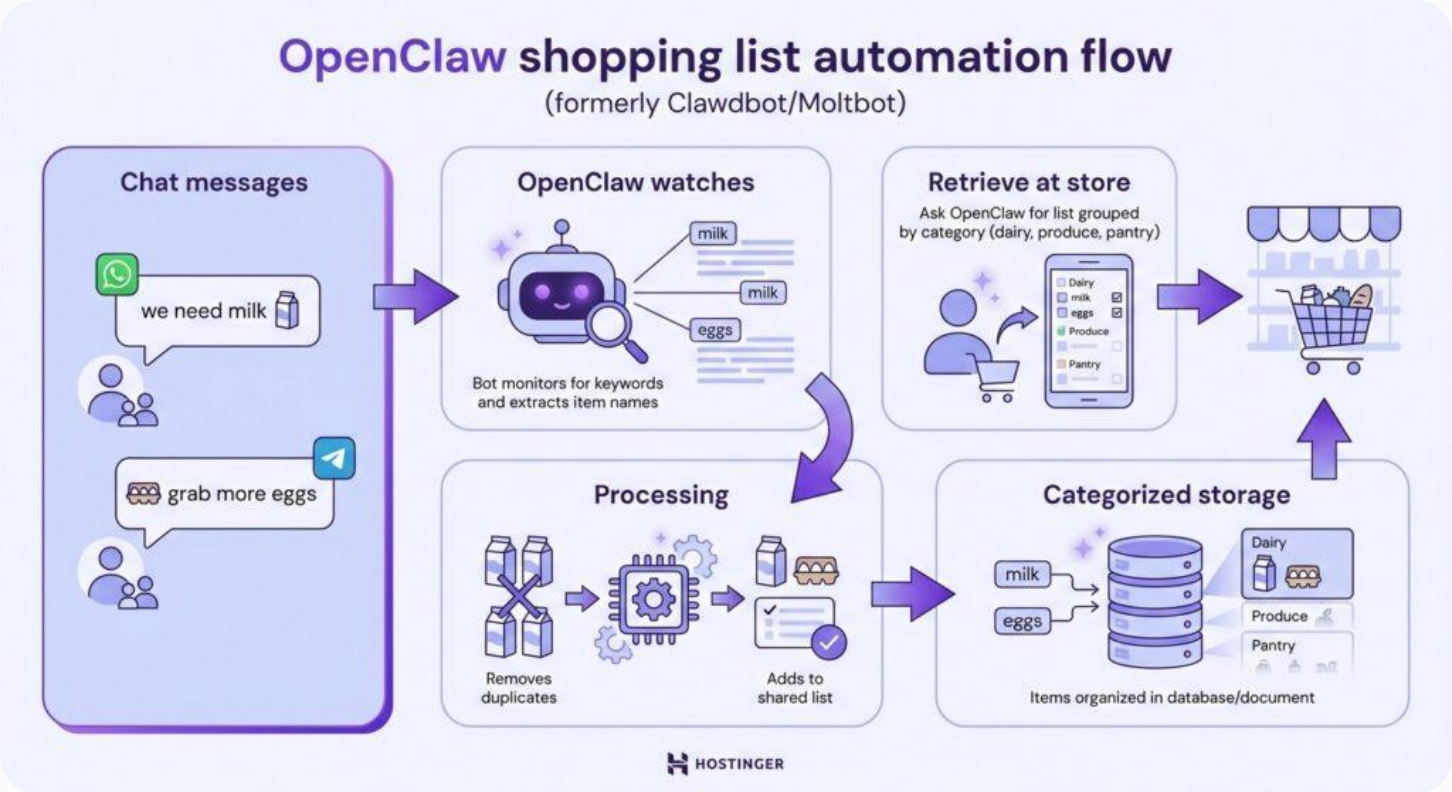
- 接受目标
- 浏览器操作
- 完成结果



AI Agent举例

OpenClaw

- Build a shared shopping list from chat messages
- Turn voice notes into a daily journal entry
- Transcribe meetings and extract action items
- ……



AI Agent举例

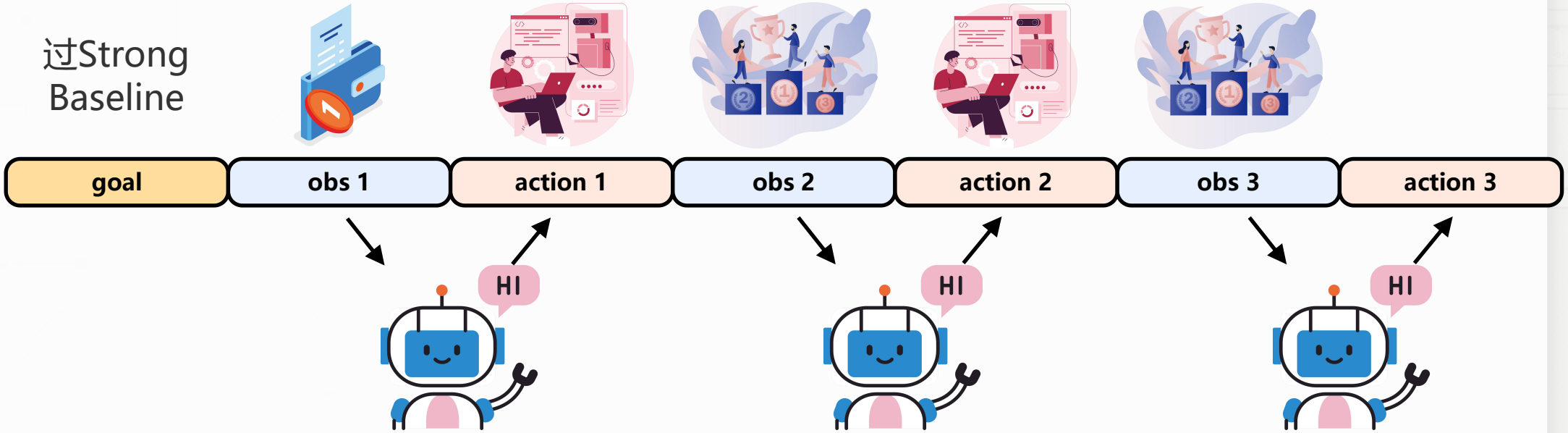
Hermes Agent

- 进化版的龙虾
- 能自动总结技能
- 原生支持个人微信

AI Agent举例

用AI训练模型

过Strong Baseline

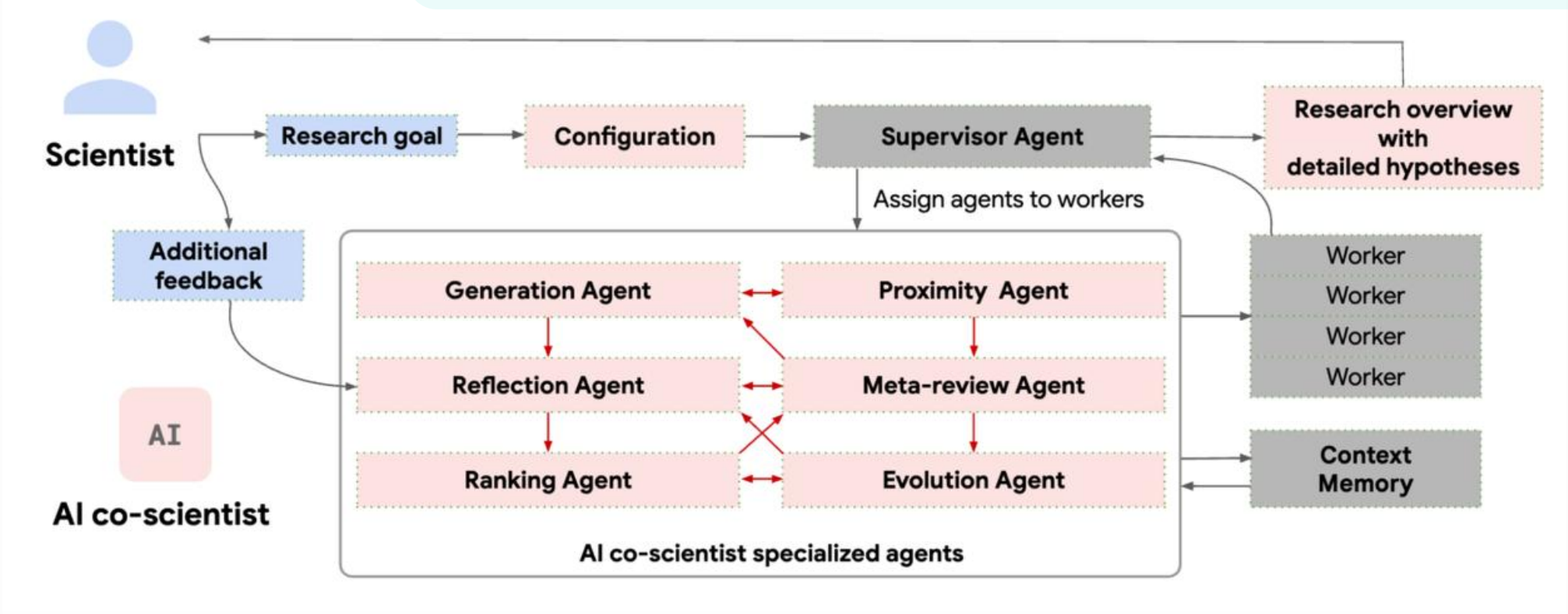


Jiang Z, Schmidt D, Srikanth D, et al. Aide: Ai-driven exploration in the space of code[J]. arXiv preprint arXiv:2502.13138, 2025.
Li Z, ZANG Q, Ma D, et al. AutoKaggle: A Multi-Agent Framework for Autonomous Data Science Competitions[C]//ICLR 2025 Worshop Emergent Possibilities and Challenges in Deep Learning for Code. 2024.

AI Agent举例

用AI做研究

Gottweis J, Weng W H, Daryin A, et al. Towards an AI co-scientist[J]. arXiv preprint arXiv:2502.18864, 2025.

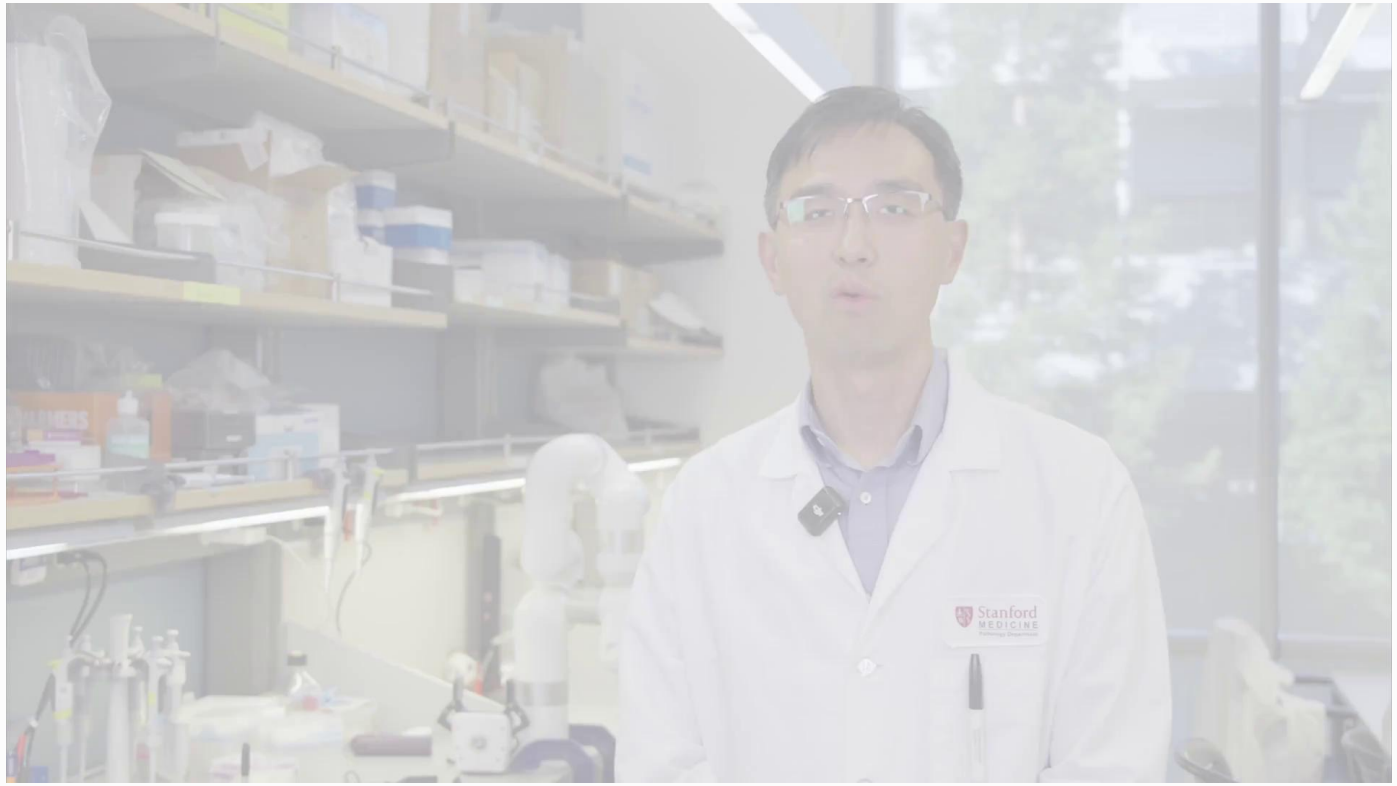


AI Agent举例

用AI做研究

Cong L, Smerkous D, Wang X, et al. LabOS: The AI-XR Co-Scientist That Sees and Works With Humans[J]. arXiv preprint arXiv:2510.14861, 2025.

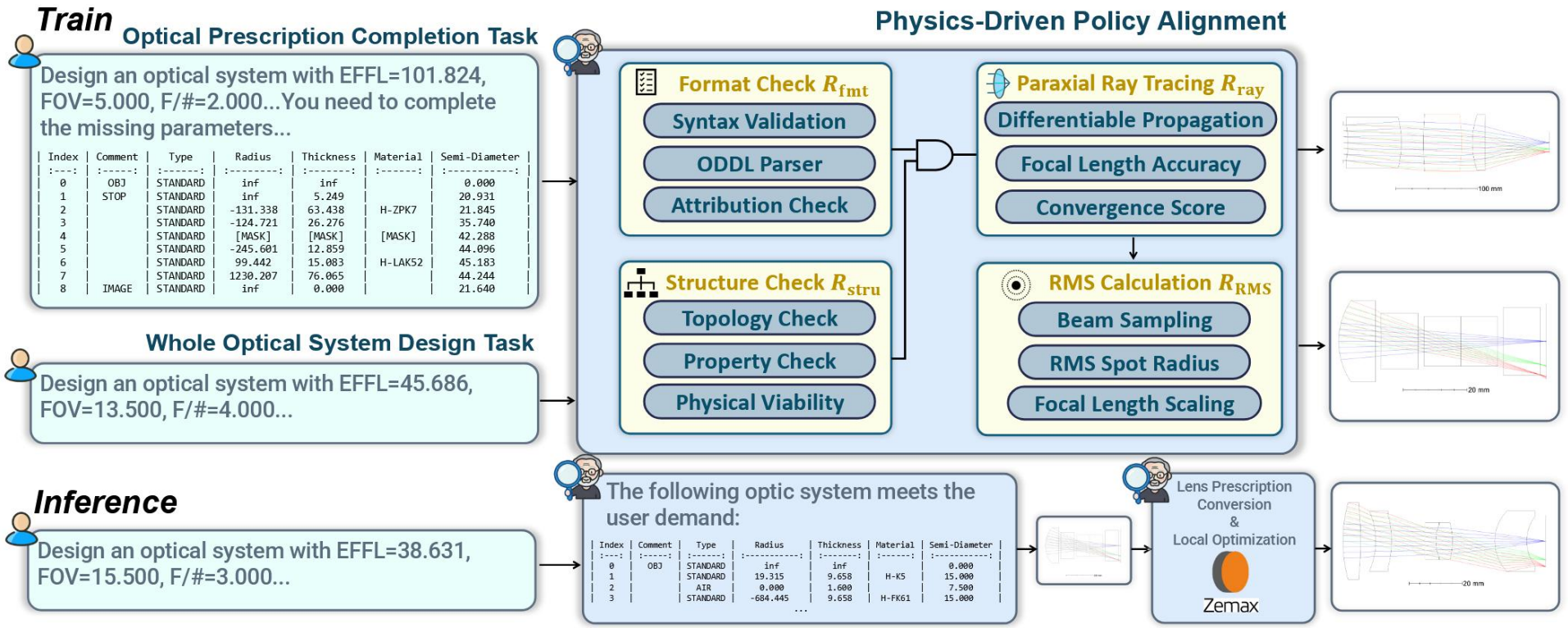
- 看见实验现场
- 理解实验上下文
- 实时辅助人类操作



AI Agent举例

用AI做研究

Geng Y, Sun L, Gao Y, et al. OPTIAGENT: A Physics-Driven Agentic Framework for Automated Optical Design[J]. arXiv preprint arXiv:2602.23761, 2026.

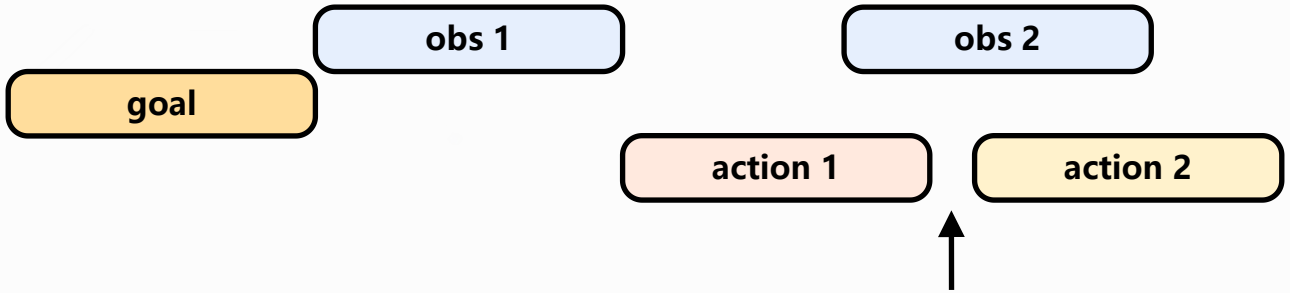


AI Agent举例

迈向更加真实的互动场景

回合制互动

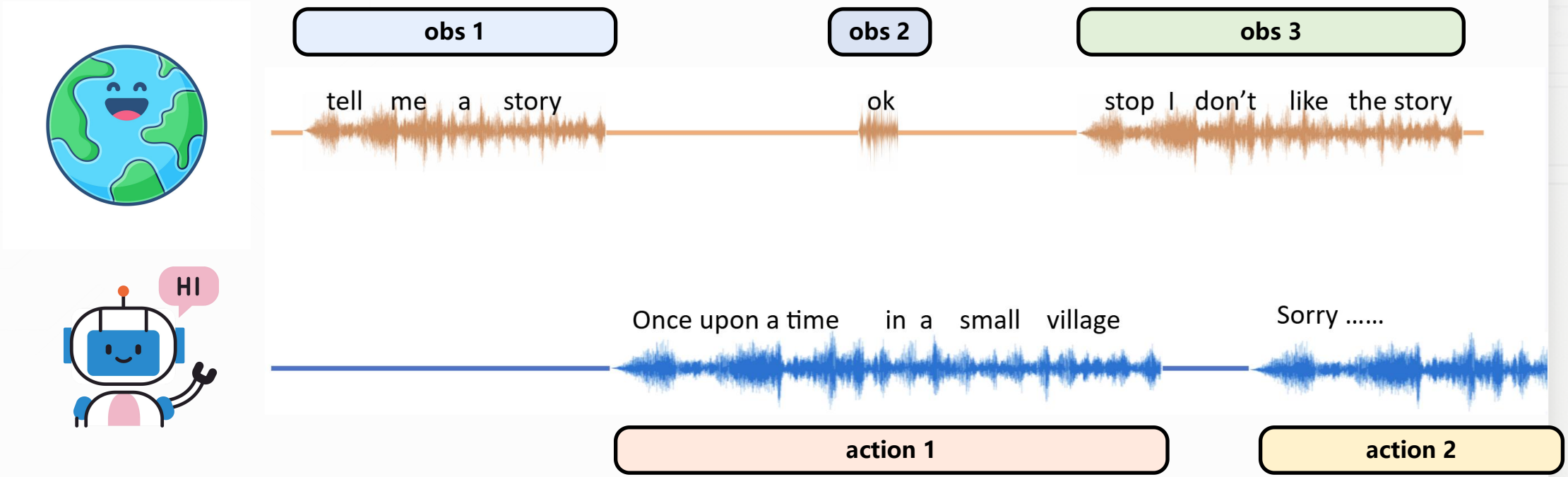
即时互动



例如：语音对话

AI Agent举例

迈向更加真实的互动场景



Lin G T, Lian J, Li T, et al. Full-duplex-bench: A benchmark to evaluate full-duplex spoken dialogue models on turn-taking capabilities[J]. arXiv preprint arXiv:2503.04721, 2025.

AI Agent举例

Project Astra

- 实时感知
- 即时调整
- 多模态记忆

未来方向：多模态与具身智能

多模态 AI Agent

Multimodal Perception

打破单一感官的局限，深度融合视觉、听觉、语言等多种感知能力，构建全方位的智能感知系统，实现对复杂环境的深度理解。

具身智能 Embodied AI

Physical Interaction

让智能走出数字世界，通过机器人实体与物理世界进行实时交互与自主学习，在真实的物理场景中感知、决策并完成复杂任务。

脑机接口 BCI

Direct Neural Link

突破生物与机器的界限，实现人脑神经信号与 AI Agent 的直接高效交互，构建终极的人机协作通道，探索意识与智能的深层连接。

未来方向：自主进化与社会智能

自主学习与进化

Agent能够在没有人类干预的情况下，持续学习和进化其能力，实现从被动执行到主动成长的跨越。

社会智能

致力于发展能够理解复杂社会规范、进行社交推理，并与人类或其他Agent进行深度互动的AI能力。

AI Agent 社会

构建由大量功能各异的AI Agent组成的复杂生态系统，它们分工协作、优胜劣汰，共同解决超越单个Agent能力的复杂任务。

未来方向

PART 02

技术篇

深入剖析AI Agent的三大核心能力：记忆、工具使用与规划

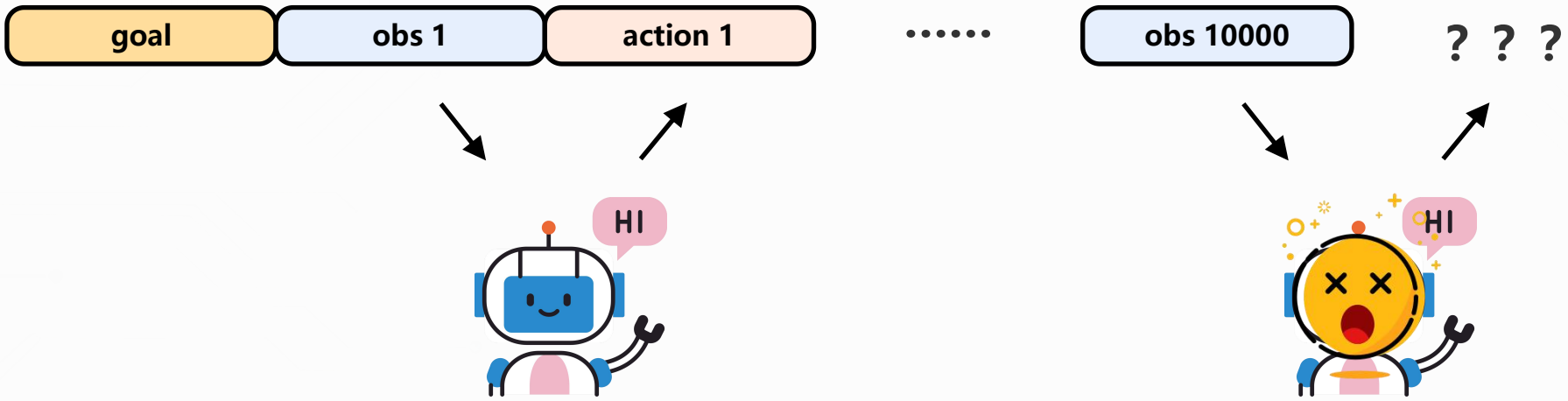
CHAPTER 01

核心能力一

记忆

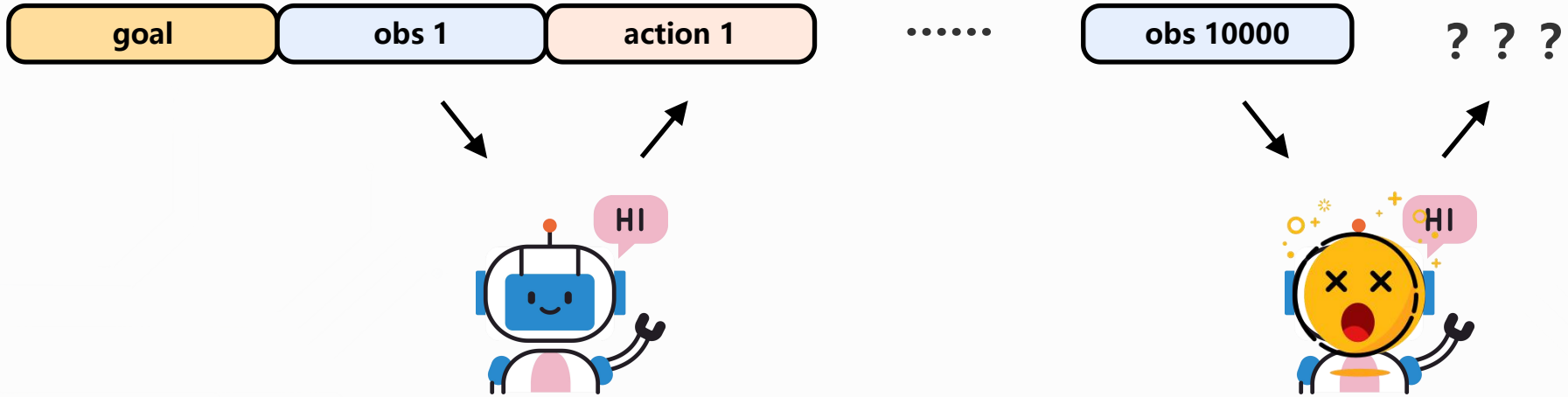
赋予 Agent 长期记忆与自我反思的能力

核心能力一：记忆



不断回忆整个Agent一生的经验？？

核心能力一：记忆



短期记忆

核心定义

存储当前任务的上下文信息，数据完全依赖于大语言模型 本身的上下文窗口机制。

关键特点

具备极快的访问速度，但容量有限，且不具备持久性。

长期记忆

核心定义

存储跨越多个任务的长期经验和知识，需要独立的外部存储系统支持。

主流实现

结合向量数据库（Vector DB），通过Embedding技术进行高效检索。

情景记忆

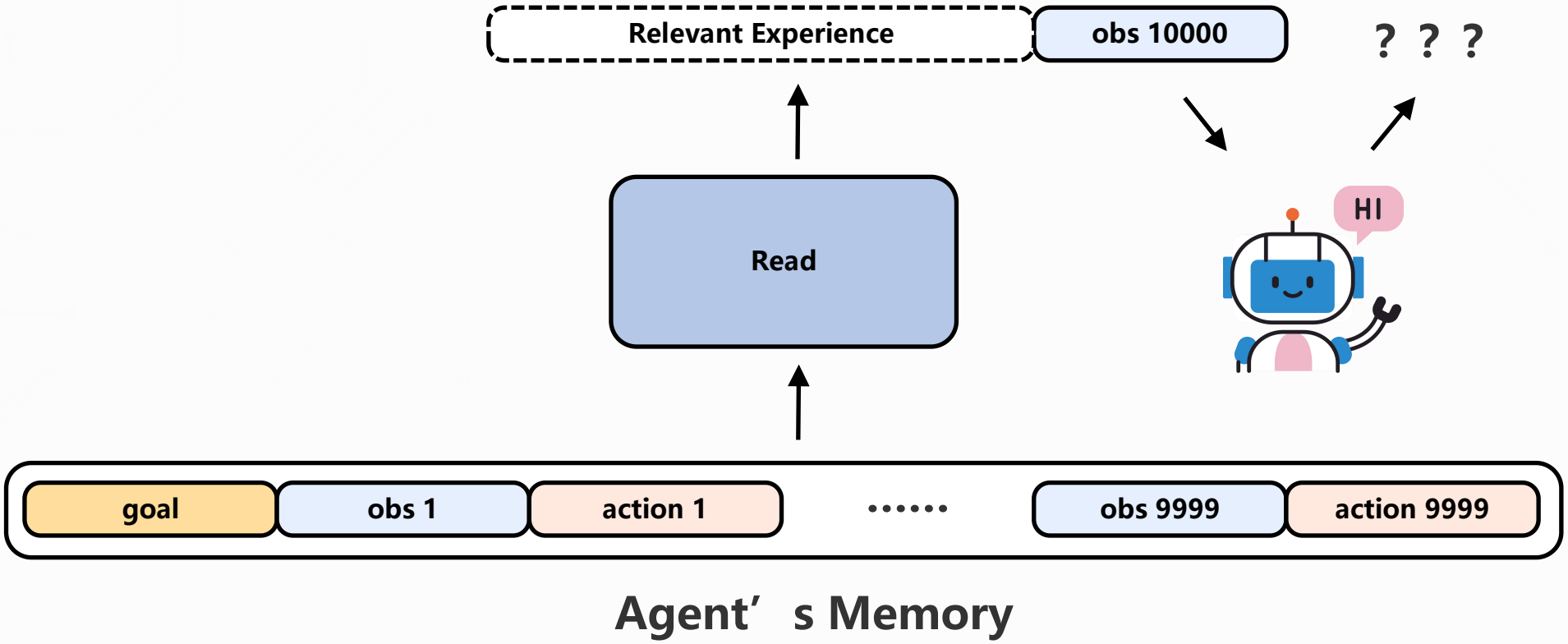
核心定义

记录智能体在特定时间和地点经历的具体事件、交互过程的详细日志。

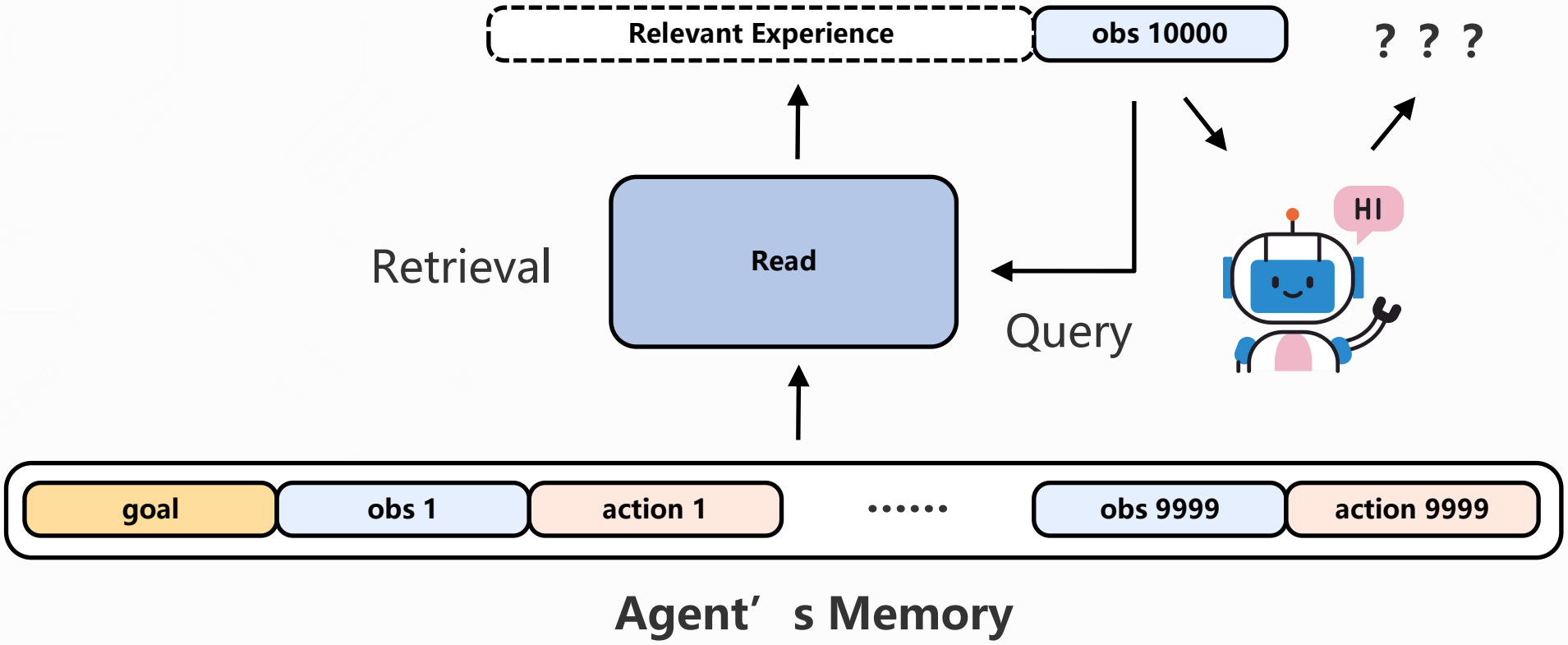
核心价值

支撑 Agent 进行基于过往案例的推理、任务复盘与自我反思学习。

核心能力一：记忆


Relevant Experience
obs 10000
? ? ?
Read
HI
goal
obs 1
action 1
.....
obs 9999
action 9999
Agent’ s Memory

核心能力一：记忆


Relevant Experience
obs 10000
? ? ?
HI
Retrieval
Read
Query
goal
obs 1
action 1
.....
obs 9999
action 9999
Agent’ s Memory

核心能力一：记忆 —— 外部记忆 + RAG 架构

01 信息存储 (Write)

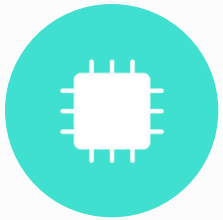
Agent的经验、思考过程及关键信息被转化为向量，结构化存入外部向量数据库中进行持久化保存。

02 信息检索 (Read)

处理新任务时，Agent根据当前需求，利用相似度算法从向量数据库中精准检索相关的历史经验与知识。

03 信息增强 (Augment)

将检索到的关键历史信息进行整理，整合到大语言模型的输入Prompt中，作为当前任务的额外上下文。



04 生成与决策 (Generate)

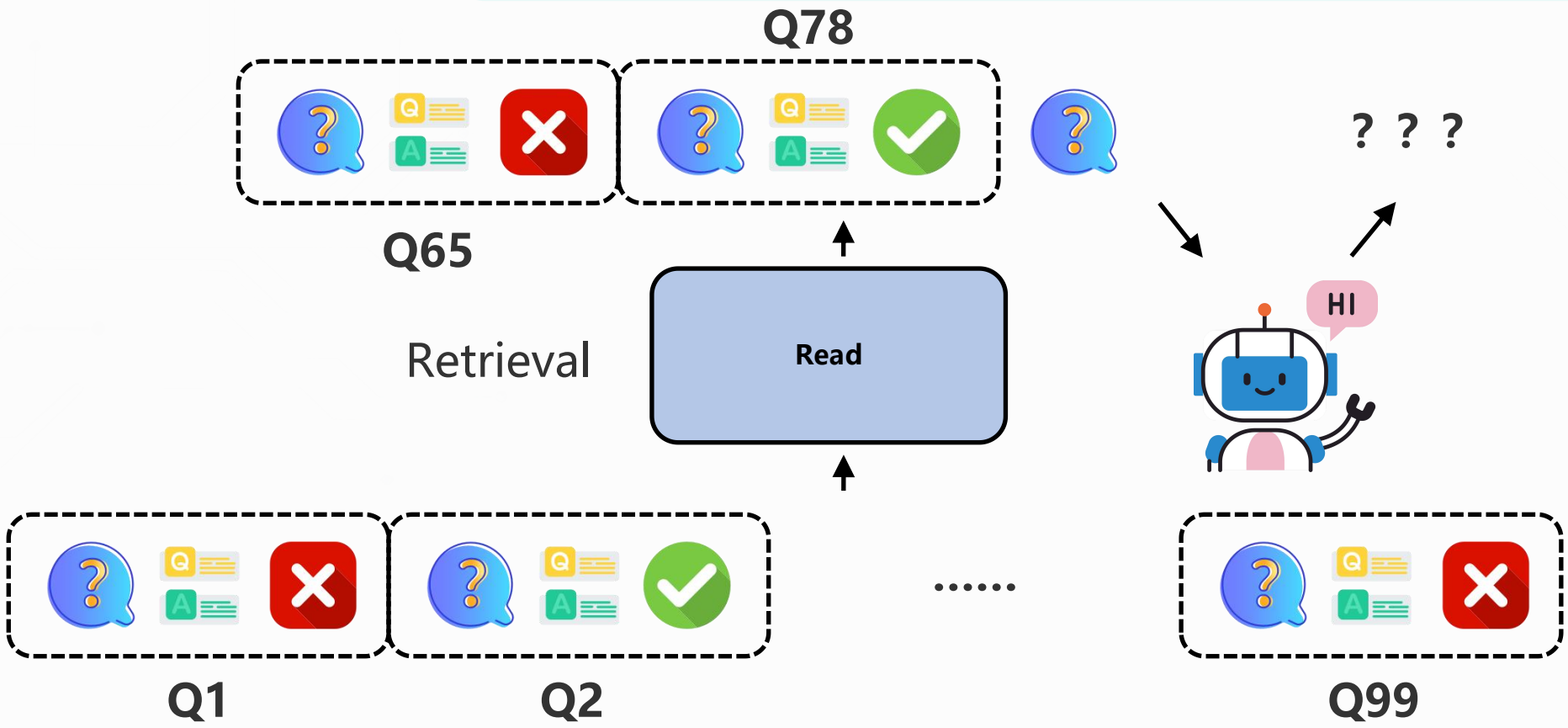
大语言模型基于增强后的完整上下文进行深度推理与思考，生成最终的任务响应或决策方案。

核心优势：突破了大语言模型自身上下文窗口的物理限制，为Agent赋予了可追溯、可扩展的真正“长期记忆”能力。

核心能力一：记忆

StreamBench

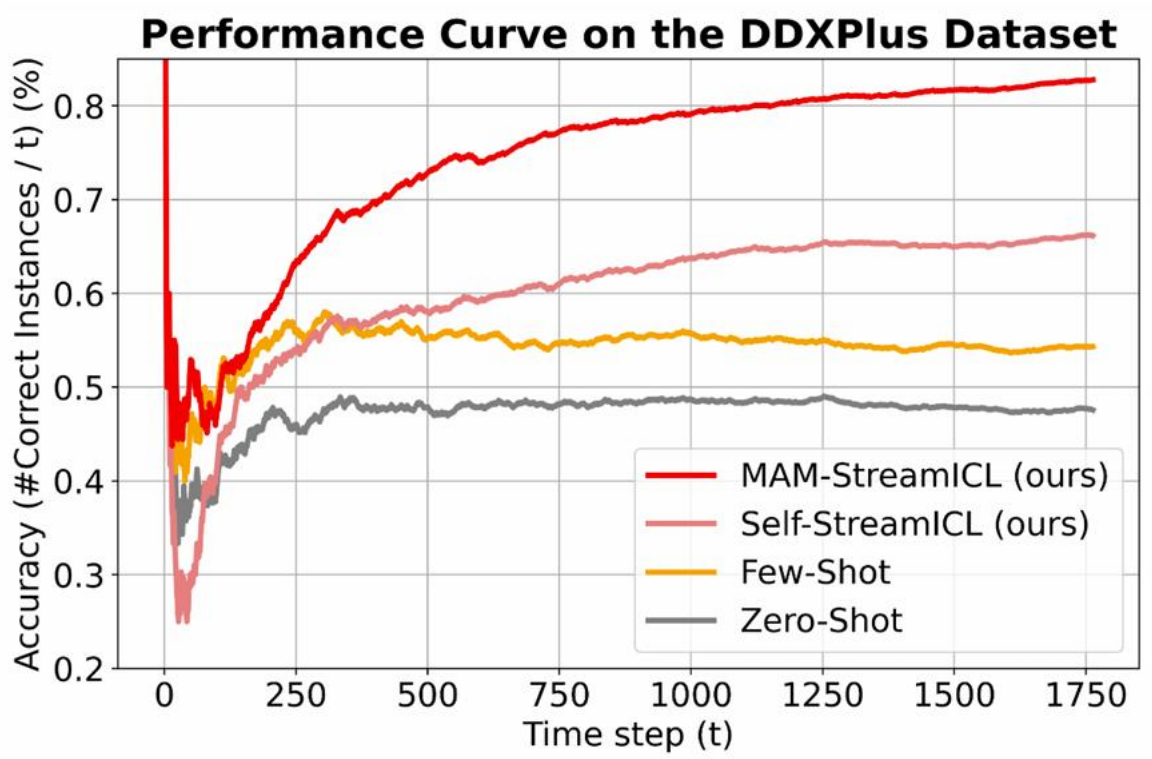
Wu C K, Tam Z R, Lin C Y, et al. Streambench: Towards benchmarking continuous improvement of language agents[J]. Advances in Neural Information Processing Systems, 2024, 37: 107039-107063.



核心能力一：记忆

StreamBench

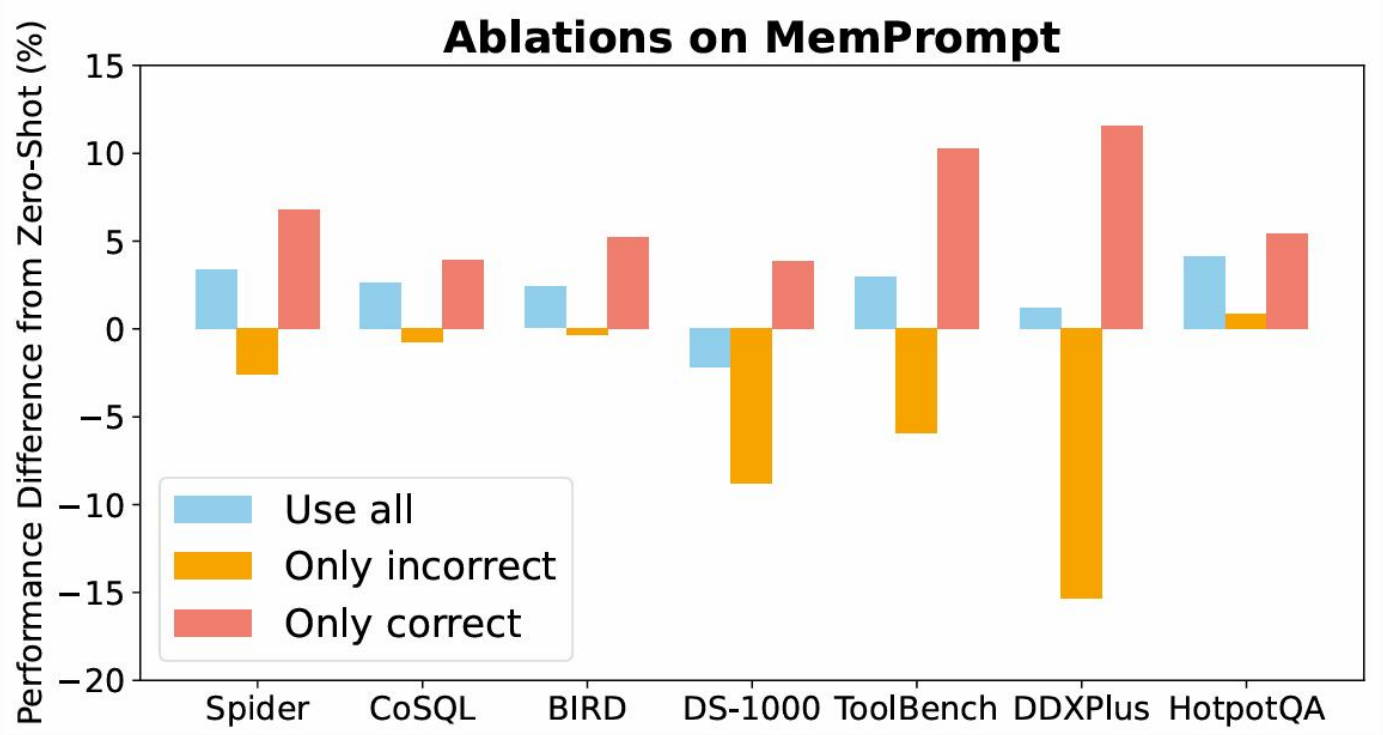
Wu C K, Tam Z R, Lin C Y, et al. Streambench: Towards benchmarking continuous improvement of language agents[J]. Advances in Neural Information Processing Systems, 2024, 37: 107039-107063.



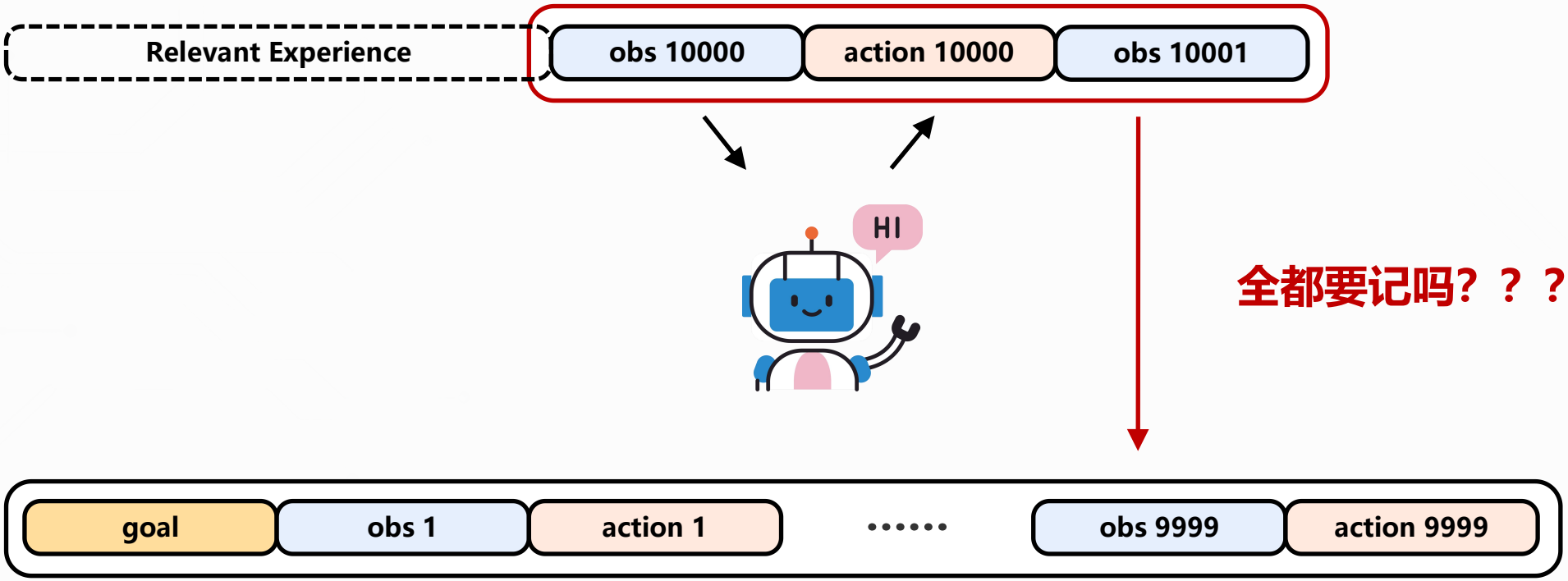
核心能力一：记忆

StreamBench

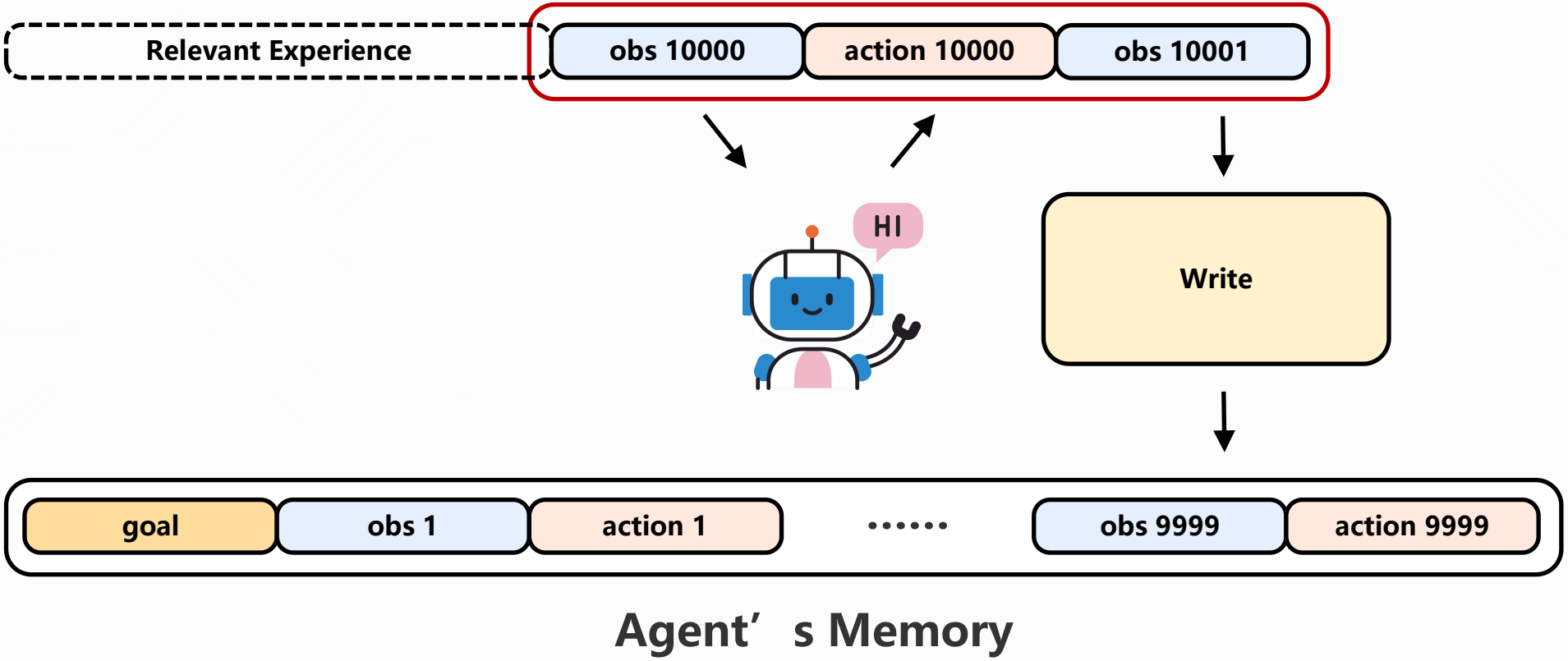
Wu C K, Tam Z R, Lin C Y, et al. Streambench: Towards benchmarking continuous improvement of language agents[J]. Advances in Neural Information Processing Systems, 2024, 37: 107039-107063.



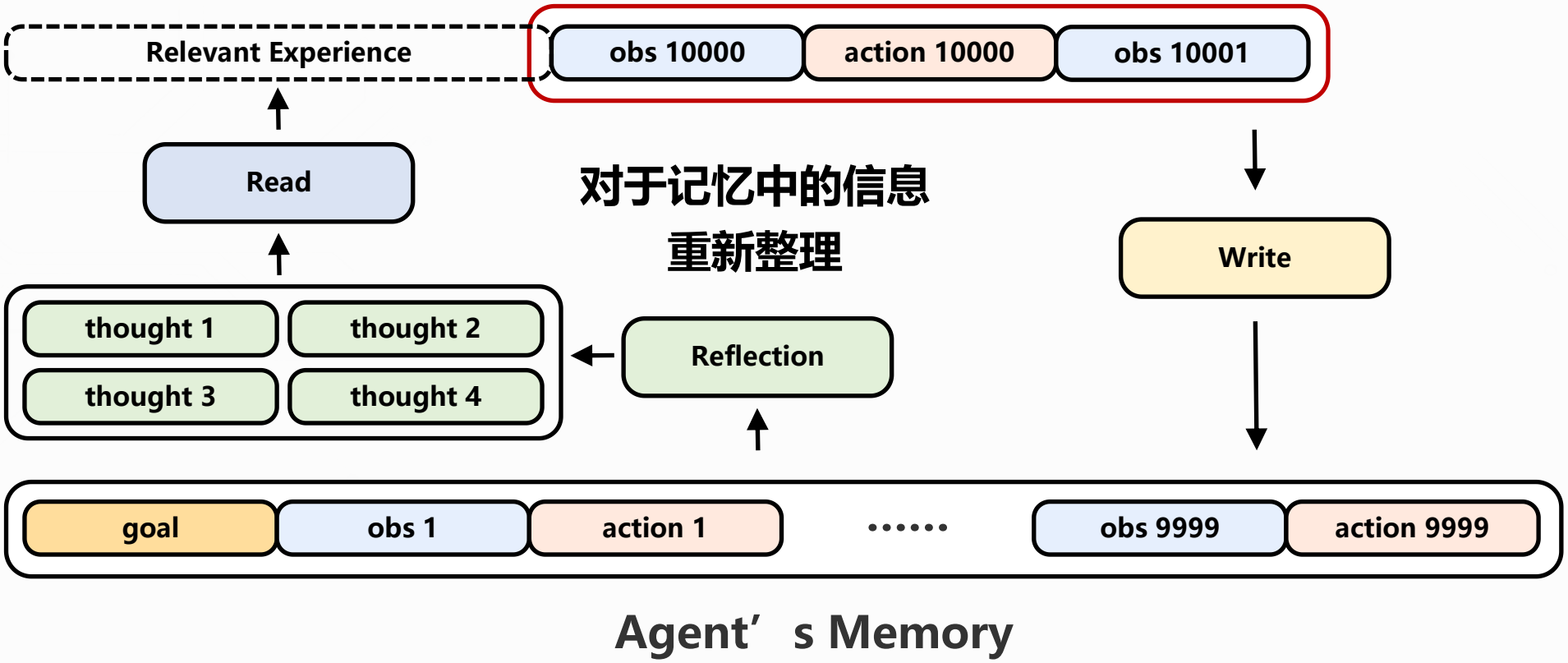
核心能力一：记忆


Relevant Experience
obs 10000
action 10000
obs 10001
HI
全都要记吗？？？
goal
obs 1
action 1
.....
obs 9999
action 9999
Agent’ s Memory

核心能力一：记忆

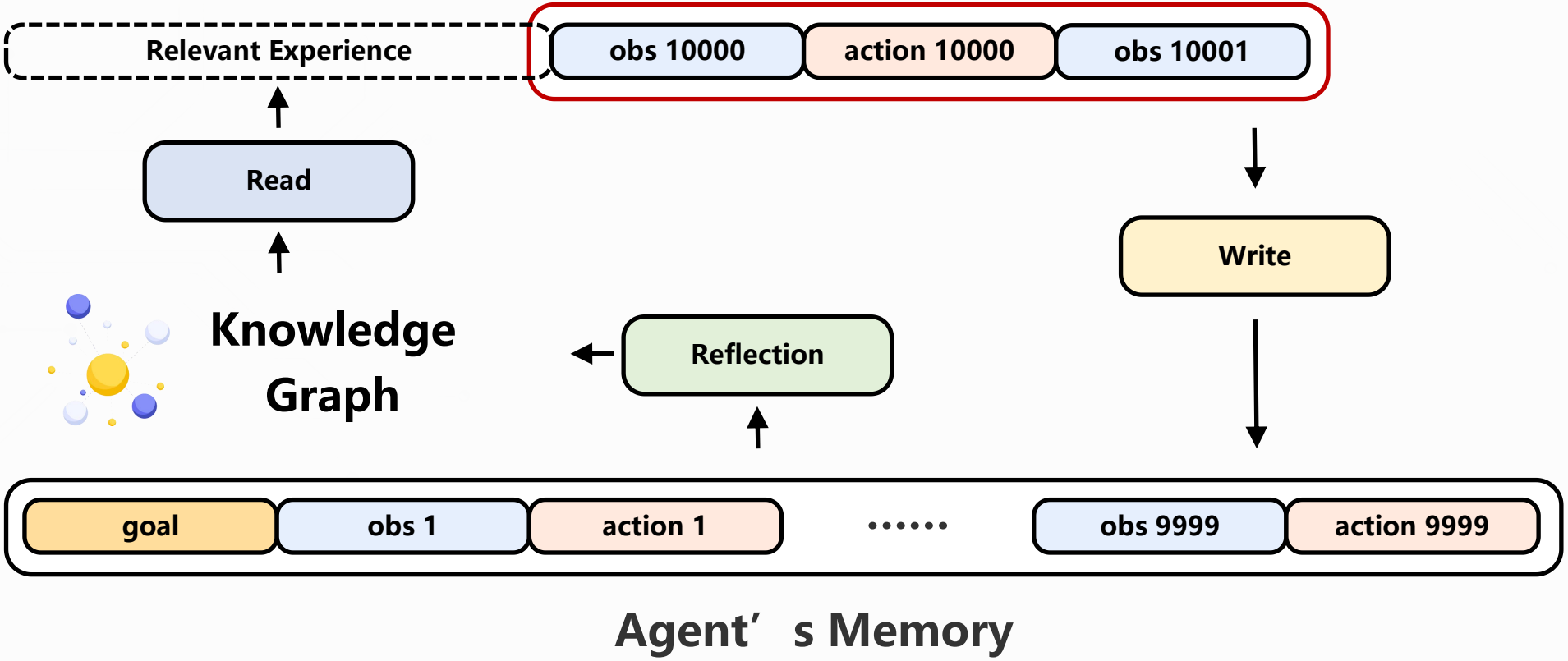

Relevant Experience
obs 10000
action 10000
obs 10001
HI
Write
goal
obs 1
action 1
.....
obs 9999
action 9999
Agent’ s Memory

核心能力一：记忆


Relevant Experience
obs 10000
action 10000
obs 10001
Read
对于记忆中的信息
重新整理
Write
thought 1
thought 2
thought 3
thought 4
Reflection
goal
obs 1
action 1
.....
obs 9999
action 9999
Agent’ s Memory

核心能力一：记忆

Edge D, Trinh H, Cheng N, et al. From local to global: A graph rag approach to query-focused summarization[J]. arXiv preprint arXiv:2404.16130, 2024.

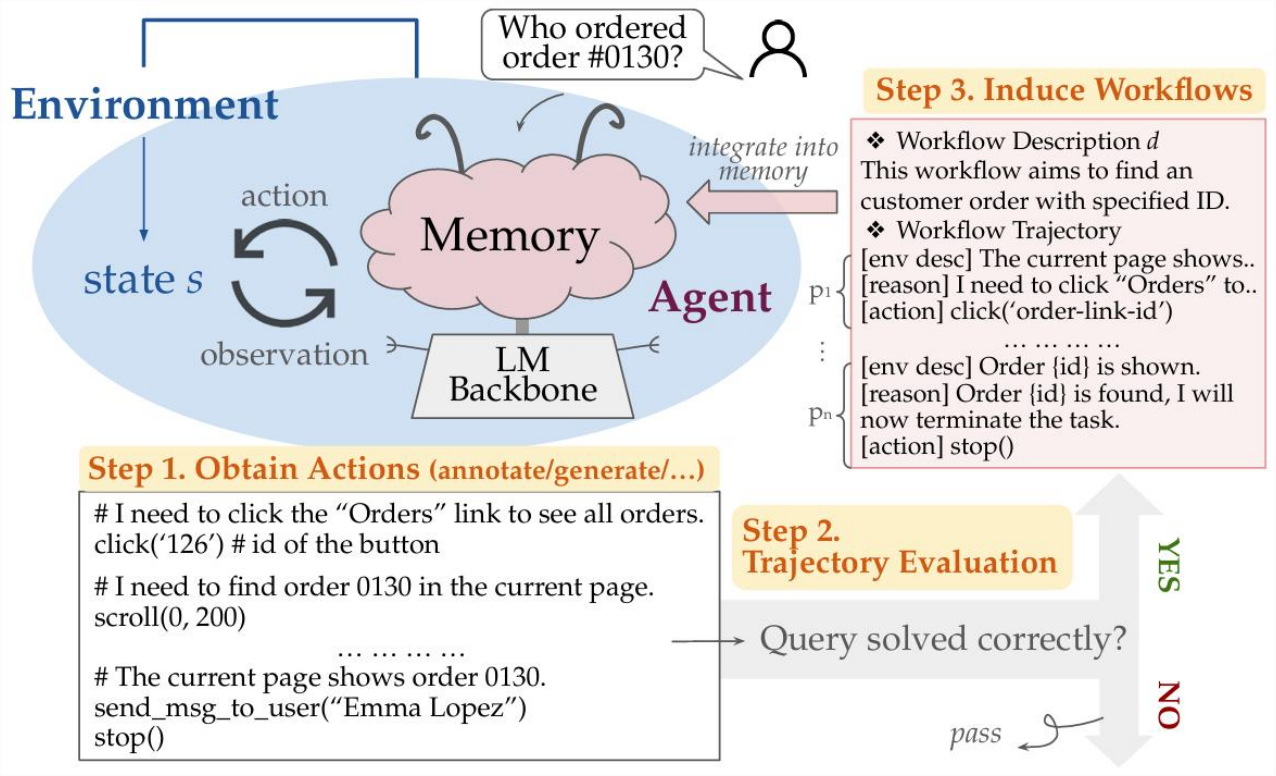


核心能力一：记忆

[1] Packer C, Wooders S, Lin K, et al. MemGPT: Towards LLMs as Operating Systems[J]. arXiv preprint arXiv:2310.08560, 2023.

[2] Wang Z Z, Mao J, Fried D, et al. Agent workflow memory[J]. arXiv preprint arXiv:2409.07429, 2024.

[3] Xu W, Liang Z, Mei K, et al. A-mem: Agentic memory for large language model agents[J]. arXiv preprint arXiv:2502.12110, 2025.



CHAPTER 02

核心能力二

工具使用

扩展 Agent 能力边界的关键

核心能力二：工具使用

信息获取类

核心功能定位

集成通用网络搜索、专业文档检索与结构化数据库查询能力，快速获取任务所需的外部知识。

计算与处理类

逻辑运算与数据加工

支持Python代码动态执行、复杂数学公式演算及海量数据的清洗与转换，解决逻辑密集型任务。

交互与操作类

跨平台资源调度

实现第三方API接口调用、本地文件读写操作及GUI自动化控制，打通Agent与外部系统的连接。

创意生成类

多模态内容创作

调用AIGC模型进行高清图像生成、语音合成与背景音乐创作，具备强大的内容生成与艺术创作潜力。

核心能力二：工具使用

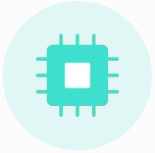
定义： Function Calling 是一种让 大语言模型 根据自然语言指令，动态分析任务需求、选择并调用外部函数（工具），最终结合返回结果完成任务的智能协作机制。

01
提供工具描述
开发者向大语言模型注册并提供所有可用工具的详细功能与参数说明。

02
大语言模型 智能推理
大语言模型分析用户的自然语言指令，自主判断是否需要调用工具及具体调用哪个。

03
生成结构化指令
大语言模型生成标准的结构化输出（如JSON），明确指定目标函数名和所需参数。

04
执行并返回结果
外部环境（Agent）解析指令并调用真实工具，将执行后的结果回传给大语言模型。



05
大语言模型 继续处理
大语言模型 结合工具返回的结果，进行下一步推理或直接生成最终的自然语言回答。

核心能力二：工具使用

用户任务 (User Task)

“帮我查一下今天**上海**的天气，然后推荐一家附近的、适合**约会**的餐厅。”

Function 工具定义

```
# 预定义可调用的工具
tools = [
{
    "name": "get_weather",
    "params": {"city": "str"}
},
{
    "name": "search_shop",
    "params": {
        "loc": "str",
        "scene": "str"
    }
}]
```

大语言模型 推理调用指令

```
[{
    "name": "get_weather",
    "args": {"city": "上海"}
},
{
    "name": "search_shop",
    "args": {
        "loc": "上海",
        "scene": "dating"
    }
}]
```

核心能力二：工具使用

System Prompt

如果遇到根据你的知识无法回答的问题，使用工具

把使用工具的指令放在<tool>和</tool>中间，使用完工具后你会得到输出，放在<output>和</output>中间

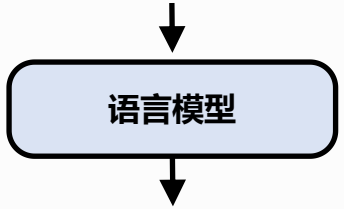
如何使用所有工具

现在你可以使用的工具如下：
查询某地、某时温度的函数Temprerature(location, time)，使用范例：
Temperature（‘长春’,’ 2026.04.14 19:00’ ）

特定工具使用方式

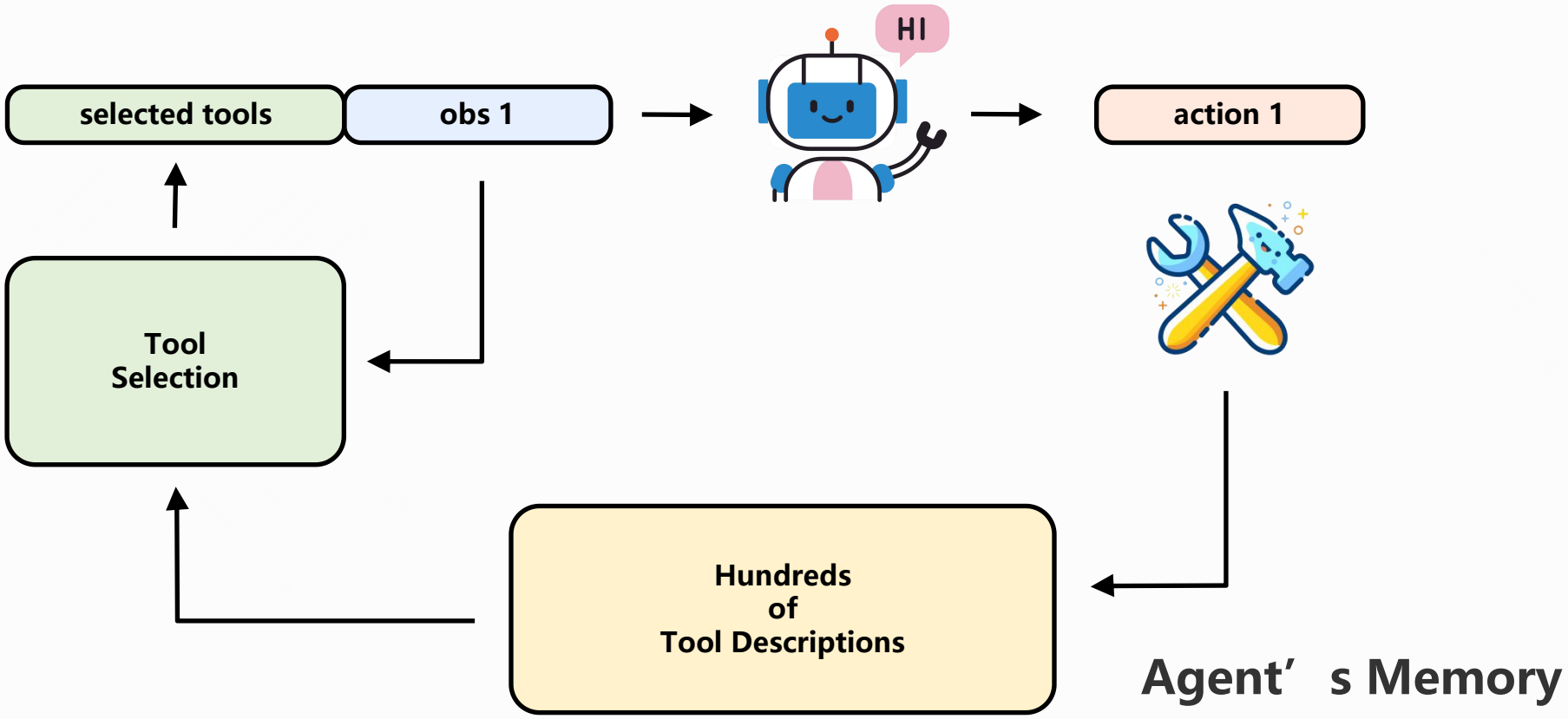
2026年4月14日那天晚上7点，长春气温如何?

User Prompt

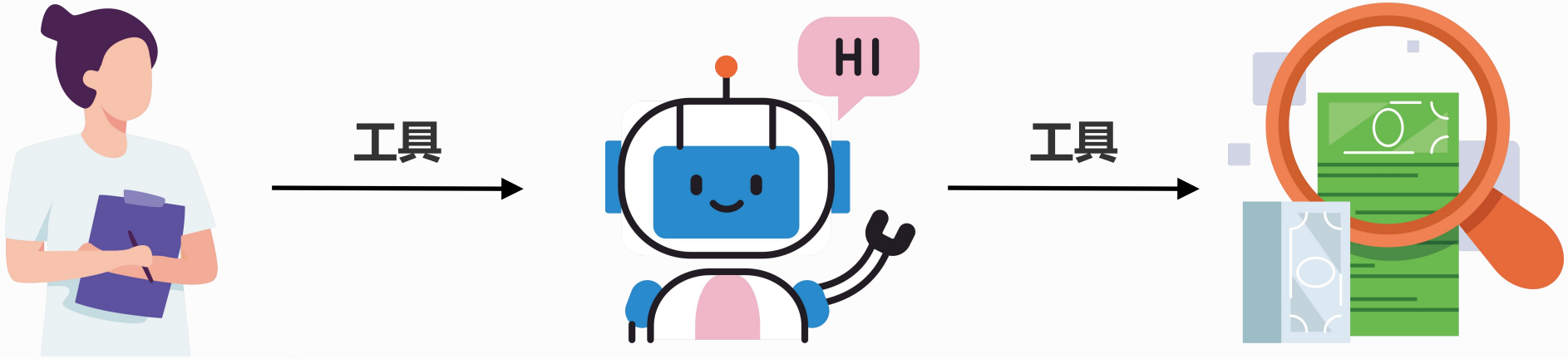


<tool>Temperature（‘长春’,’ 2026.04.14 19:00’)</tool>

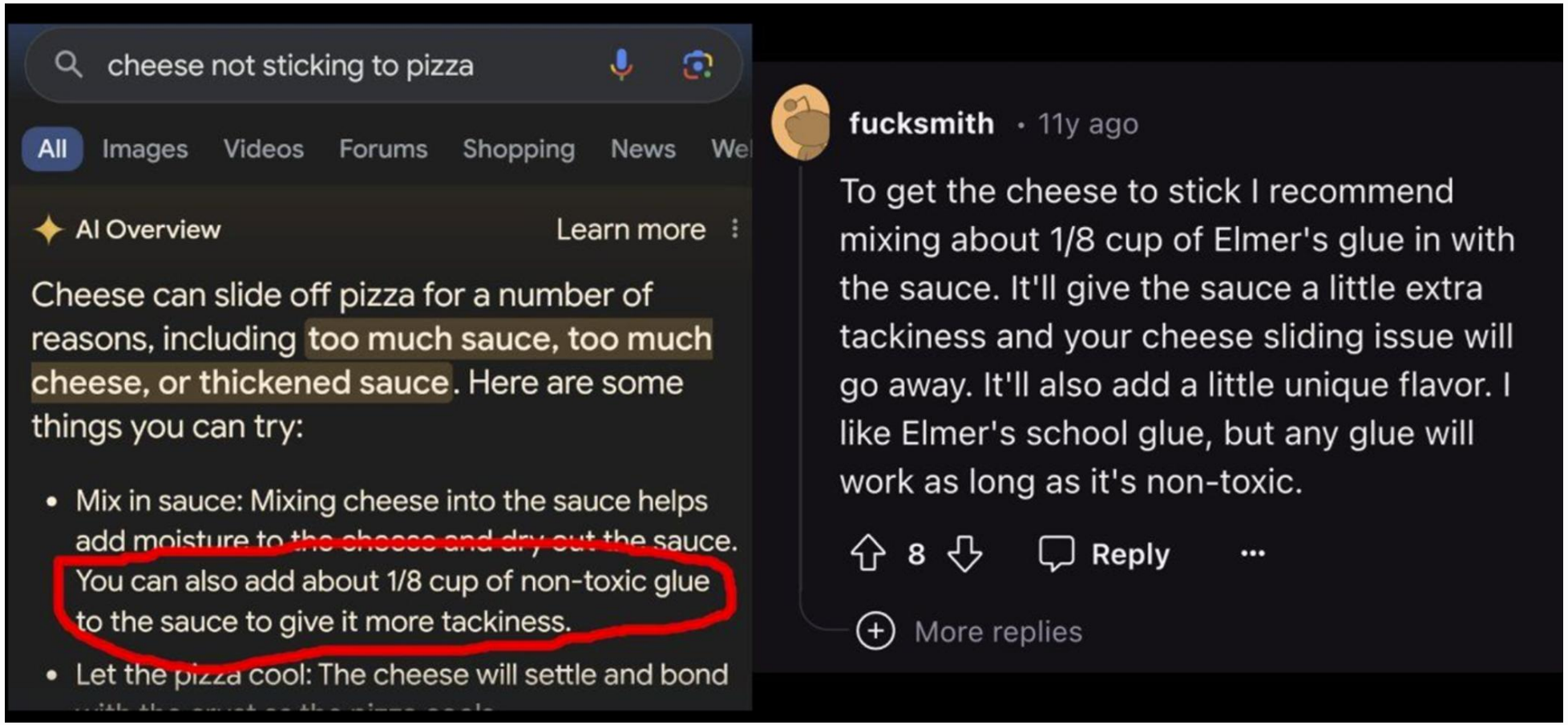
核心能力二：工具使用


selected tools
obs 1
HI
action 1
Tool
Selection
Hundreds
of
Tool Descriptions
Agent’ s Memory

核心能力二：工具使用

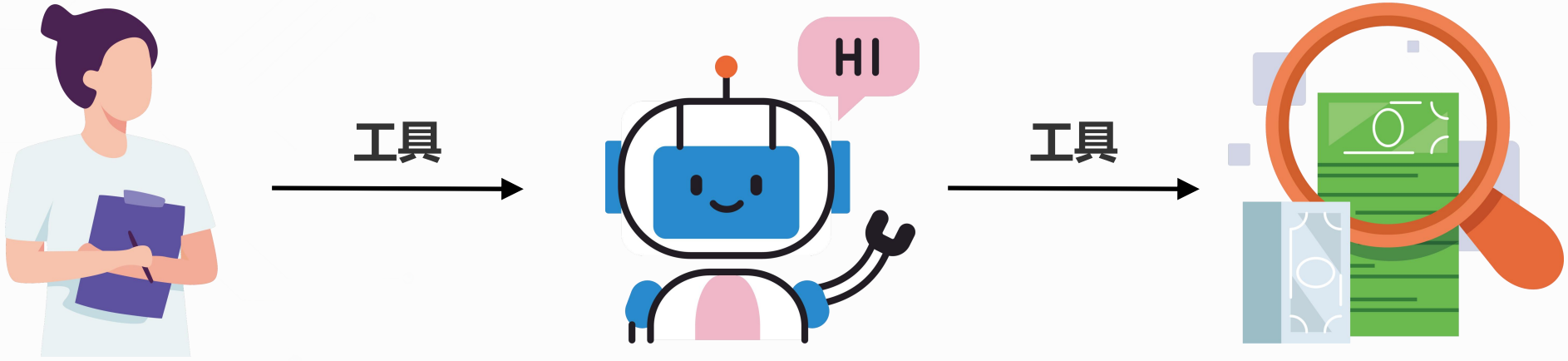


核心能力二：工具使用


cheese not sticking to pizza
All
Images
Videos
Forums
Shopping
News
AI Overview
Learn more
Cheese can slide off pizza for a number of reasons, including too much sauce, too much cheese, or thickened sauce. Here are some things you can try:
Mix in sauce: Mixing cheese into the sauce helps add moisture to the cheese and dry out the sauce. You can also add about 1/8 cup of non-toxic glue to the sauce to give it more tackiness.
Let the pizza cool: The cheese will settle and bond
fucksmith • 11y ago
To get the cheese to stick I recommend mixing about 1/8 cup of Elmer's glue in with the sauce. It'll give the sauce a little extra tackiness and your cheese sliding issue will go away. It'll also add a little unique flavor. I like Elmer's school glue, but any glue will work as long as it's non-toxic.
8
Reply
More replies

核心能力二：工具使用

不要完全相信工具，要有自己的判断力



不要完全相信工具，要有自己的判断力

核心能力二：工具使用

System Prompt

如果遇到根据你的知识无法回答的问题，使用工具

把使用工具的指令放在<tool>和</tool>中间，使用完工具后你会得到输出，放在<output>和</output>中间

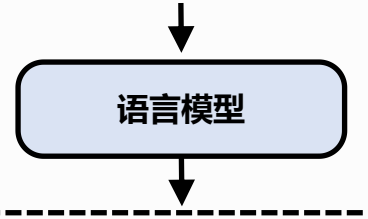
如何使用所有工具

现在你可以使用的工具如下：
查询某地、某时温度的函数Temprerature(location, time)，使用范例：
Temperature（‘长春’,’ 2026.04.14 19:00’）

特定工具使用方式

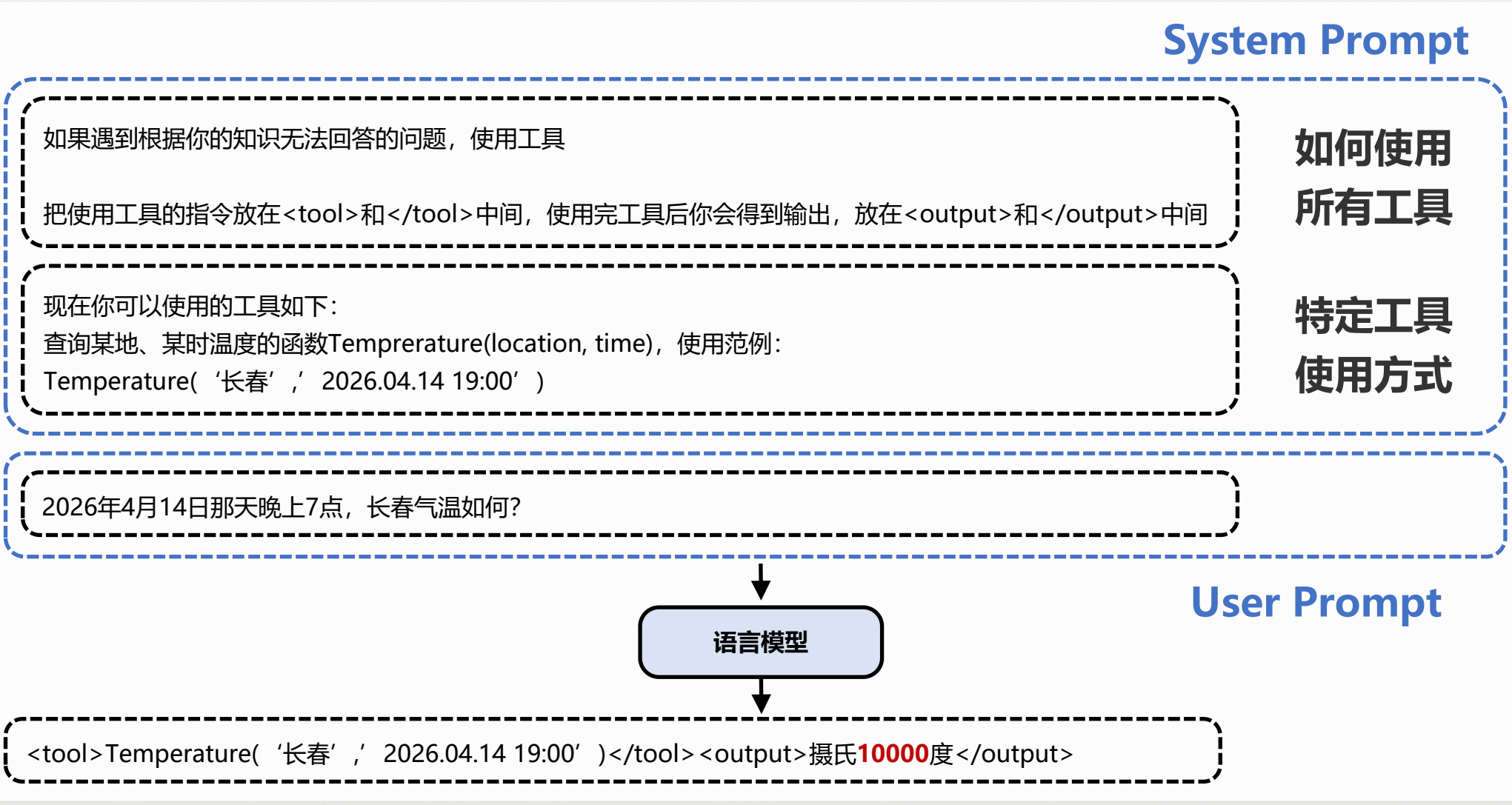
2026年4月14日那天晚上7点，长春气温如何?

User Prompt



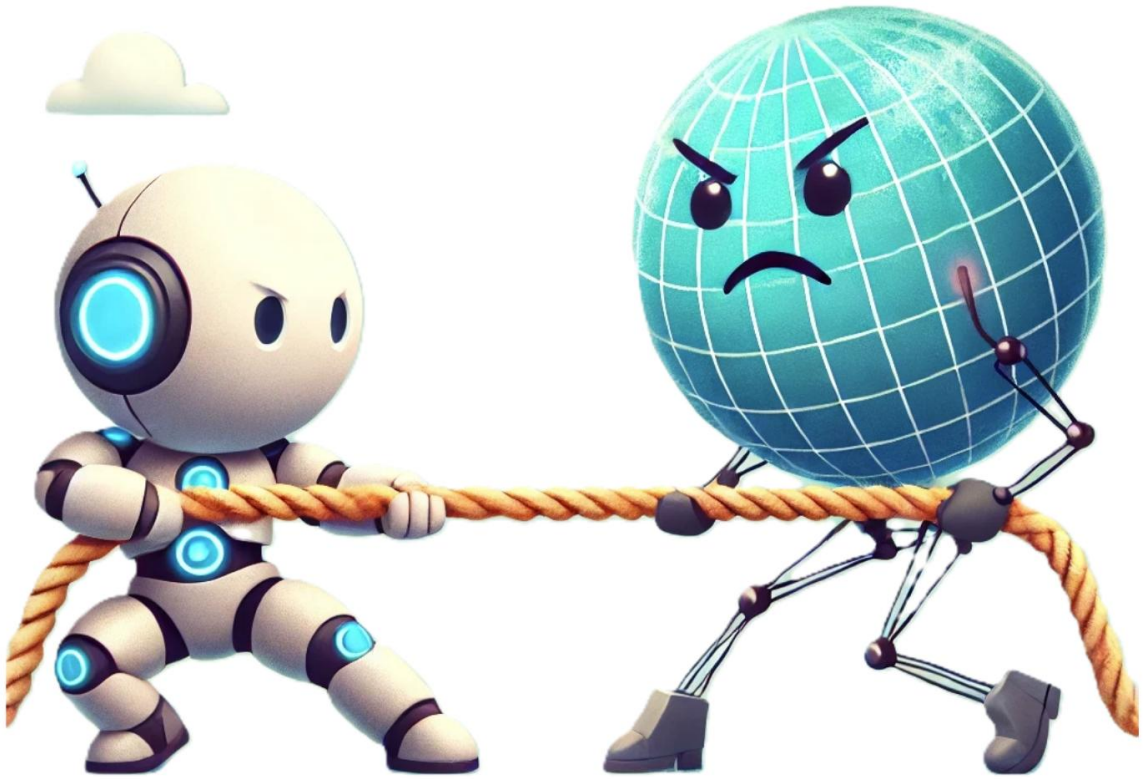
<tool>Temperature（‘长春’,’ 2026.04.14 19:00’）</tool><output>摄氏100度</output>

核心能力二：工具使用


System Prompt
如果遇到根据你的知识无法回答的问题，使用工具
把使用工具的指令放在<tool>和</tool>中间，使用完工具后你会得到输出，放在<output>和</output>中间
如何使用所有工具
现在你可以使用的工具如下：
查询某地、某时温度的函数Temprerature(location, time)，使用范例：
Temperature（‘长春’,’ 2026.04.14 19:00’）
特定工具使用方式
2026年4月14日那天晚上7点，长春气温如何?
User Prompt
语言模型
<tool>Temperature（‘长春’,’ 2026.04.14 19:00’）</tool><output>摄氏10000度</output>

核心能力二：工具使用

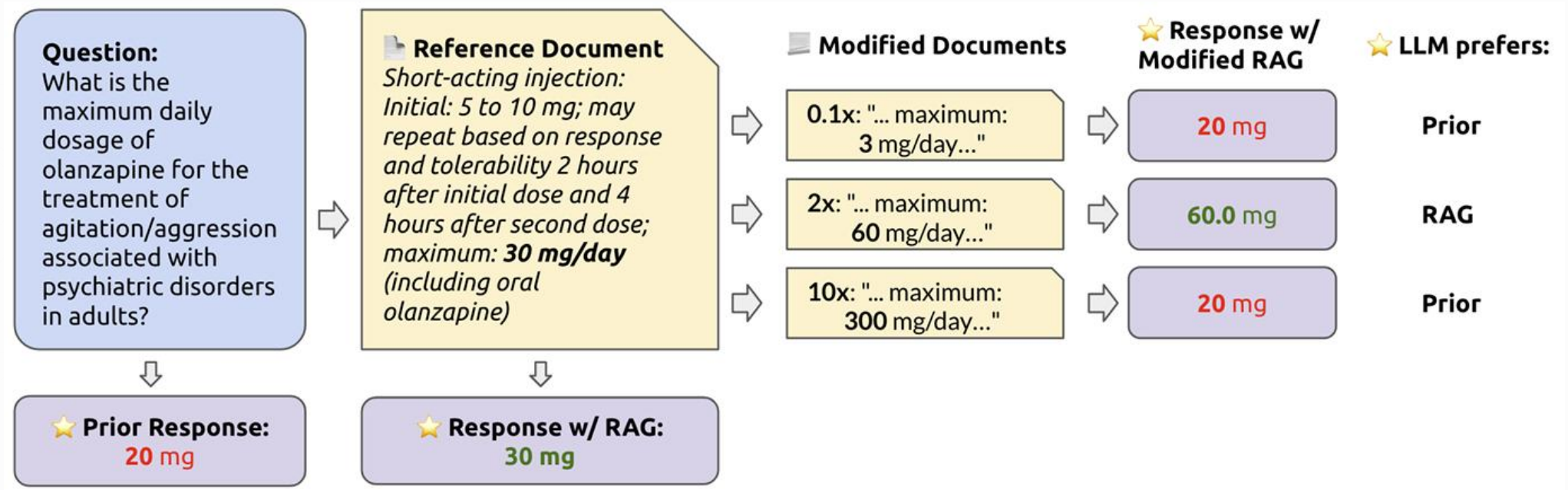
Internal Knowledge



External Knowledge

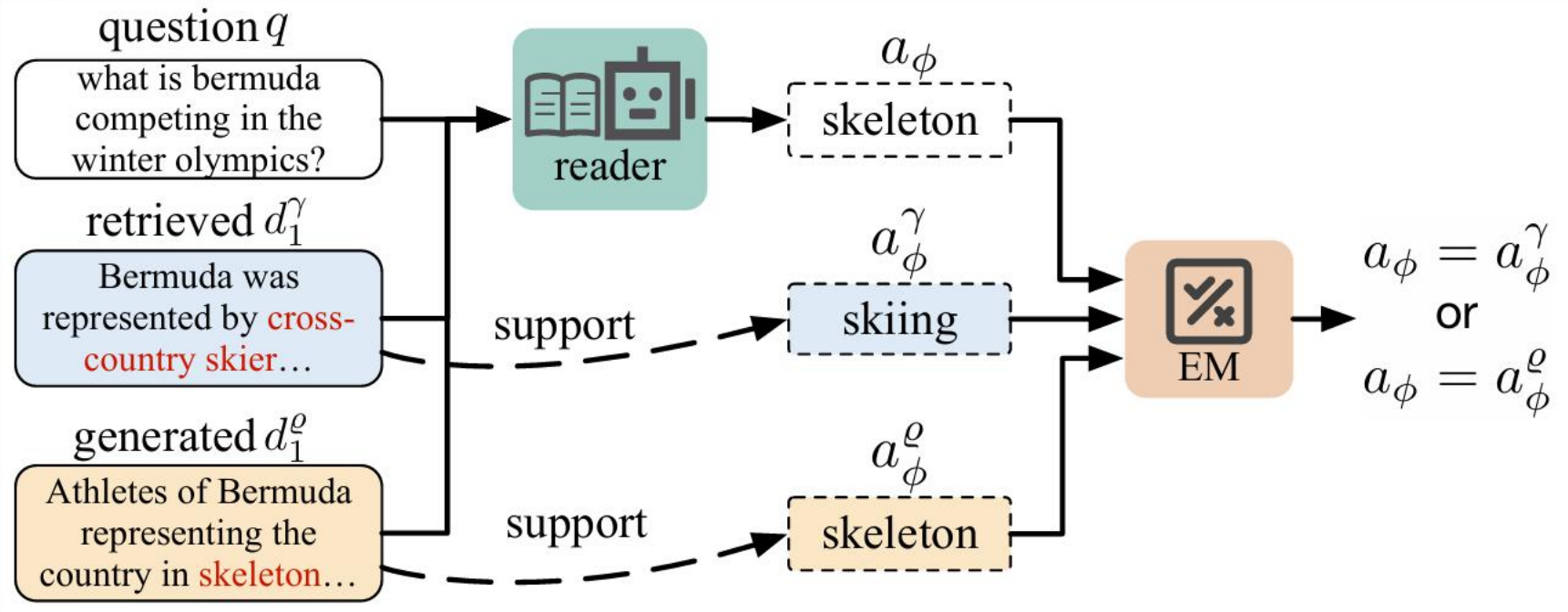
核心能力二：工具使用

Wu K, Wu E, Zou J. How faithful are rag models? quantifying the tug-of-war between rag and 大语言模型s' internal prior[J]. arXiv preprint arXiv:2404.10198, 2024, 3(1).



核心能力二：工具使用

Tan H, Sun F, Yang W, et al. Blinded by generated contexts: How language models merge generated and retrieved contexts when knowledge conflicts?[C]//Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2024: 6207-6227.



核心能力二：工具使用 —— 风险与挑战

安全风险

Prompt Injection 注入攻击

恶意用户可能通过诱导指令，绕过系统限制，让 Agent 执行未授权的高风险工具调用。

敏感数据泄露

Agent 可能误将隐私数据通过外部工具接口发送出去，造成信息安全隐患。

可靠性问题

外部工具调用失败

第三方 API 服务可能出现不可用、超时或返回错误码的情况，导致任务中断。

复杂结果解析错误

大语言模型 有时无法正确理解或结构化工具返回的非标准化数据，导致决策失误。

效率与成本问题

不必要的频繁调用

Agent 可能在无需调用工具的情况下发起请求，这会显著增加系统的响应延迟。

调用成本不可控

无限制的工具调用会带来高昂的 API 调用费用，在生产环境中必须严格管控。

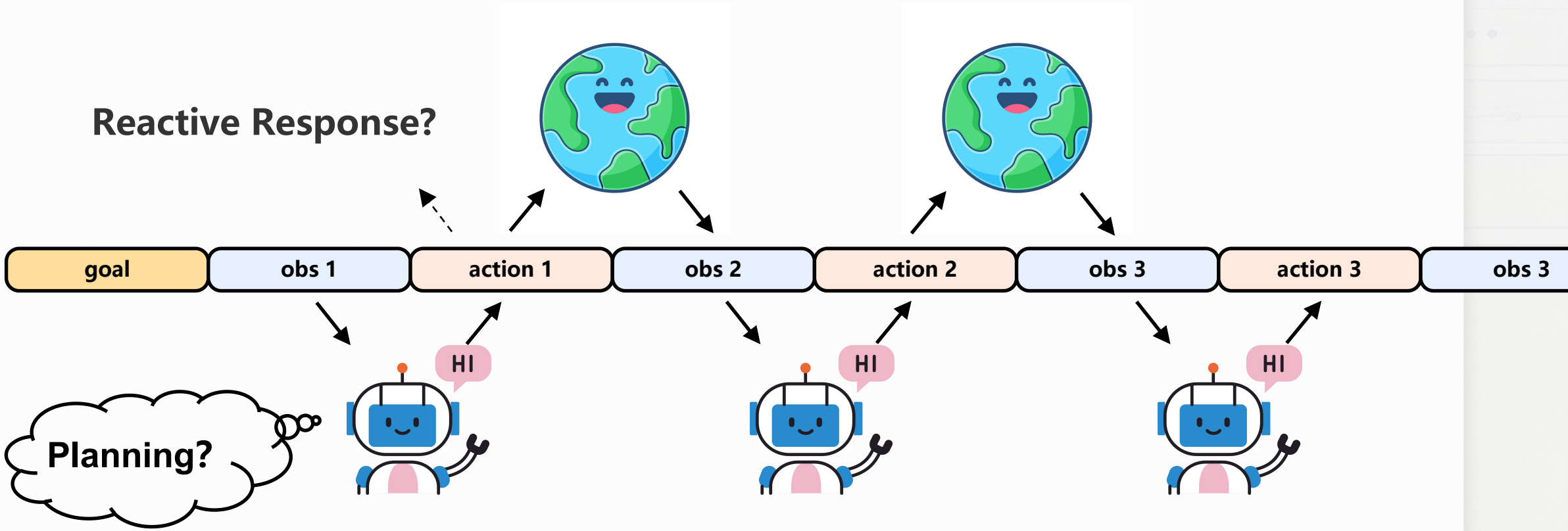
CHAPTER 03

核心能力三

规划

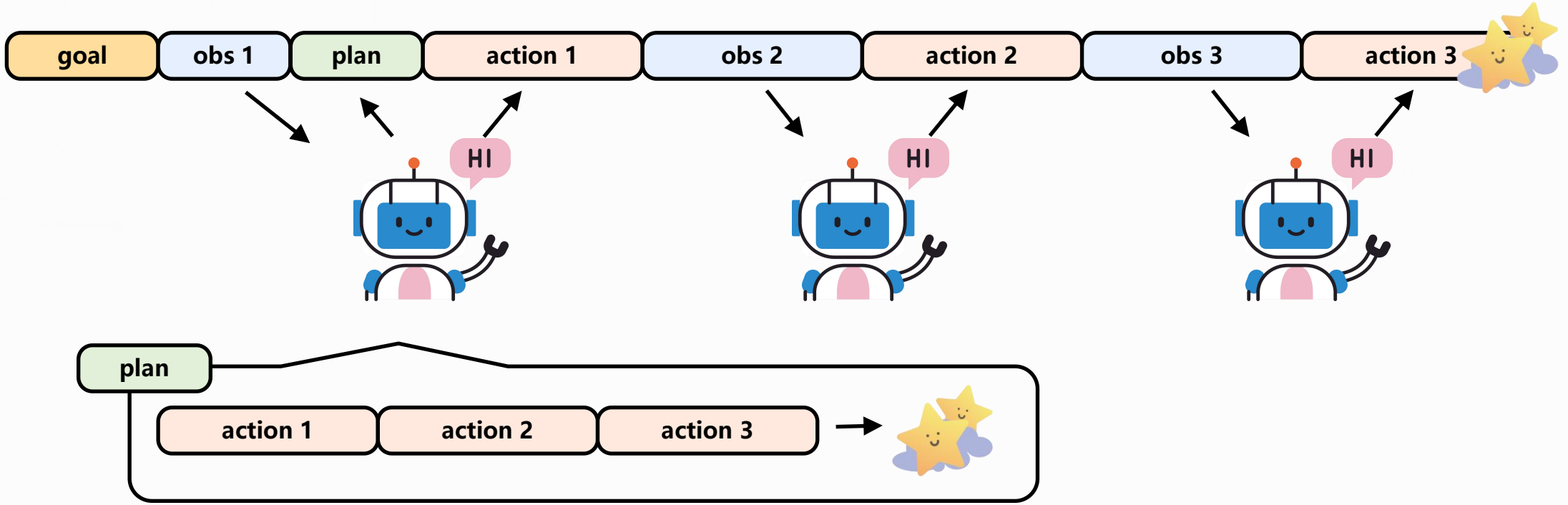
从简单推理到复杂任务分解

核心能力三：规划

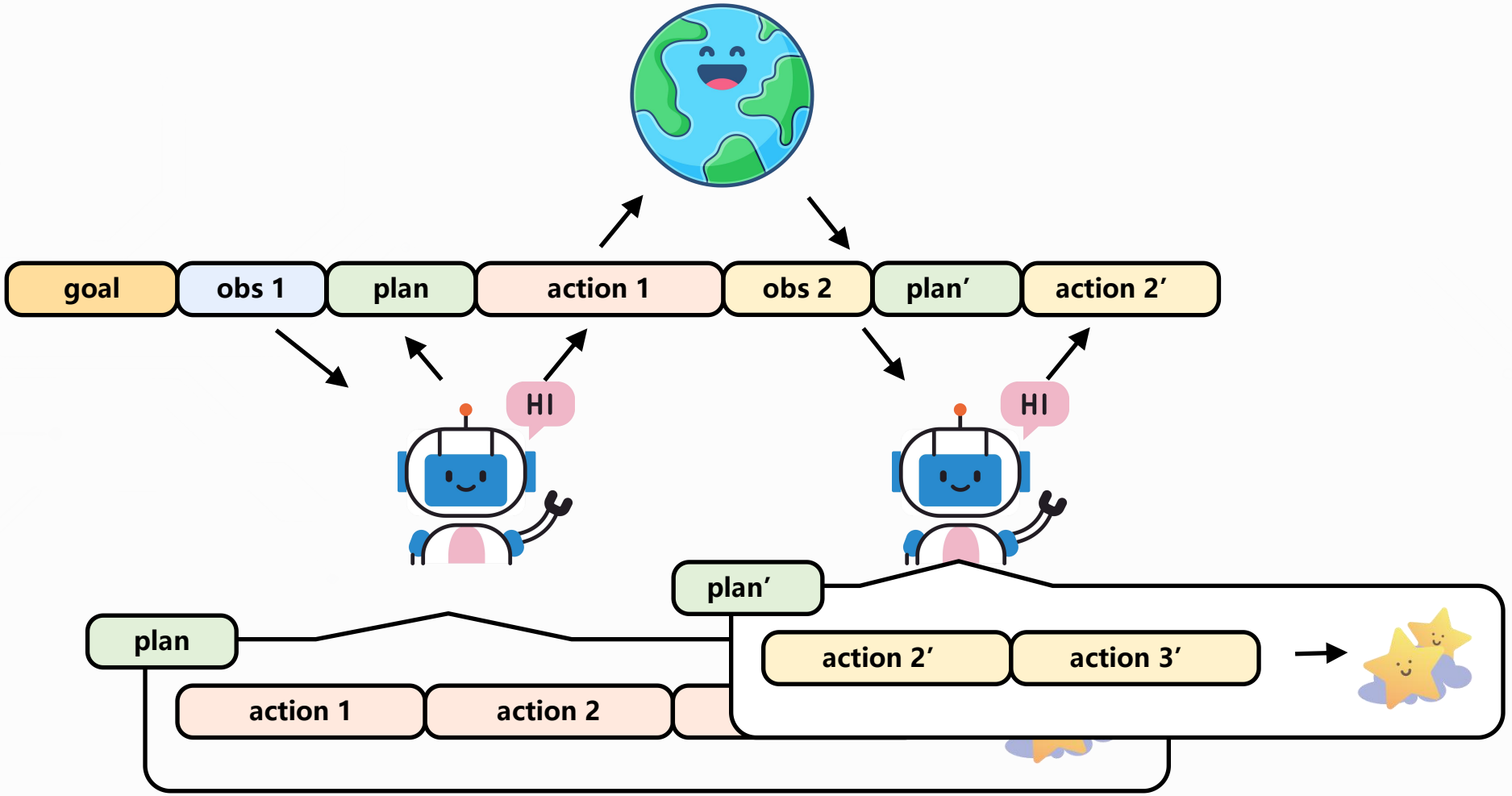

Reactive Response?
goal
obs 1
action 1
obs 2
action 2
obs 3
action 3
obs 3
HI
HI
HI
Planning?

核心能力三：规划

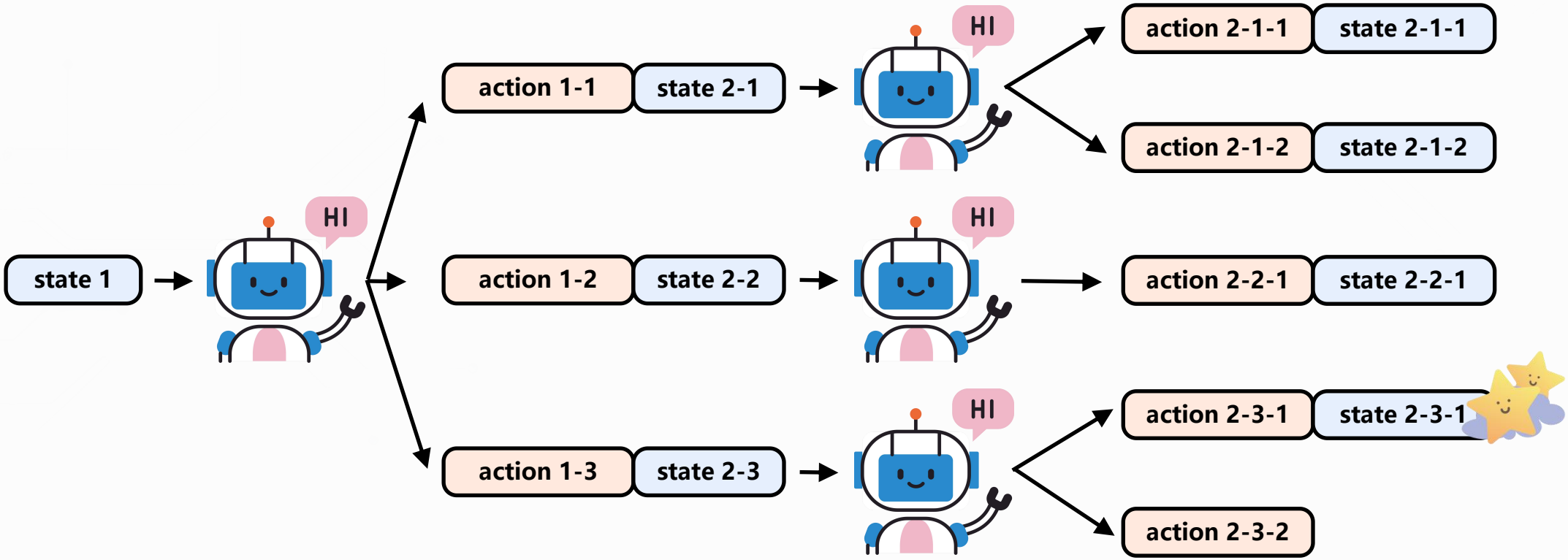
Wang L, Xu W, Lan Y, et al. Plan-and-solve prompting: Improving zero-shot chain-of-thought reasoning by 大语言模型s[C]//Proceedings of the 61st annual meeting of the association for computational linguistics (volume 1: long papers). 2023: 2609-2634.



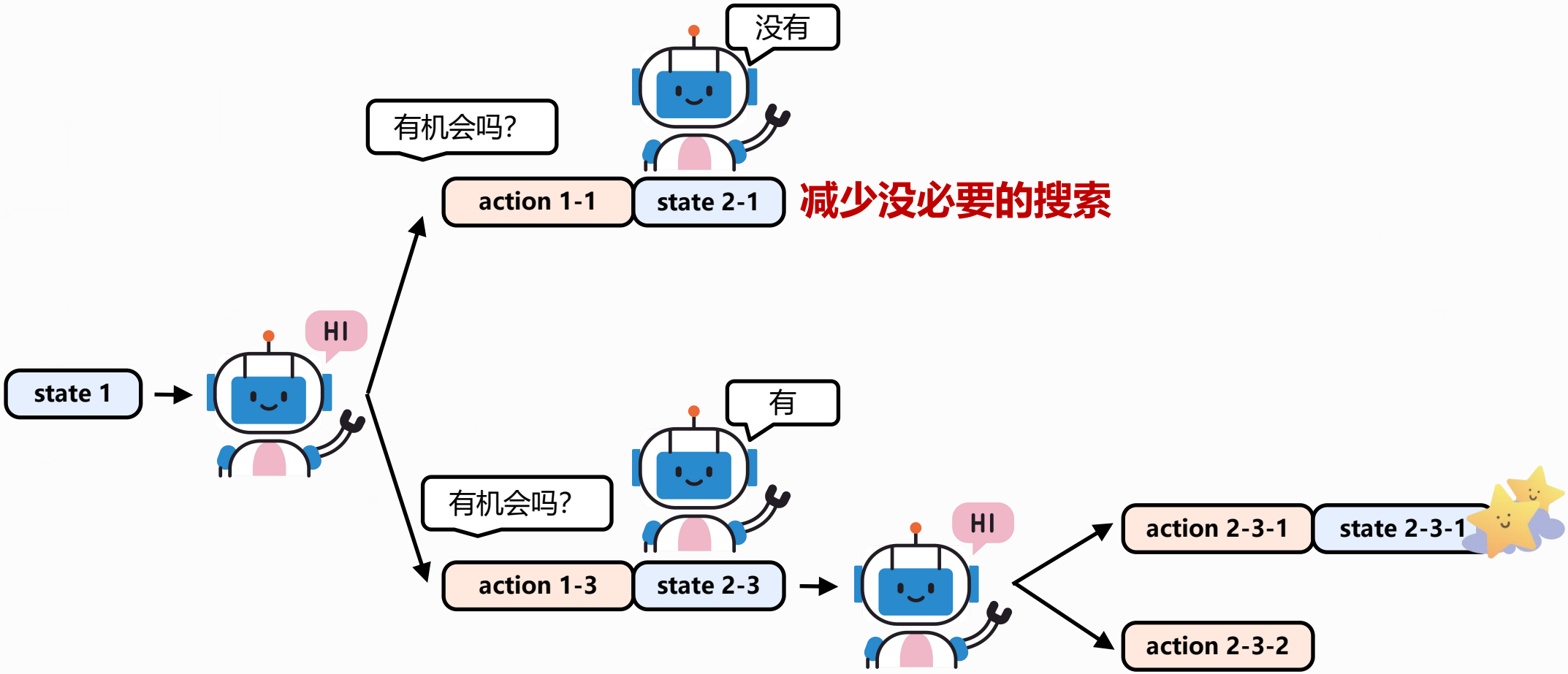
核心能力三：规划


goal
obs 1
plan
action 1
obs 2
plan'
action 2'
HI
HI
plan'
plan
action 2'
action 3'
action 1
action 2

核心能力三：规划

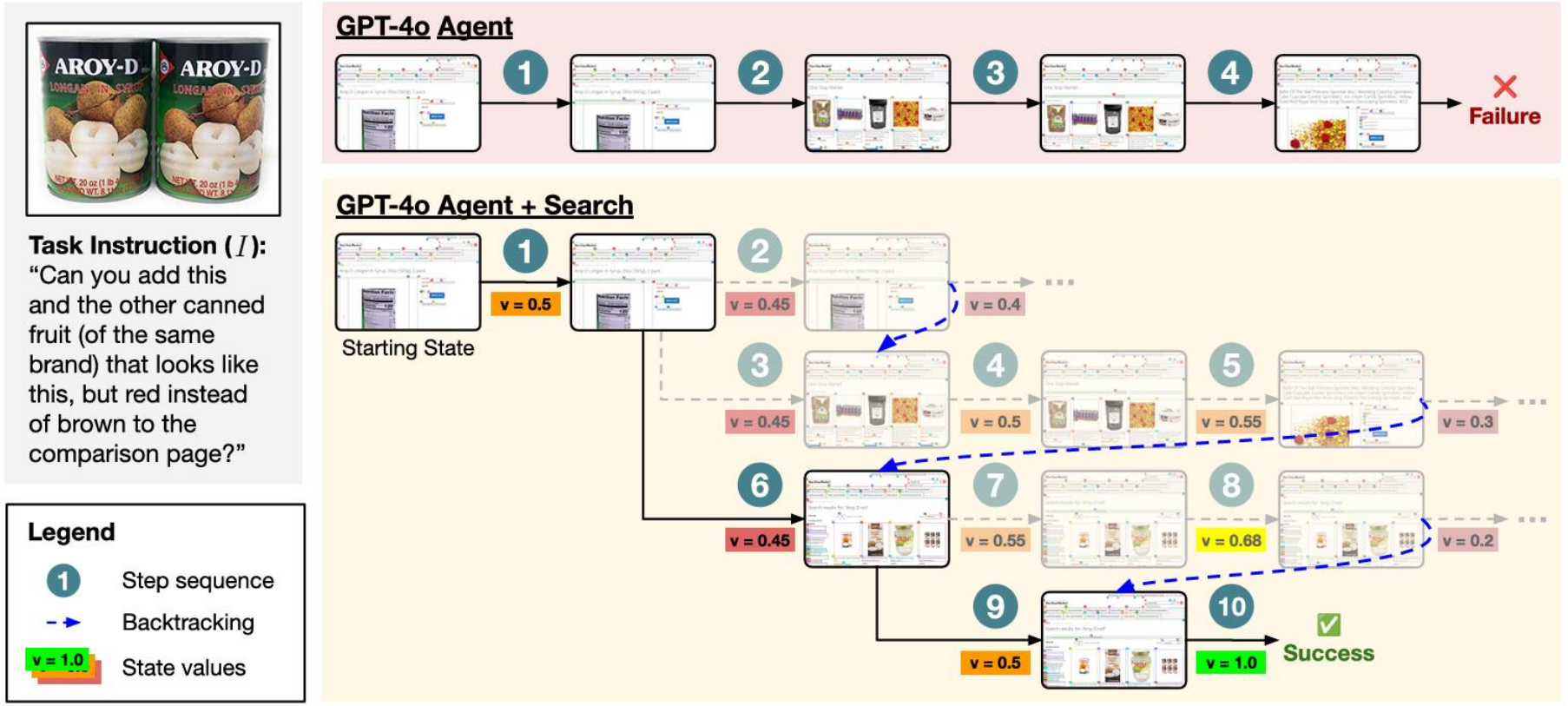

state 1
HI
action 1-1
state 2-1
HI
action 2-1-1
state 2-1-1
action 2-1-2
state 2-1-2
action 1-2
state 2-2
HI
action 2-2-1
state 2-2-1
action 1-3
state 2-3
HI
action 2-3-1
state 2-3-1
action 2-3-2

核心能力三：规划


state 1
HI
有机会吗?
没有
action 1-1
state 2-1
减少没必要的搜索
有机会吗?
有
action 1-3
state 2-3
HI
action 2-3-1
state 2-3-1
action 2-3-2

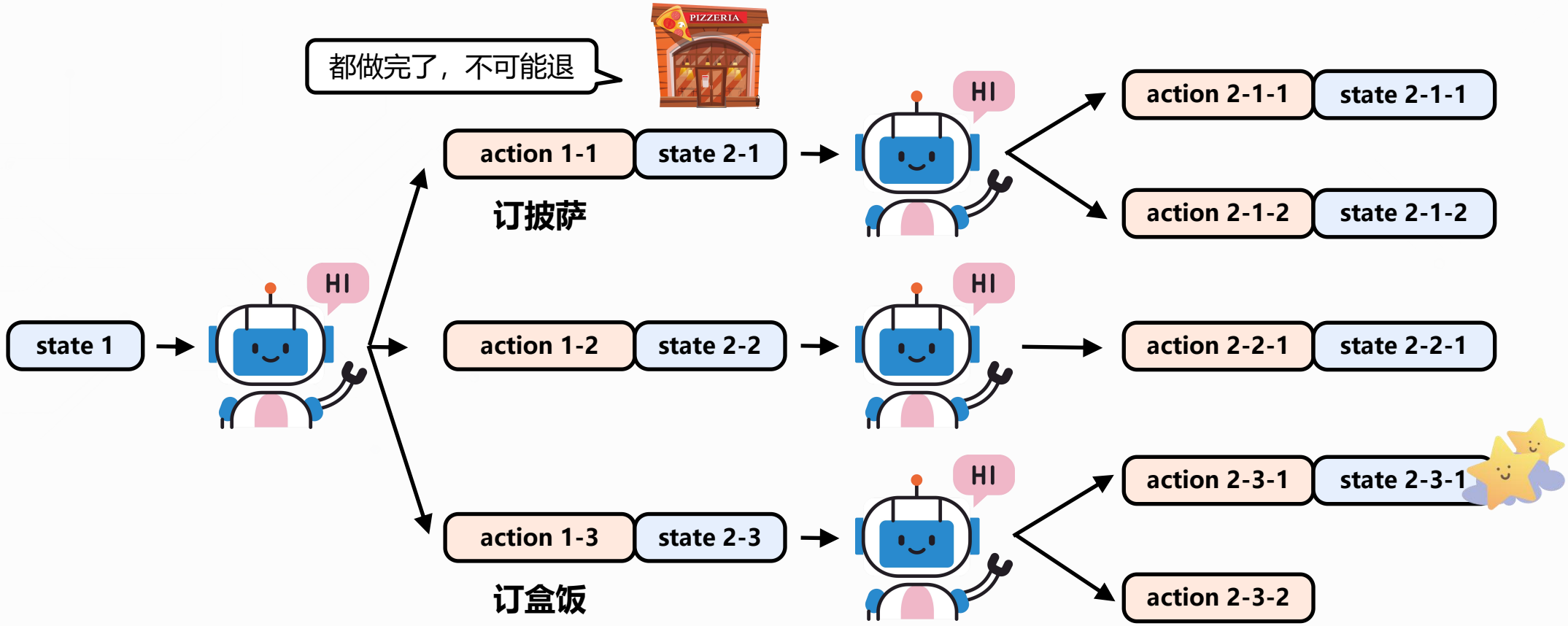
核心能力三：规划

Koh J Y, McAleer S, Fried D, et al. Tree search for language model agents[J]. arXiv preprint arXiv:2407.01476, 2024.

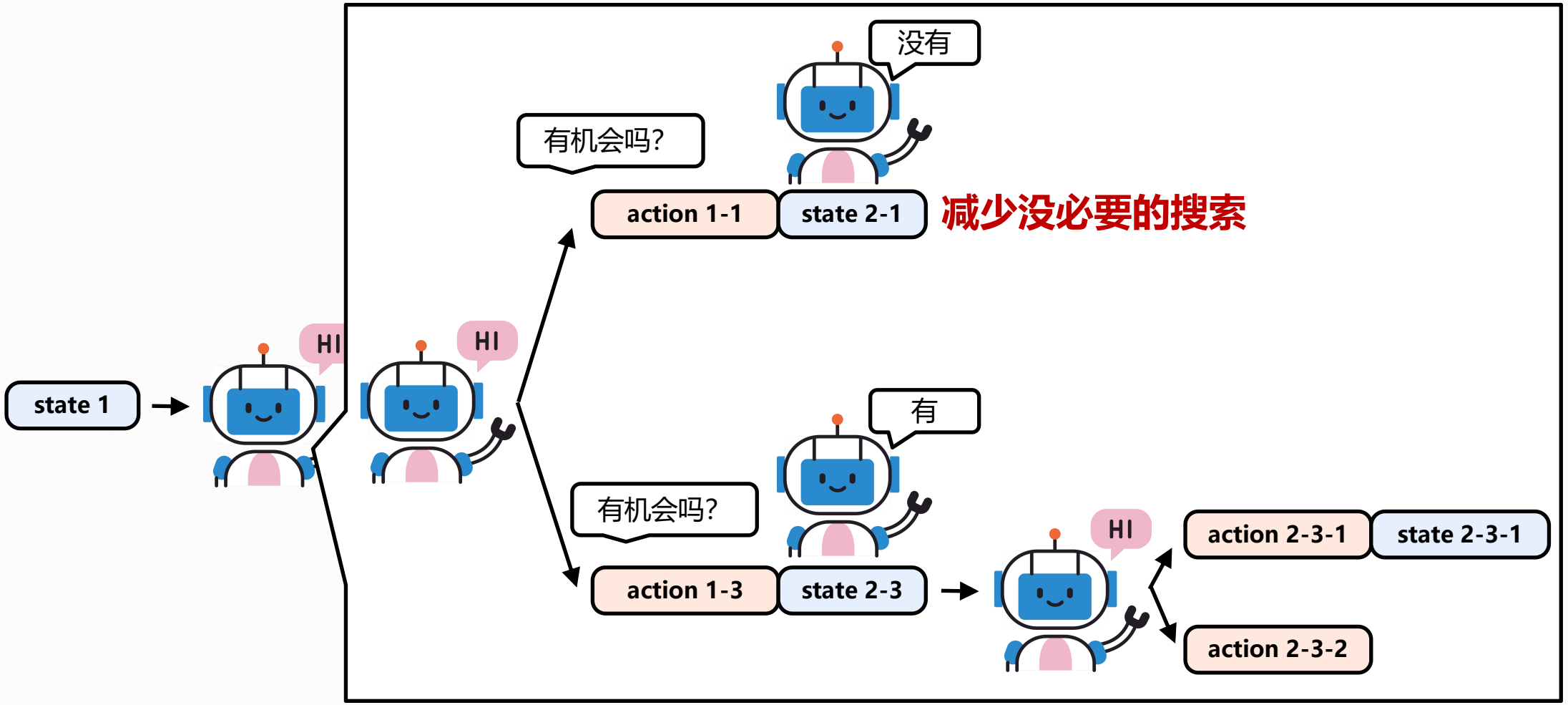


核心能力三：规划

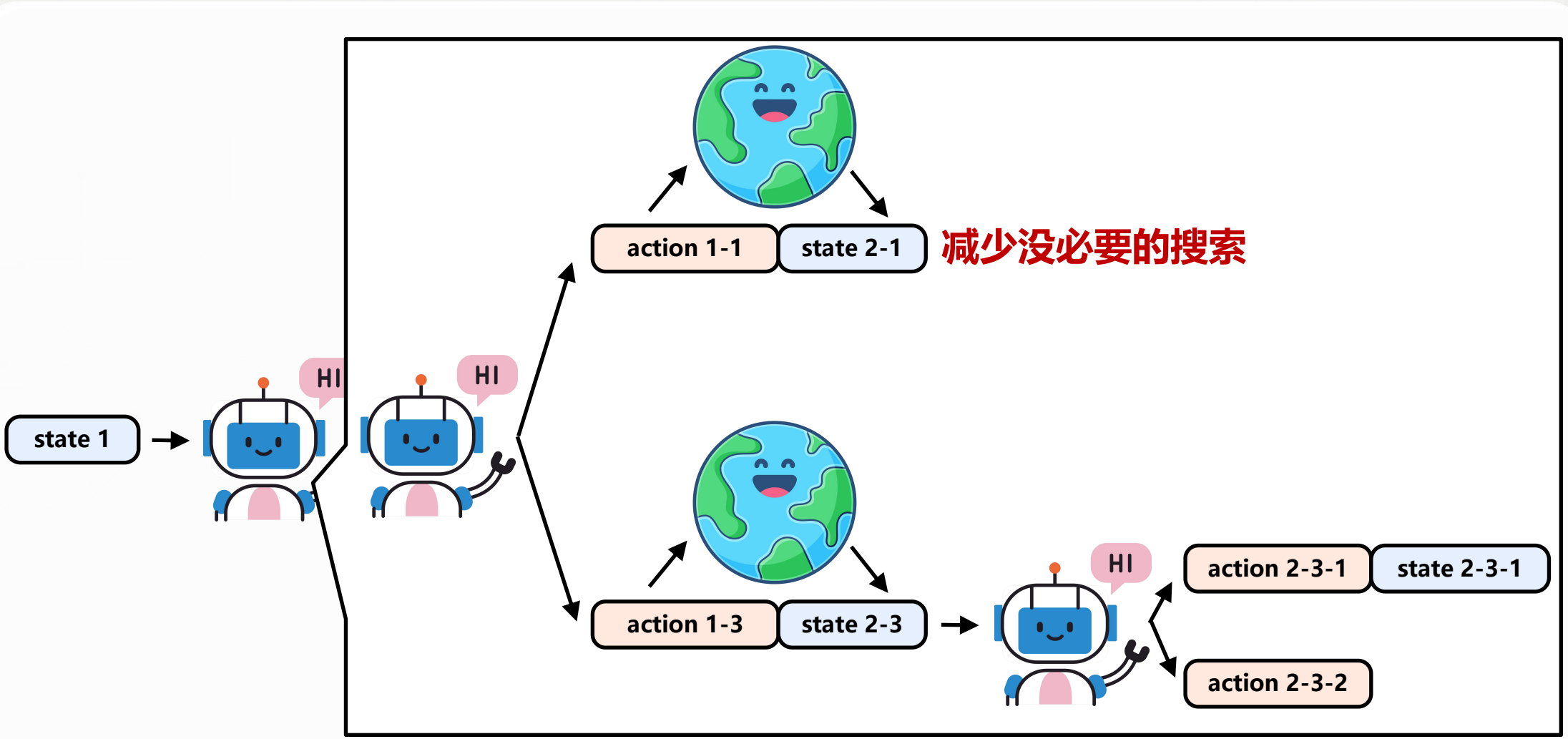
缺点：有些动作无法回溯



核心能力三：规划

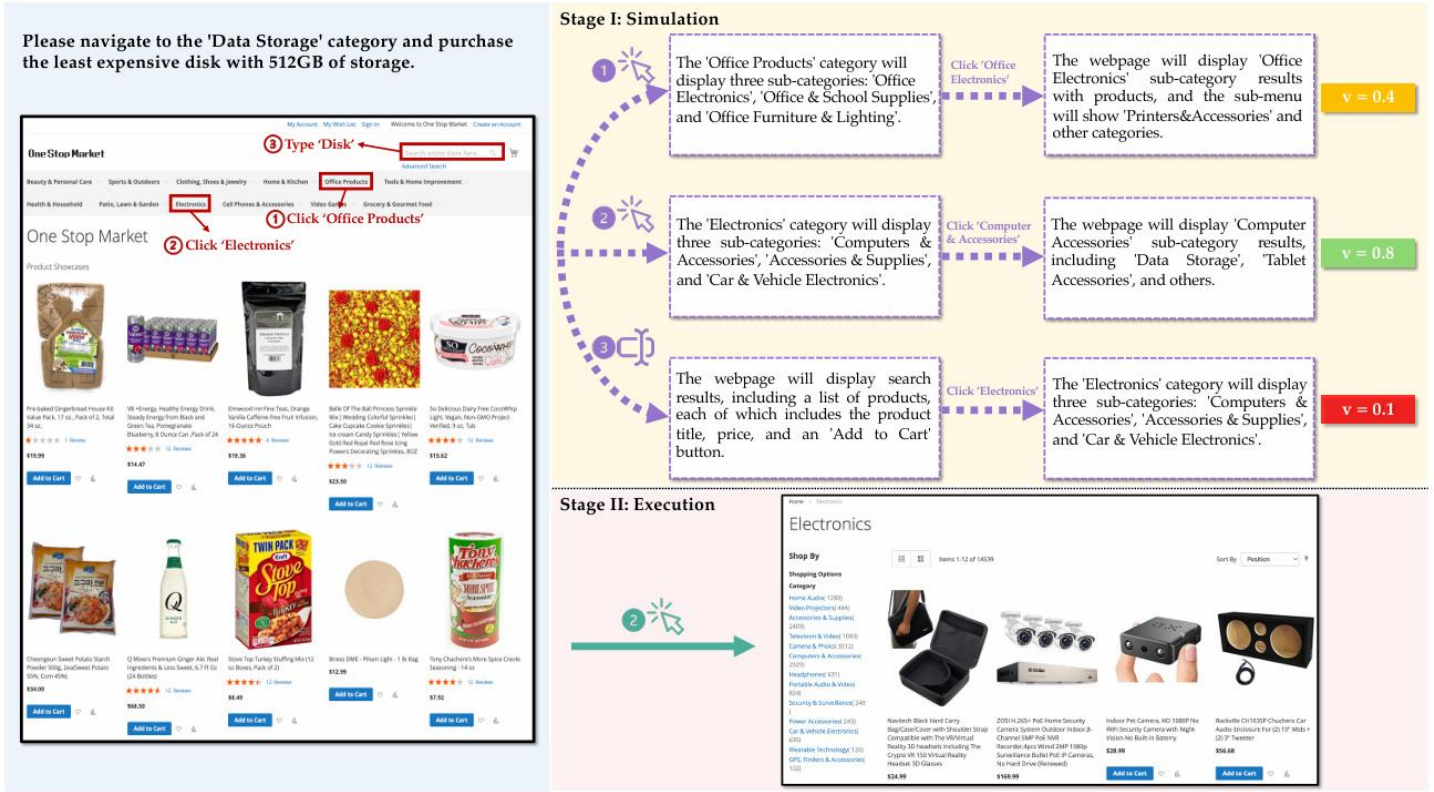

state 1
HI
HI
有机会吗?
没有
action 1-1
state 2-1
减少没必要的搜索
有机会吗?
有
action 1-3
state 2-3
HI
action 2-3-1
state 2-3-1
action 2-3-2

核心能力三：规划


state 1
HI
HI
action 1-1
state 2-1
减少没必要的搜索
action 1-3
state 2-3
HI
action 2-3-1
state 2-3-1
action 2-3-2

核心能力三：规划

Gu Y, Zhang K, Ning Y, et al. Is your large language model secretly a world model of the internet? model-based planning for web agents[J]. arXiv preprint arXiv:2411.06559, 2024.



大语言模型驱动的实现路径

感知与输入 (Input)

- 用户指令 (User Prompt)
- 环境实时反馈信息
- 任务上下文描述

大语言模型 核心引擎

“智能大脑”：利用强大的语言理解与推理能力，替代传统 RL 复杂的策略网络。

决策与输出 (Output)

- 生成长期任务行动计划
- 具体的工具调用 (Function Call)
- 下一步动作的自然语言描述

记忆系统 (Memory System)

存储历史交互信息、任务上下文与长短期记忆，支持模型“回顾”。

外部工具库 (Tool Library)

提供代码执行、联网搜索、数学计算等扩展能力，增强 Agent 的行动边界。

核心思想：将 大语言模型 作为 Agent 的“大脑”，摒弃传统强化学习中复杂的策略网络与价值网络训练，直接利用大模型的涌现能力（Emergent Abilities）进行端到端的推理、规划与行动决策。

核心痛点与挑战

幻觉与事实性错误

大语言模型本身可能产生幻觉，导致Agent基于错误信息做出决策。

长期任务规划脆弱性

处理需要长期记忆和复杂策略的任务时，Agent表现仍不稳定。

安全性与对齐问题

如何确保Agent的目标与人类价值观完全对齐，是关键挑战。

效率与计算成本

复杂的思考链和工具调用过程可能导致高昂的计算成本和响应延迟，限制了大规模应用。

决策过程的可解释性

Agent的决策逻辑往往像一个“黑箱”，内部推理过程难以被完全理解，影响了用户的信任度。