NANJING UNIVERSITY

Learning And Mining from DatA

# Neuro-Symbolic Agent:

# 从反应式决策到可验证决策

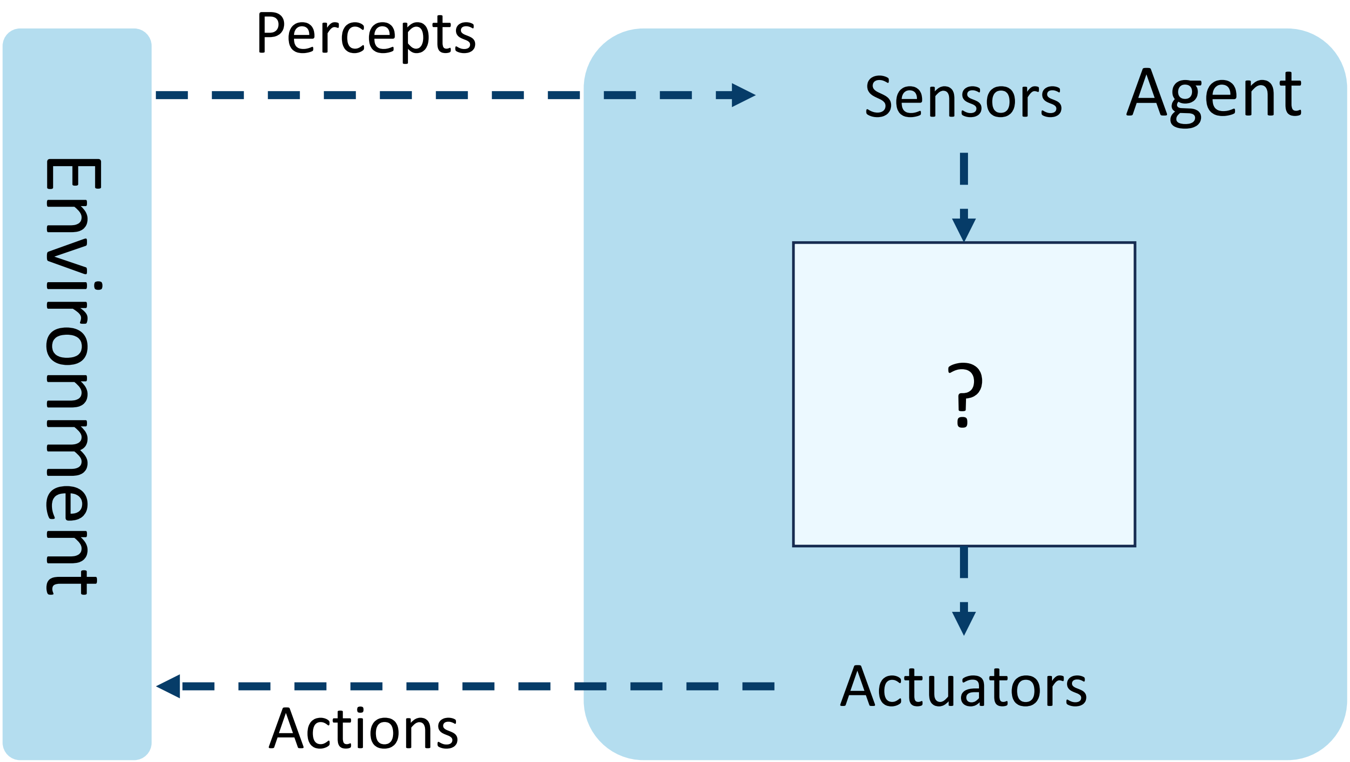

郭兰哲

南京大学

https://www.lamda.nju.edu.cn/guolz/


2026.06.13 @ 智源大会

# Agent




An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators.

— Russell & Norvig, AI: A Modern Approach

# Two Views of Agent

## LLM-first view

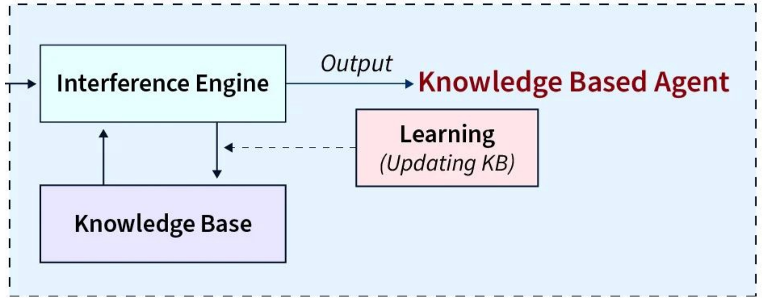
- Idea: 让我们把LLM变成Agent
- 关键：如何在LLM的基础上一步步升级，从Chatbot到Reasoner再到Agent？

## Agent-first view

- Idea: 让我们把LLM放到Agent架构中
- 关键：Agent本身有哪些挑战？LLM的引入可以帮助缓解哪些挑战，还存在哪些新的挑战？

# Evolution of AI Agents：Agent-First View

## Symbolic Agent

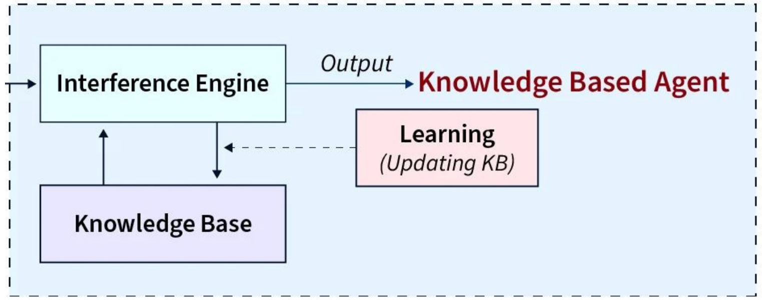



显式符号规则+推理规划算法

| | | | |
|---|---|---|---|
| 感知能力 | 缺乏符号接地能力 | | |
| 推理能力 | 严谨符号逻辑推理 | | |
| 适应能力 | 依赖预先定义的符号空间 | | |

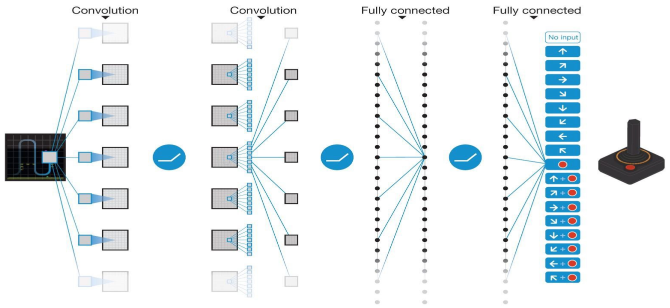
# Evolution of AI Agents：Agent-First View

**Symbolic Agent**




显式符号规则+推理规划算法

**Neural & RL Agent**

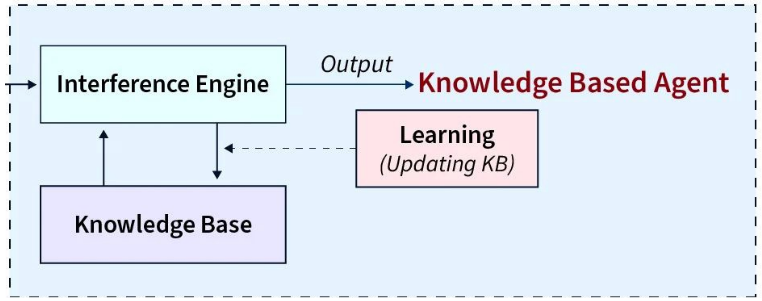



神经网络感知+强化学习策略

| | | | |
|---|---|---|---|
| 感知能力 | 缺乏符号接地能力 | 神经网络的感知能力 | |
| 推理能力 | 严谨符号逻辑推理 | 强化学习策略 | |
| 适应能力 | 依赖预先定义的符号空间 | 样本效率低，难以在不同场景迁移 | |

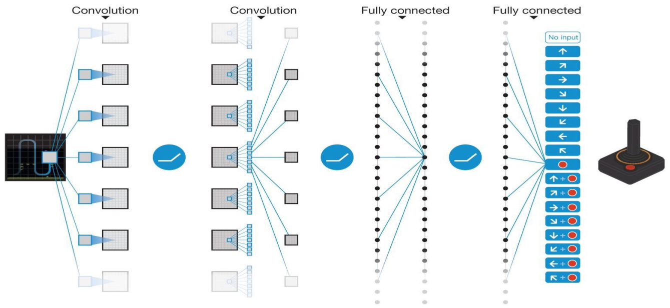
# Evolution of AI Agents：Agent-First View

**Symbolic Agent**



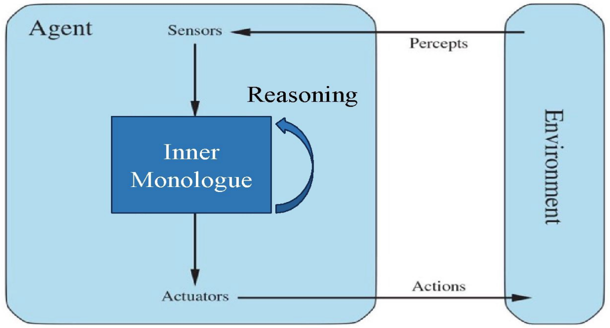

显式符号规则+推理规划算法

**Neural & RL Agent**




神经网络感知+强化学习策略

**LLM Agent**

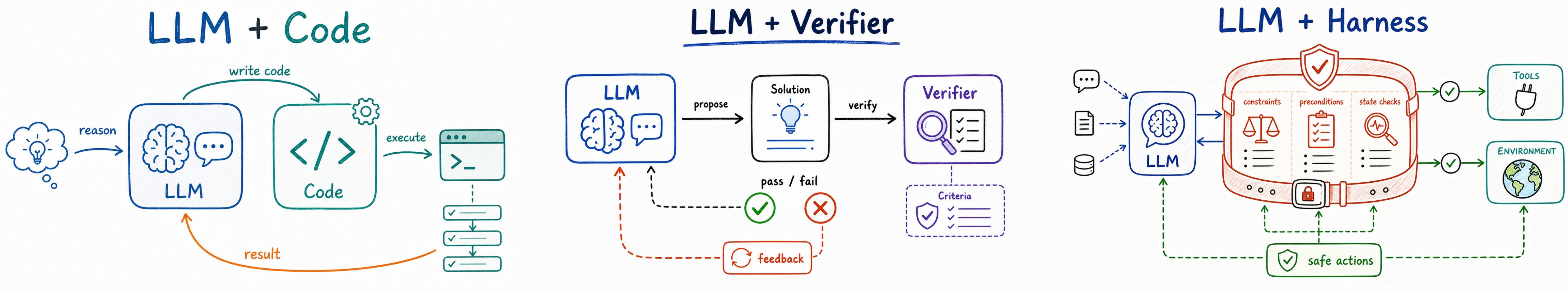



LLM通用感知与推理

| | | | |
|---|---|---|---|
| 感知能力 | 缺乏符号接地能力 | 神经网络的感知能力 | 高，几乎任意模态数据 |
| 推理能力 | 严谨符号逻辑推理 | 强化学习策略 | 基于语言的思维链推理 |
| 适应能力 | 依赖预先定义的符号空间 | 样本效率低，难以在不同场景迁移 | 大模型自身的通用性 |

# Is LLM All We Need?

单纯依赖 **LLM** 驱动的 **Agent** 面临长程规划不可靠、动作不可验证、经验难以复用等问题，将 **LLM** 与形式化语言、求解器、验证器、**Graph**等符号系统结合，已成为提升 **Agent** 可靠、可控性的重要共识



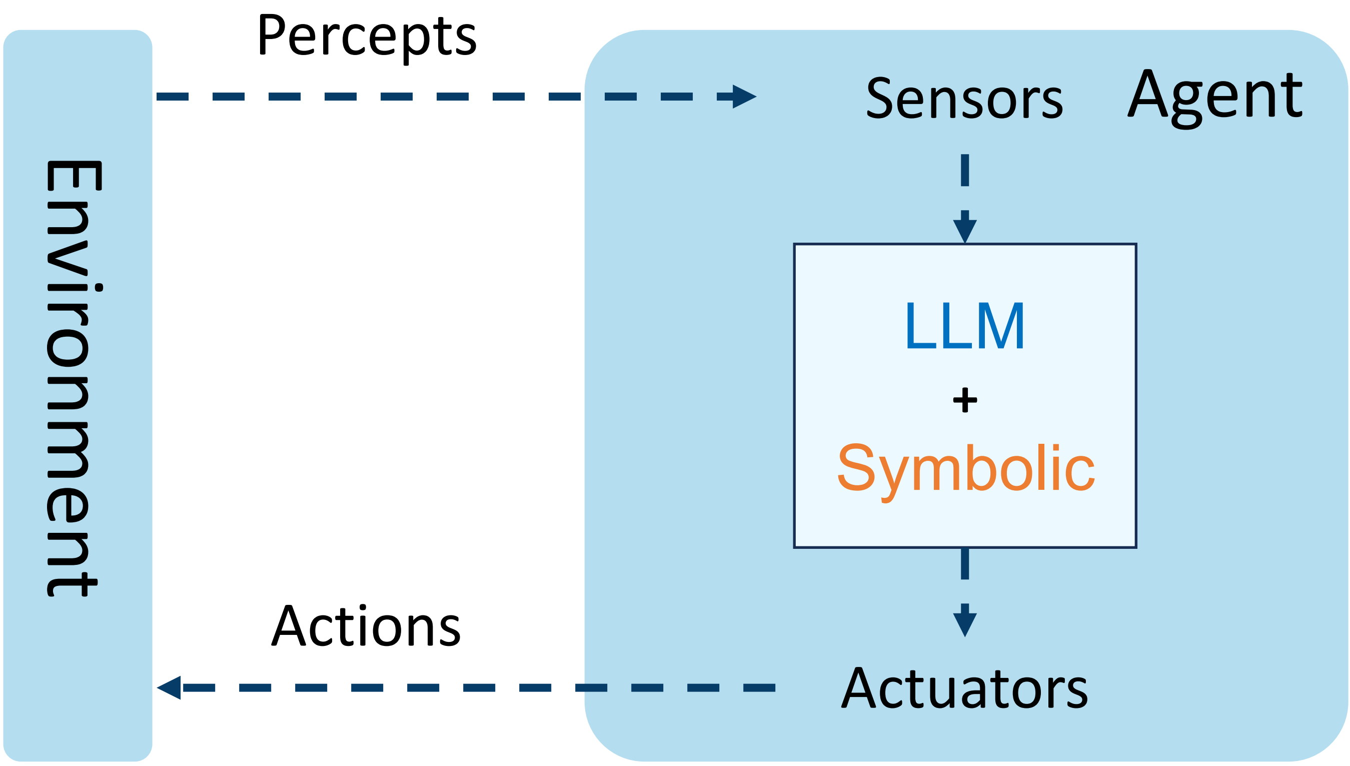

借助形式化语言的严谨推理能力

求解器、验证器的约束与反馈

从交互经验中归纳规则与技能

# Neuro-Symbolic Agent

**Neuro-Symbolic Agent**是在 **LLM-Agent**框架中，融合形式化逻辑推理、可验证的状态/动作约束、可归纳的规则技能、可组合的符号记忆等，使 **Agent** 从反应式决策走向可验证、可持续进化的可靠智能体




**Where Symbolic Helps**

- 辅助生成规划
- 验证动作执行
- 归纳可复用技能

**LLM + Symbolic = Powerful Planning, Reliable Execution**

# Outline

**Neuro-Symbolic Agent**

**for Travel Planning**

Symbolic helps planning

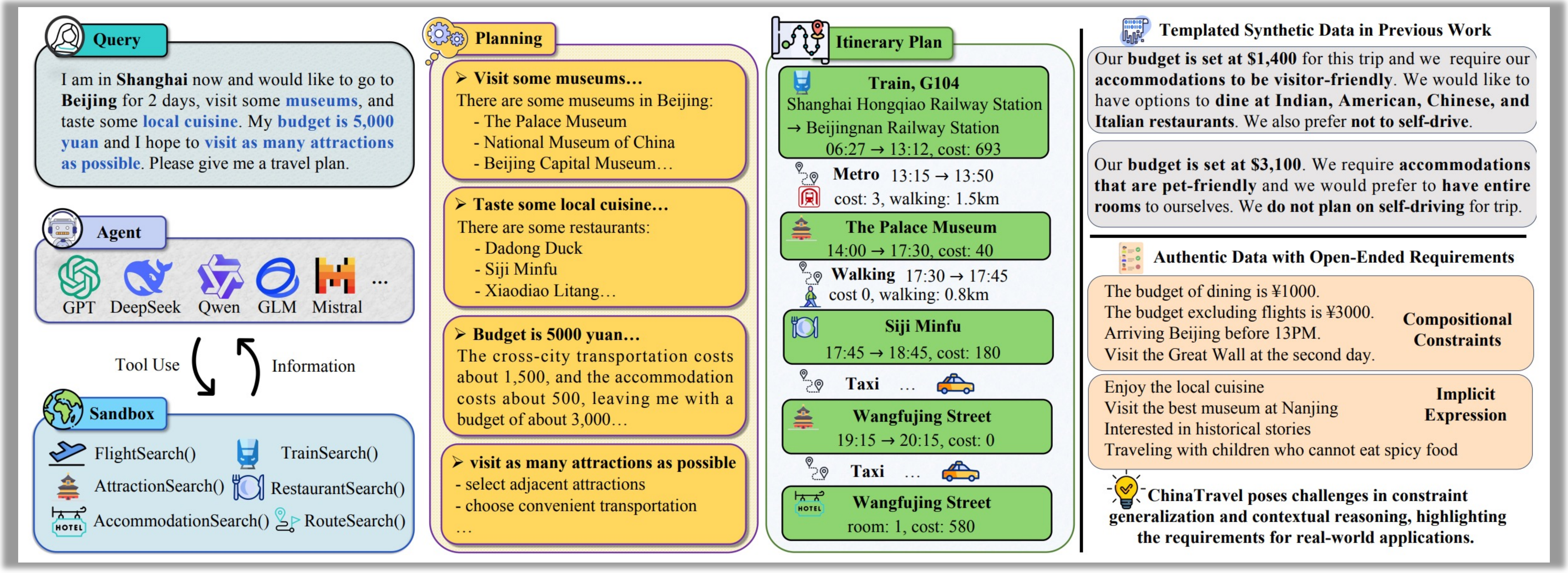
**for Embodied Planning**

Symbolic helps execute action

# Neuro-Symbolic Agent for Travel Planning

# Travel Planning Tasks

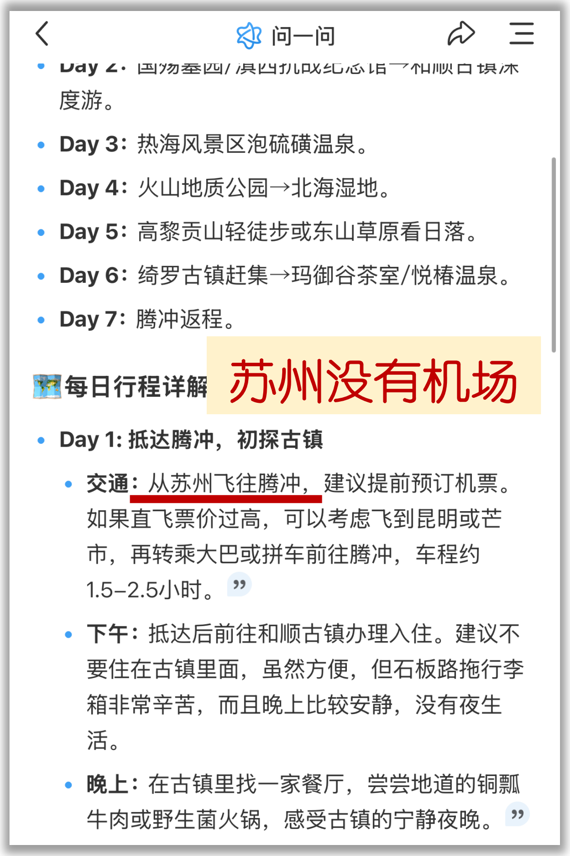
## 挑战：意图理解、工具调用、长程规划、约束满足

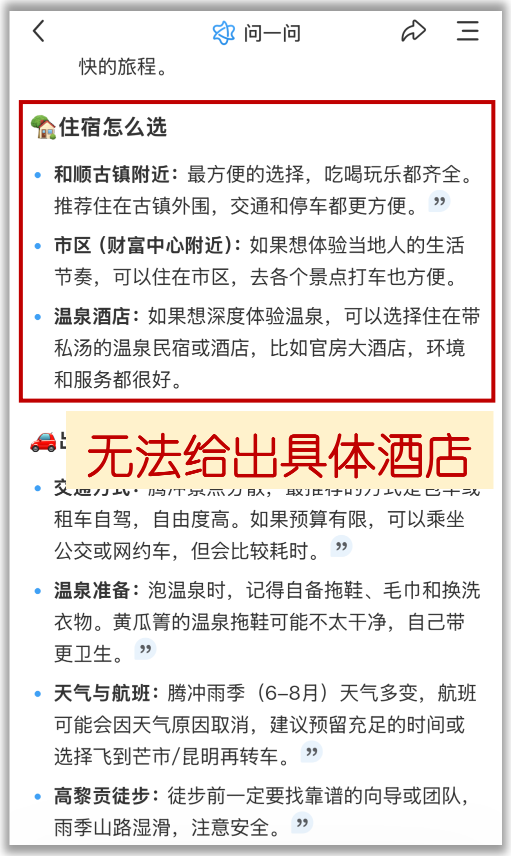


# Failure Case

用户：帮我规划一个苏州到云南腾冲、芒市5日游的规划

## 信息错误、仅能粗粒度规划

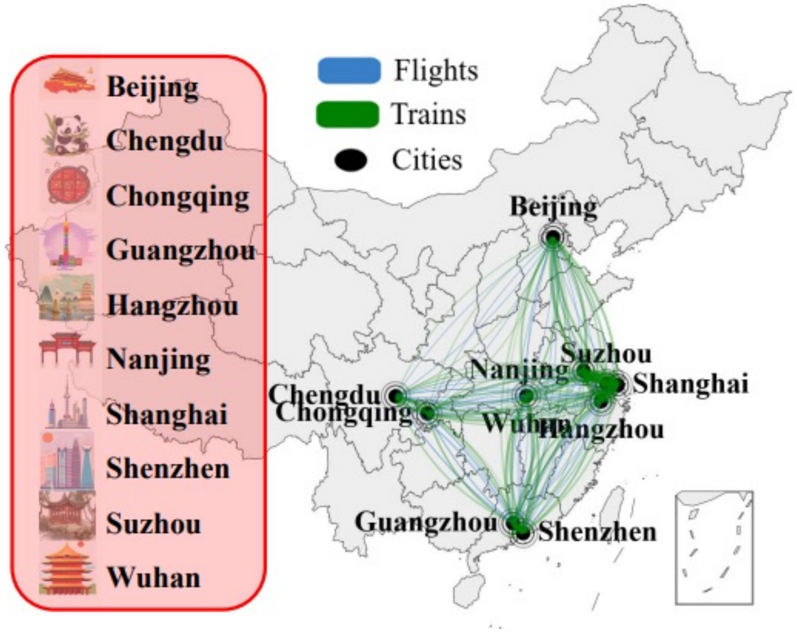


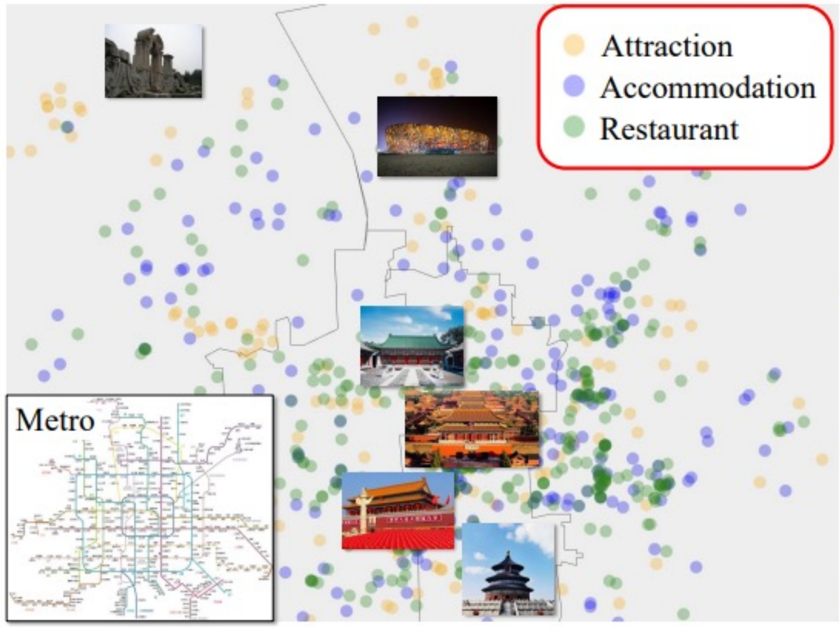



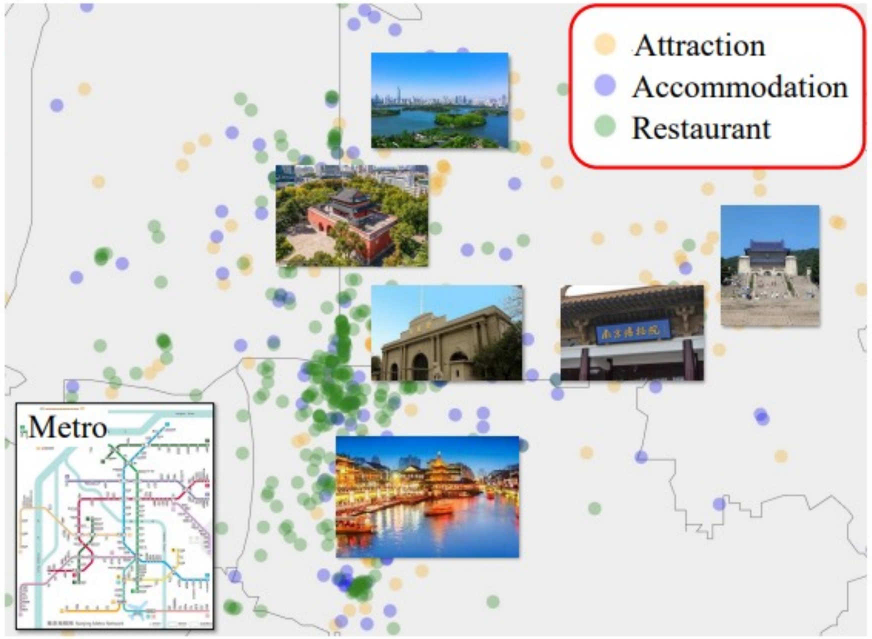

## 规划方案不可行、不合理


# ChinaTravel: An Open-Ended Travel Planning Benchmark

- GDP前10城市：上海、北京、深圳、重庆、广州、苏州、成都、杭州、武汉、南京
- 6.5K+ 跨城交通：720架次飞机、5770班次火车……
- 3K+ 景点：人文景观、自然景观、美术馆、博物馆、纪念馆、公园、游乐园……
- 4K+ 宾馆：洗衣服务、充电桩、停车场、宠物友好、会议厅、泳池、健身房……



# Real Travel Requirements

- **需求收集**：团队收集来自**300+人的真实自然语言需求**，构建了以16个符号概念为基础的约束需求
- **指标定义**：包含 **游、食、住、行**四个基础视角，构建了**23条可行性指标**，基本覆盖日常出游需要

| 信息类型 | 具体信息 |
|---|---|
| **基础信息** | 游玩天数、游玩人数、总开销 |
| **游** | 景点类型、景点名称 |
| **食** | 食物类型、餐馆名称、<br>人均单次餐饮价格 |
| **住** | 酒店特征、酒店名称、<br>房间个数、房间类型、酒店房间价格 |
| **行** | 市内交通方式、<br>城际交通方式、火车类型 |

| 可行性指标 | |
|---|---|
| **基础信息** | 游玩天数、游玩人数、总开销 |
| **游** | 选择景点在目标城市存在、在景点的开放时间进行访问、在旅程中访问景点不重复、景点的门票信息正确…… |
| **食** | 选择餐馆在目标城市存在、在餐馆的开放时间进行访问、在旅程中访问餐馆不重复、餐馆的人均价格信息正确…… |
| **住** | 选择酒店在目标城市存在、房间信息满足人数需求…… |
| **行** | 选择交通方式是存在的、给出交通方案的价格正确、给出交通方案的时间正确、给出交通方案的距离正确…… |
| **时间** | 给出的规划方案中每个事件按时间顺序发生，无冲突 |
| **空间** | 在不同地点的事件安排都需要保证给出可行的交通方式 |

# ChinaTravel: Domain-Specific Language (DSL)

## LLM友好、可组合、可验证

- **类 Python 语法**：可组合的原子概念（空间、时间、成本、类型……）
- **程序化验证**：可行性 ∧ 逻辑约束 ∧ 优先级排序
- **可组合泛化**：新约束可由少量基本概念组合而成
- **可扩展性强**：数百种人类需求约束，可通过约 20 个原子概念表达

判断餐饮预算是否超过1000

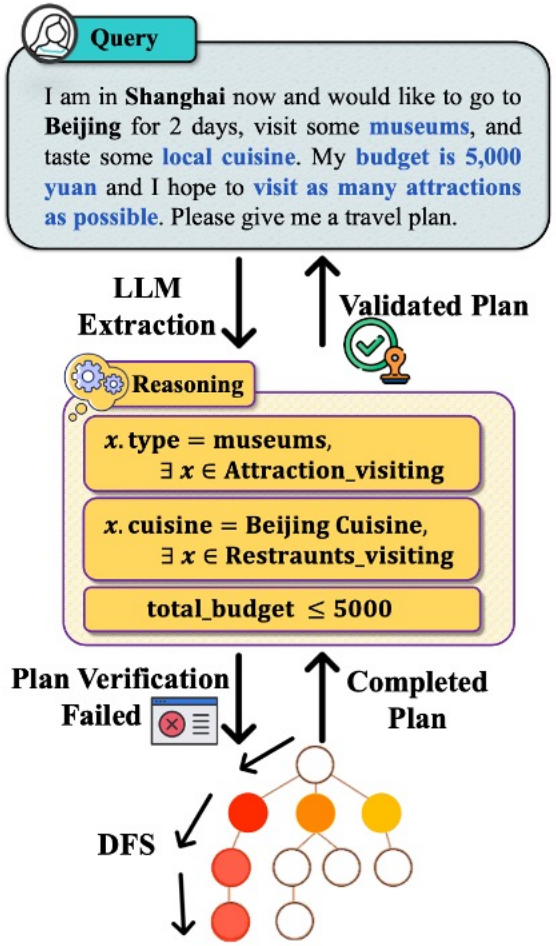
```
# Dining expenses <= 1000 CNY.
dining_cost = 0
for act_i in allactivities(plan):
  typ = activity_type(act_i)
  if typ=="breakfast" or typ=="lunch" or
        typ=="dinner": dining_cost =
      dining_cost + activity_cost(act_i)
return dining_cost <= 1000
```

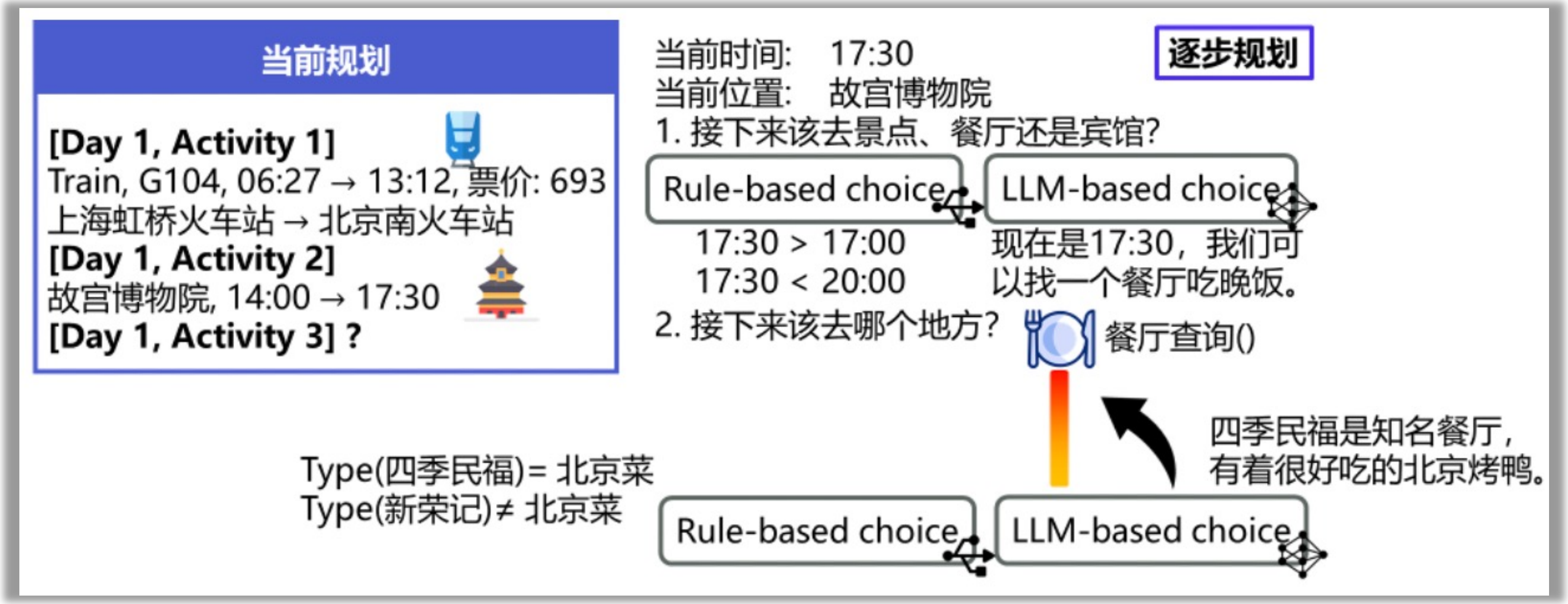
判断参观的景点总数量

```
# Arriving in Shanghai should be before
   6 PM on the second day.
return_time = 0
for act_i in day_activities(plan, 2):
  typ = activity_type(act_i)
  dest = transport_destination(act_i)
  if (typ=="train" or typ=="airplane")
      and des=="Shanghai": return_time
      == activity_endtime(act_i)
return return_time < "18:00"
```

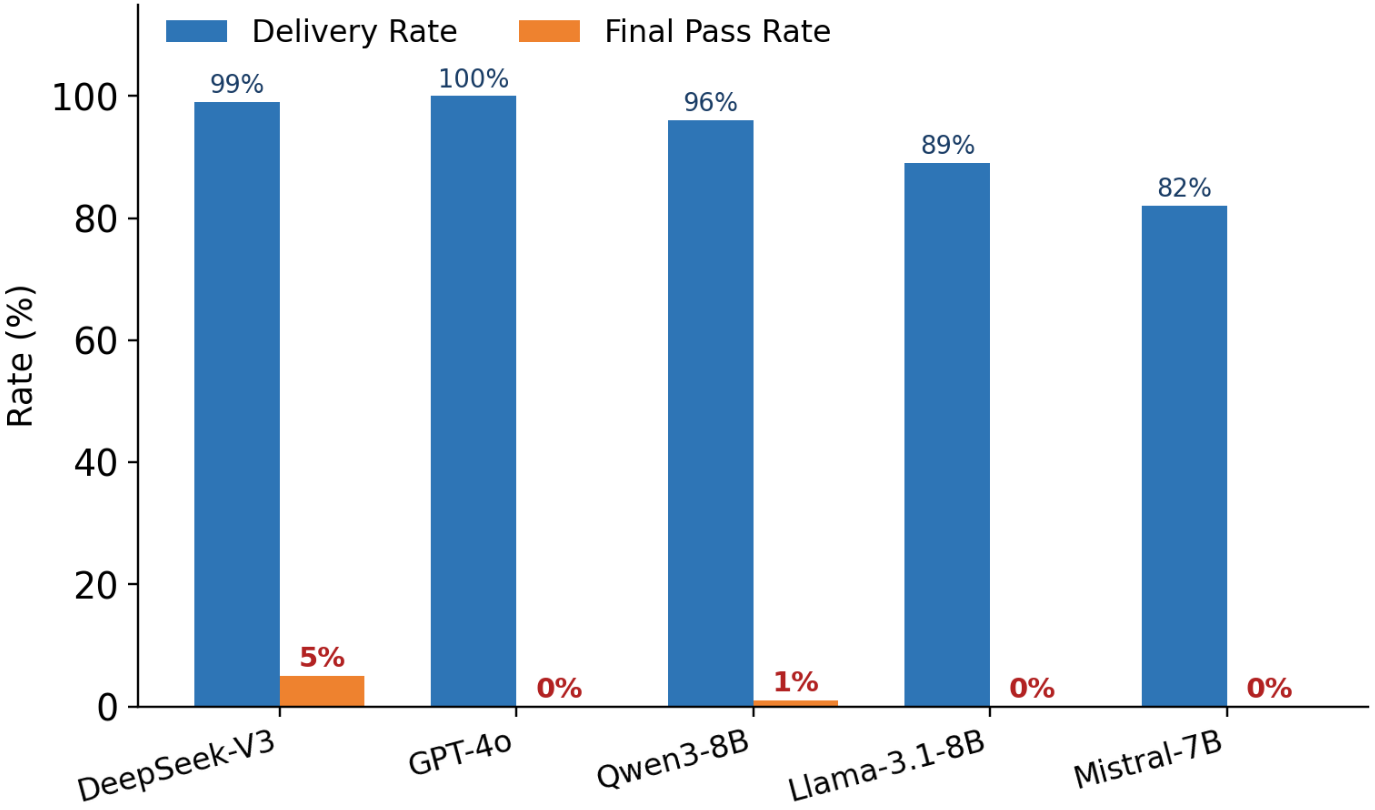
# Baseline Neuro-Symbolic Method

Natural Language -> DSL → LLM Guided Symbolic Search → DSL Verify LLM Revise



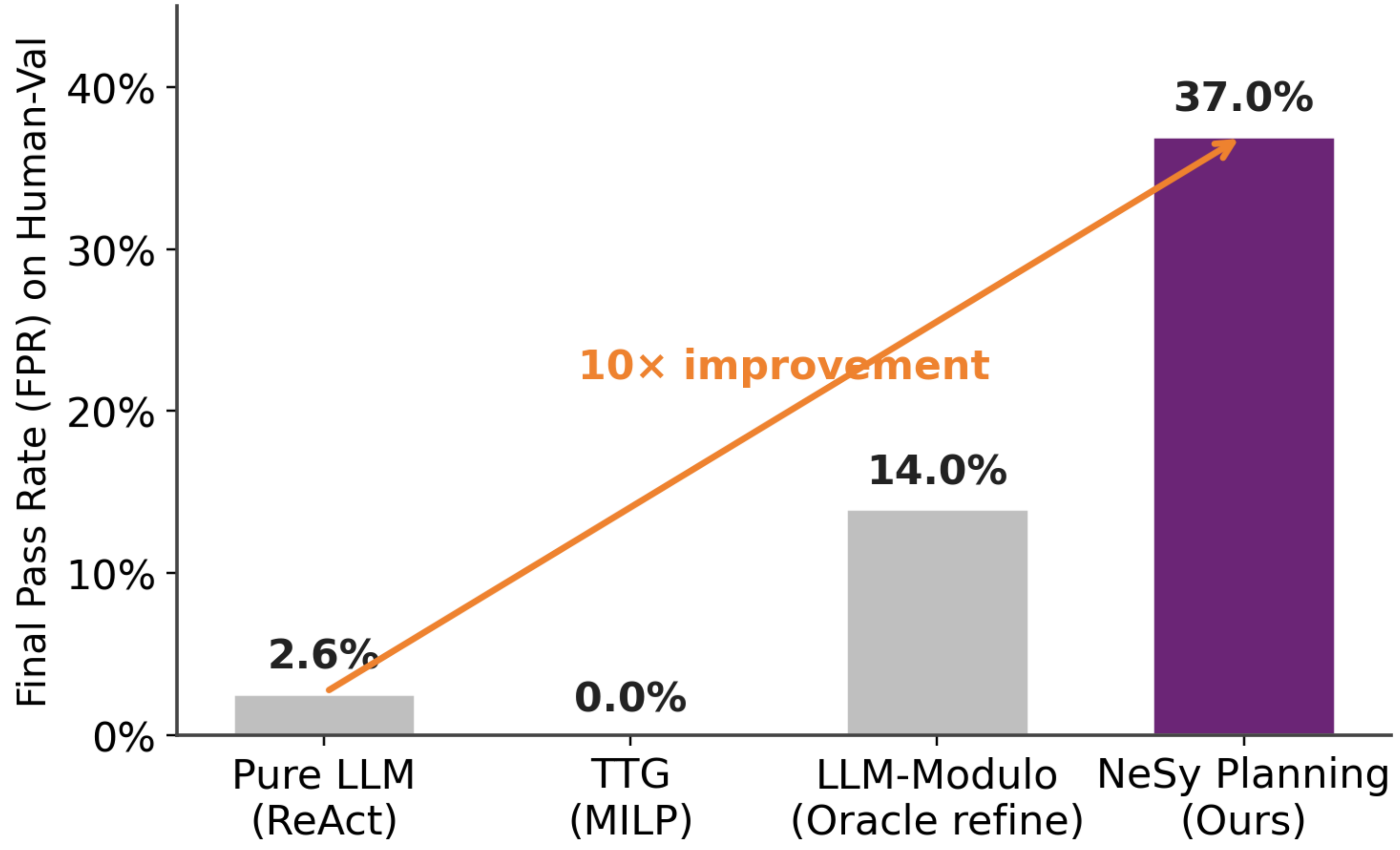
# Results of LLM Agent




- Structured output ≠ valid plan: **High DR masks low FPR**
- GPT-4o: **0% FPR**
- DeepSeek-V3 (ReAct): 5% on Easy, **~0% on human splits**

**LLM能生成看似合理的规划，但在长程、多约束的真实任务上，约束满足率几乎为0！**

# Results of Neuro-Symbolic Agent



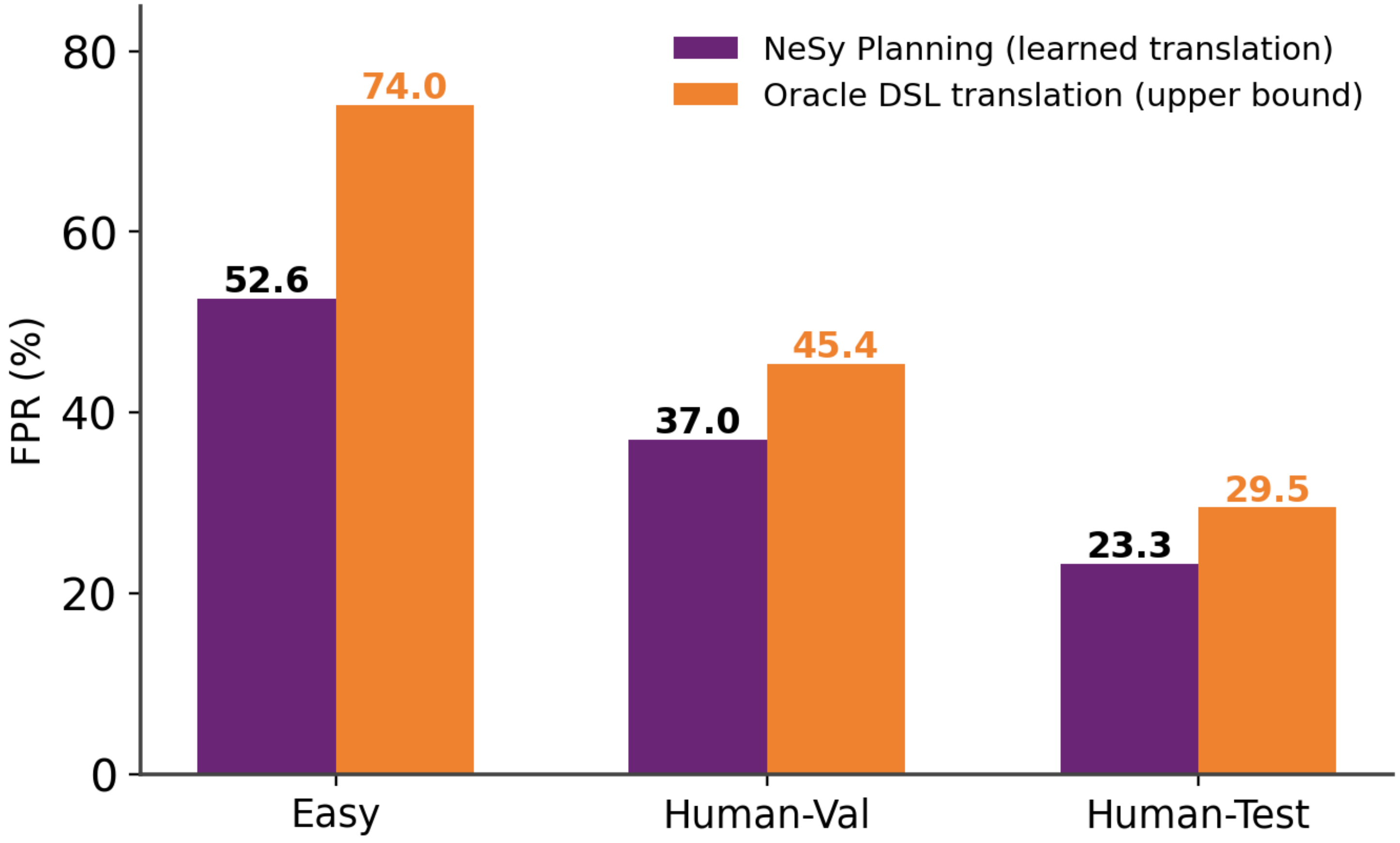

**2.6% → 37.0%**

**Pure LLM → NeSy Planning**

神经符号方法取得**15**倍性能提升

**观点** 让 **LLM** 做语言理解、让符号层处理长程搜索、让验证器闭环纠错，三者各司其职，效果远好于把所有事情都丢给 **LLM**

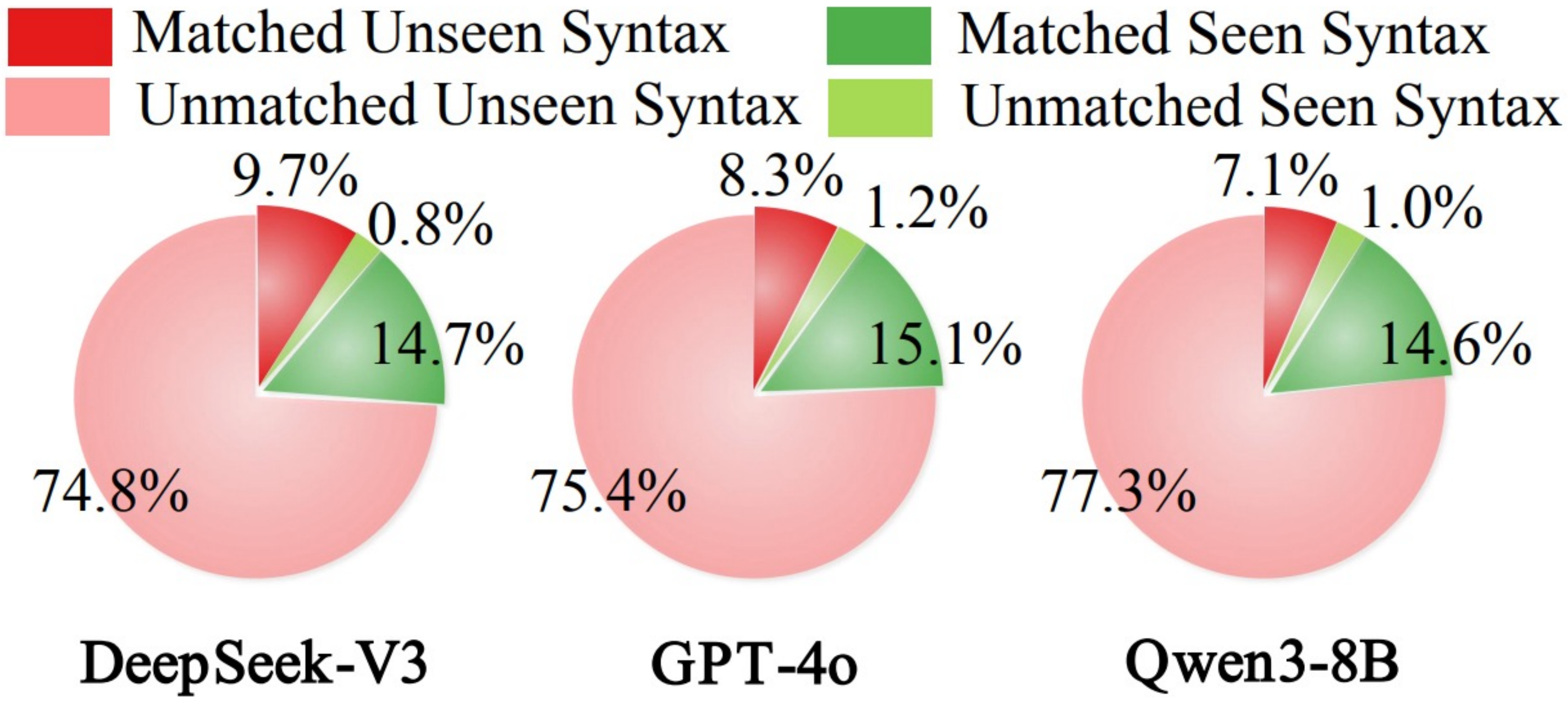
# Challenge




- 给出正确的**DSL**语言，成功率可显著提升
- **LLM**在理解开放用户需求并将其翻译成符号语言上仍存在挑战
- 即便给定标准的**DSL**语言，符号求解器约束通过率也并非**100%**，搜索效率仍然是挑战

# Compositional Generalization

在开放式旅行规划中，用户会把时间、预算、类型、偏好等基础概念组合为训练阶段未见过的新约束，**LLM**难以实现稳定的组合泛化



# TPC Challenge @ IJCAI 2206

## 第二届智能旅行规划挑战赛 @ IJCAI 2026 正式开赛！

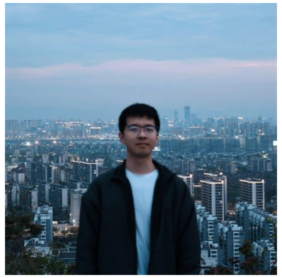
比赛官网

### Timeline

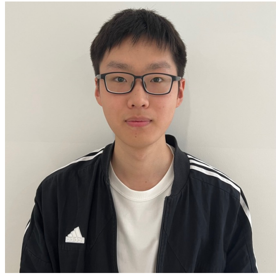
- 比赛正式启动：2026.06.01
- 第一阶段提交：2026.08.01
- 第二阶段评测：2026.08.01-07
- 技术报告提交：2026.08.07
- 最终结果公布：2026.08.10

https://chinatravel-competition.github.io/IJCAI2026

# Summary



Jie-Jing Shao



Bo-Wen Zhang

Jin Ye

- LLM对于长程规划、约束满足问题仍面临显著挑战
- 让LLM做语言理解、符号层处理长程搜索、验证器闭环纠错，效果更好
- 如何实现自然语言 -> 符号语言转换、如何提升组合泛化能力仍是挑战

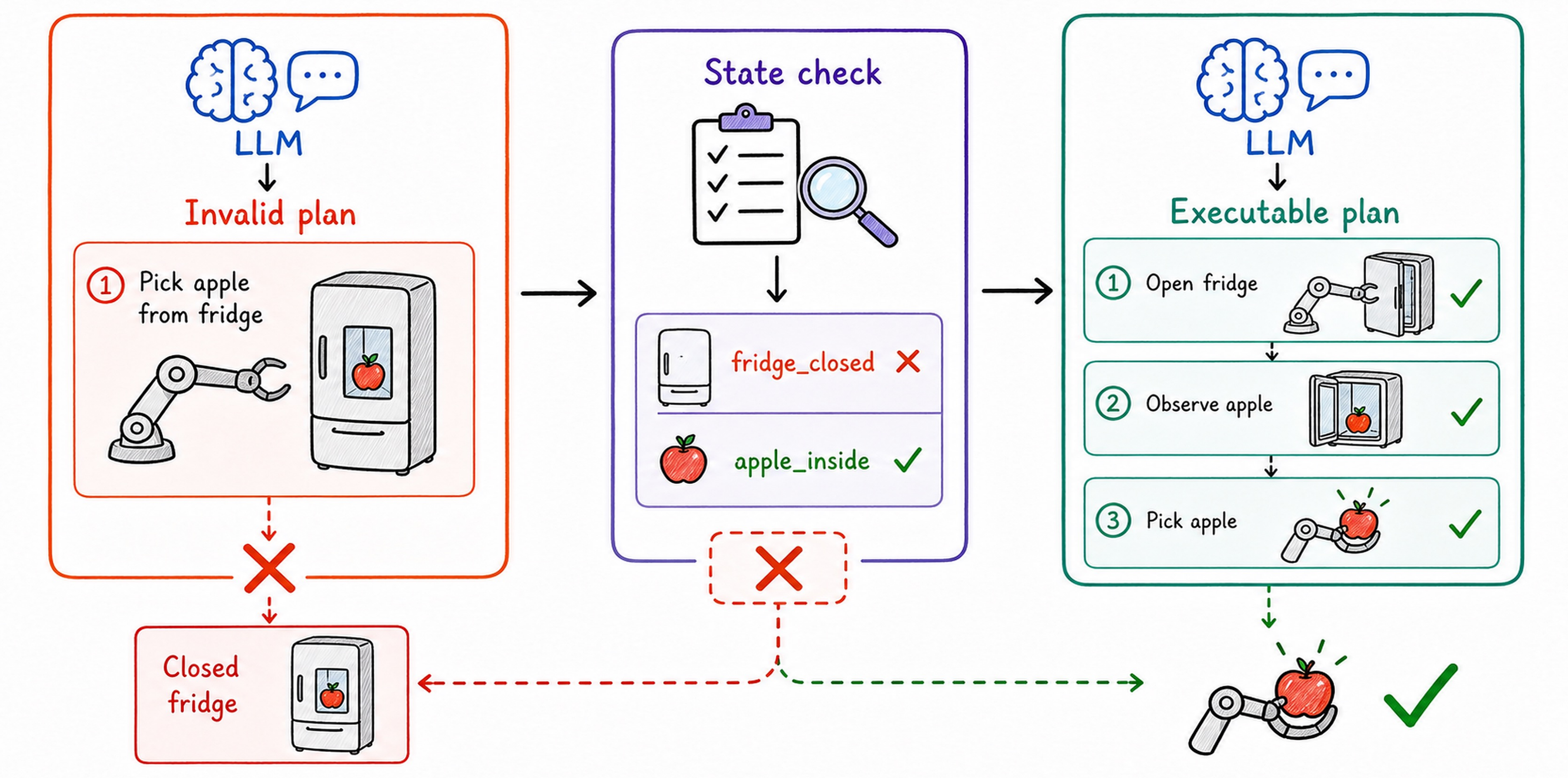
相关文献：


- Jie-Jing Shao, Bo-Wen Zhang, Xiao-Wen Yang, Bai-Zhi Chen, Si-Yu Han, Jing-Hao Pang, Wen-Da Wei, Guo-Hao Cai, Zhen-Hua Dong, Lan-Zhe Guo, Yu-Feng Li. **ChinaTravel: A Real-World Benchmark for Language Agent in Chinese Travel Planning**. **ICLR 2026**.
- Bo-Wen Zhang, Jin Ye, Jie-Jing Shao, Yu-Feng Li, Lan-Zhe Guo. **Mind the Gap to Trustworthy LLM Agents: A Systematic Evaluation on Constraint Satisfaction for Real-World Travel Planning**. **AAAI 2026 Trust Agent Workshop Best Student Paper Award**.
- Bo-Wen Zhang, Jin Ye, Jie-Jing Shao, Xiao-Wen Yang, Kun-Yang Yu, Zhi Zhou, Yu-Feng Li, Lan-Zhe Guo. **A Review of the IJCAI 2025 Travel Planning Challenge: Benchmarking LLM Agents with Compositional Constraints**. **Frontiers of Computer Science**.

# Neuro-Symbolic Agent for Embodied Planning

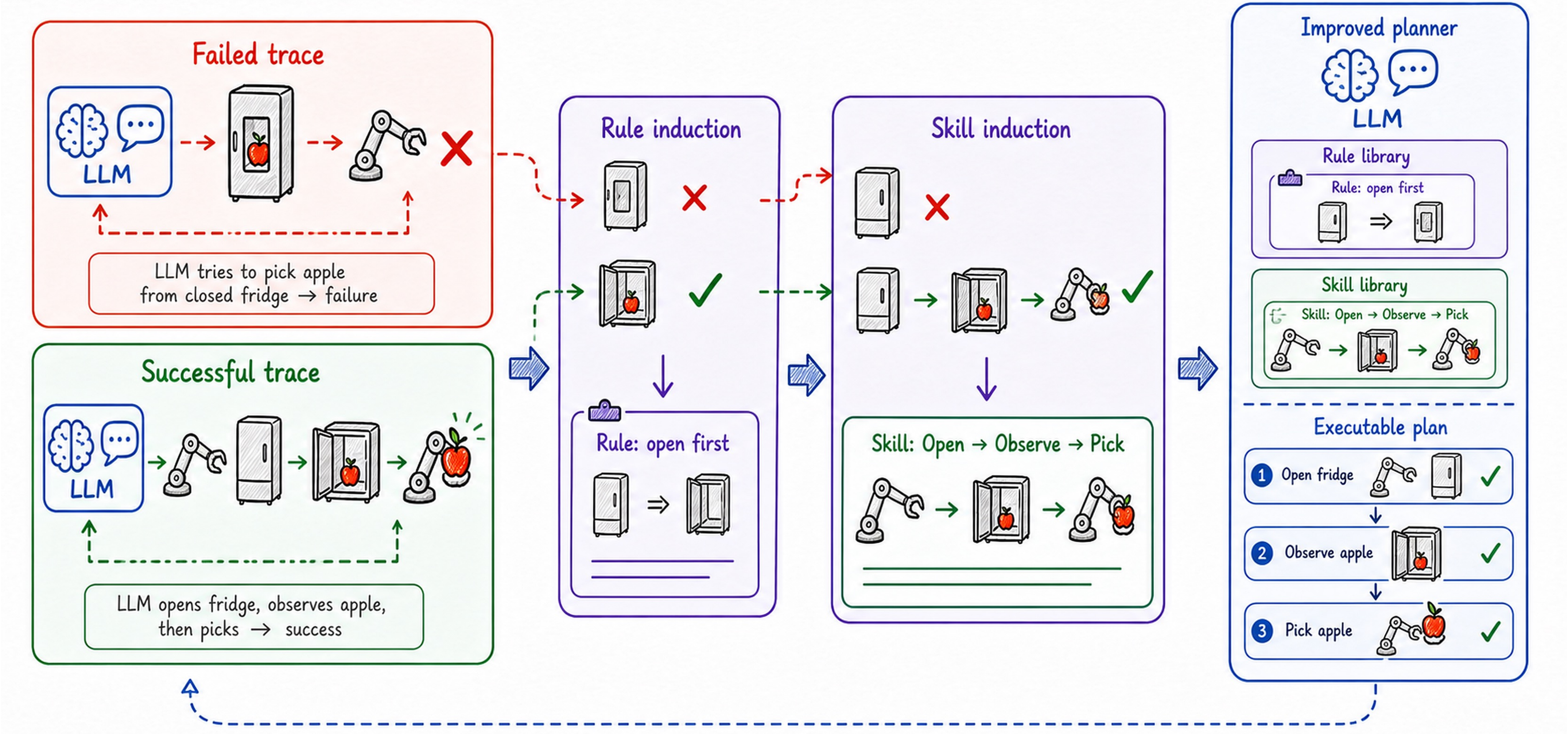
# Challenges of Embodied Planning

**Embodied Planning**：模型不仅需要生成合理的规划，还需要满足环境状态与物理约束，例如动作前提、可达性、**affordance**、安全约束等



# Rule/Skill Induction

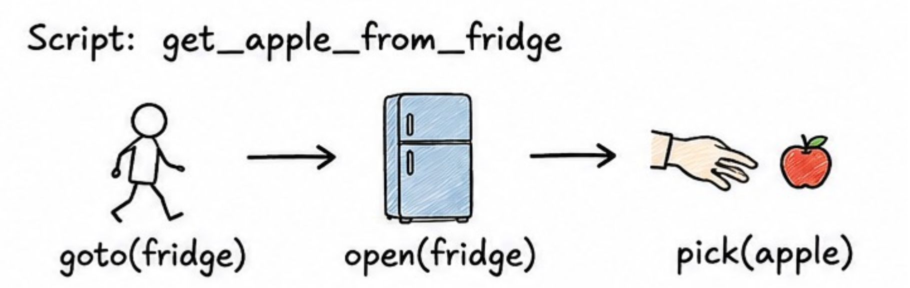
规则/技能学习：从成功与失败的交互轨迹中，归纳出可复用的规则与技能，进而约束LLM动作选择，指导生成可执行的规划



# Embodied Agent Skill

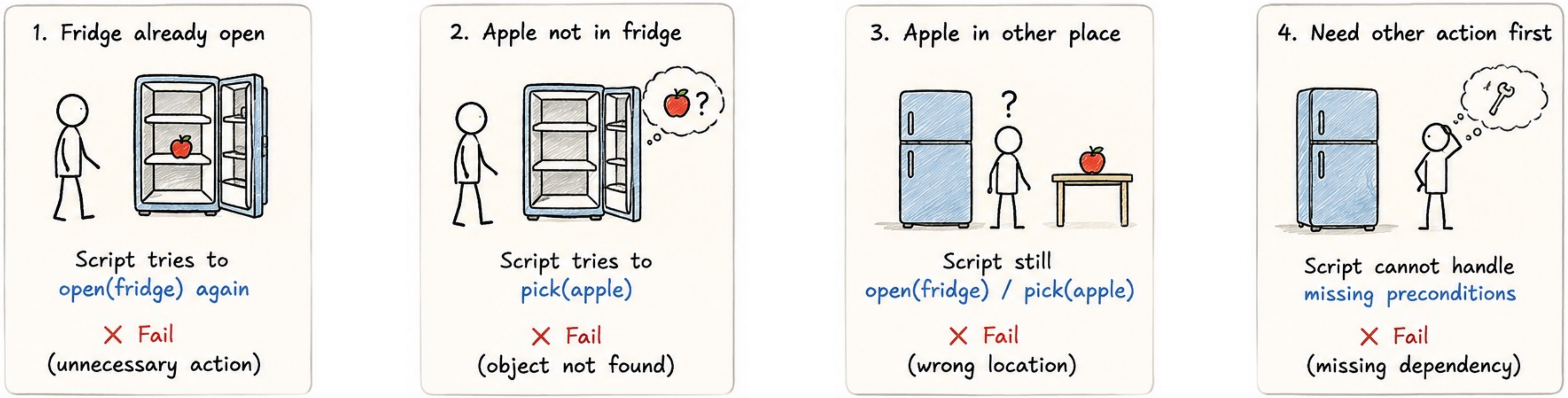
## 既有Skill通常仅依赖大模型归纳，表示形式为参数化的动作脚本

已有方法通常把交互轨迹压缩成**参数化动作脚本**



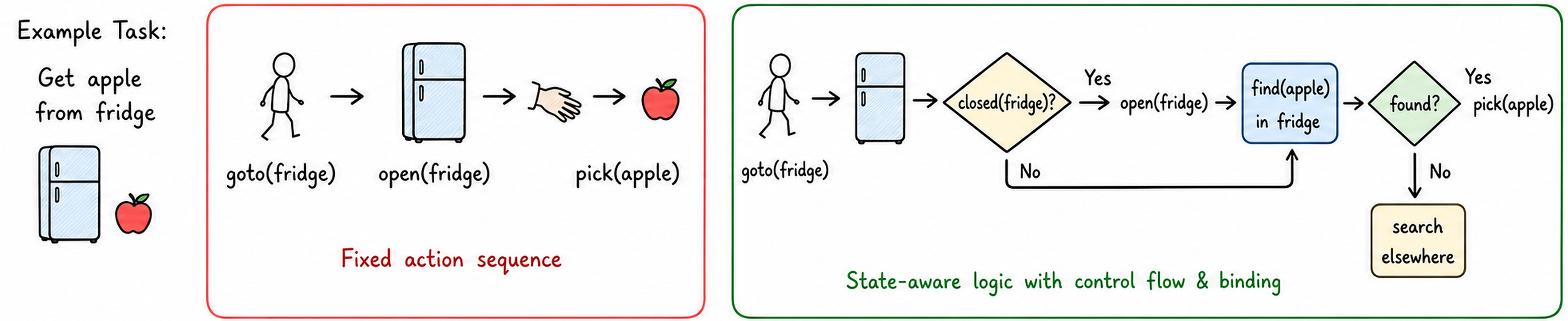

**State-Blind Skill**

- 缺少显式条件判断：不知道何时 open / skip
- 缺少动态变量绑定：目标对象依赖当前状态
- 环境偏移时容易失败，长程任务错误累积

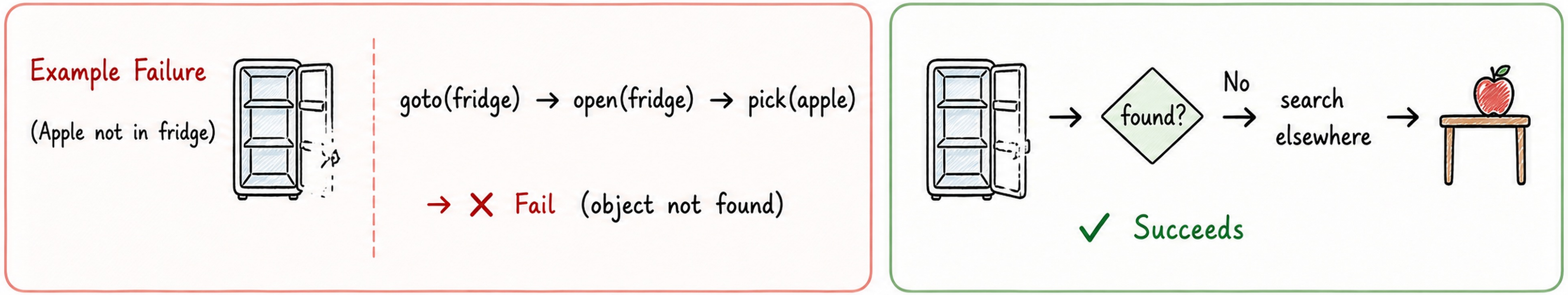


# Neuro-Symbolic Skill

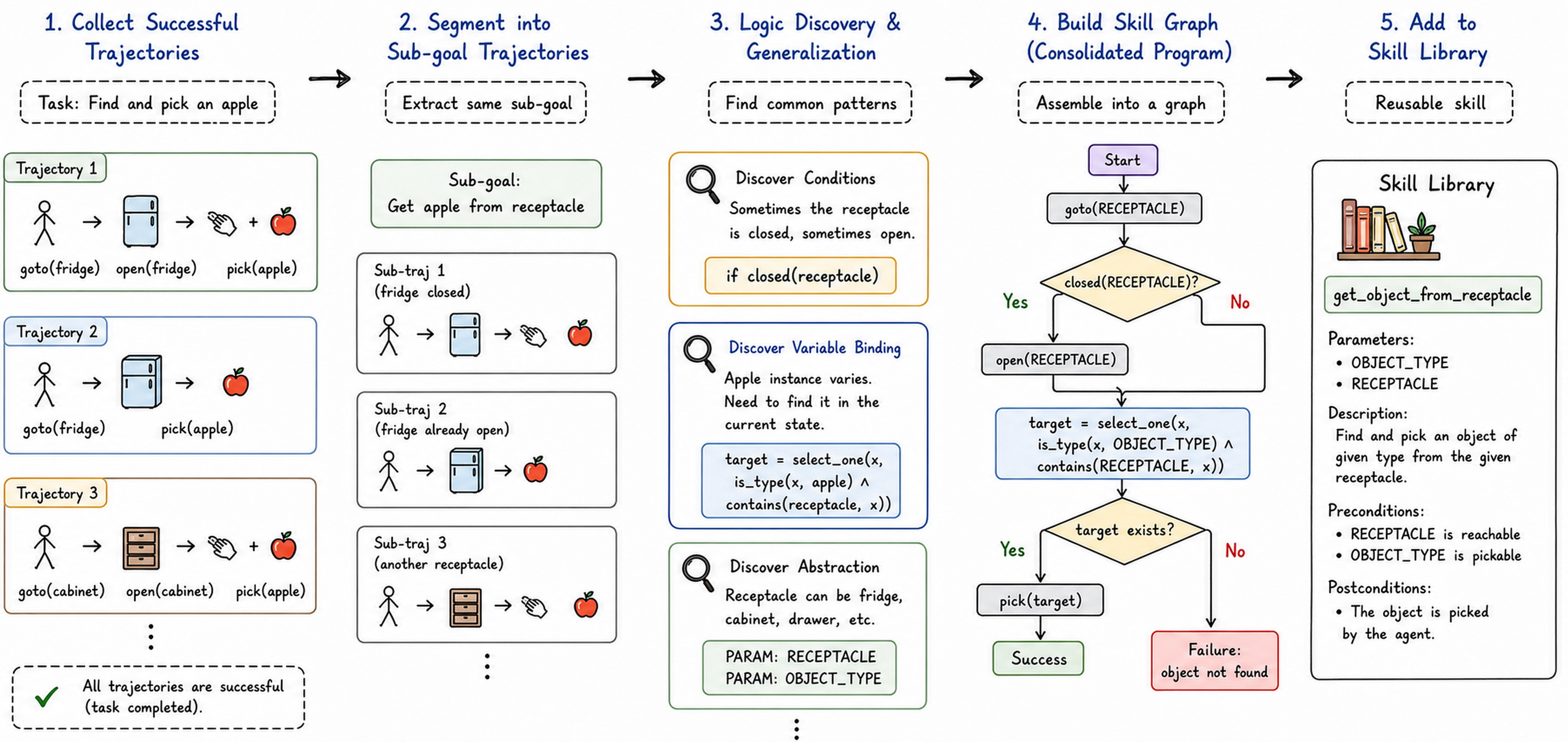
## Neuro-Symbolic Skill：“状态谓词 + 控制流 + 动态变量绑定”的可执行图



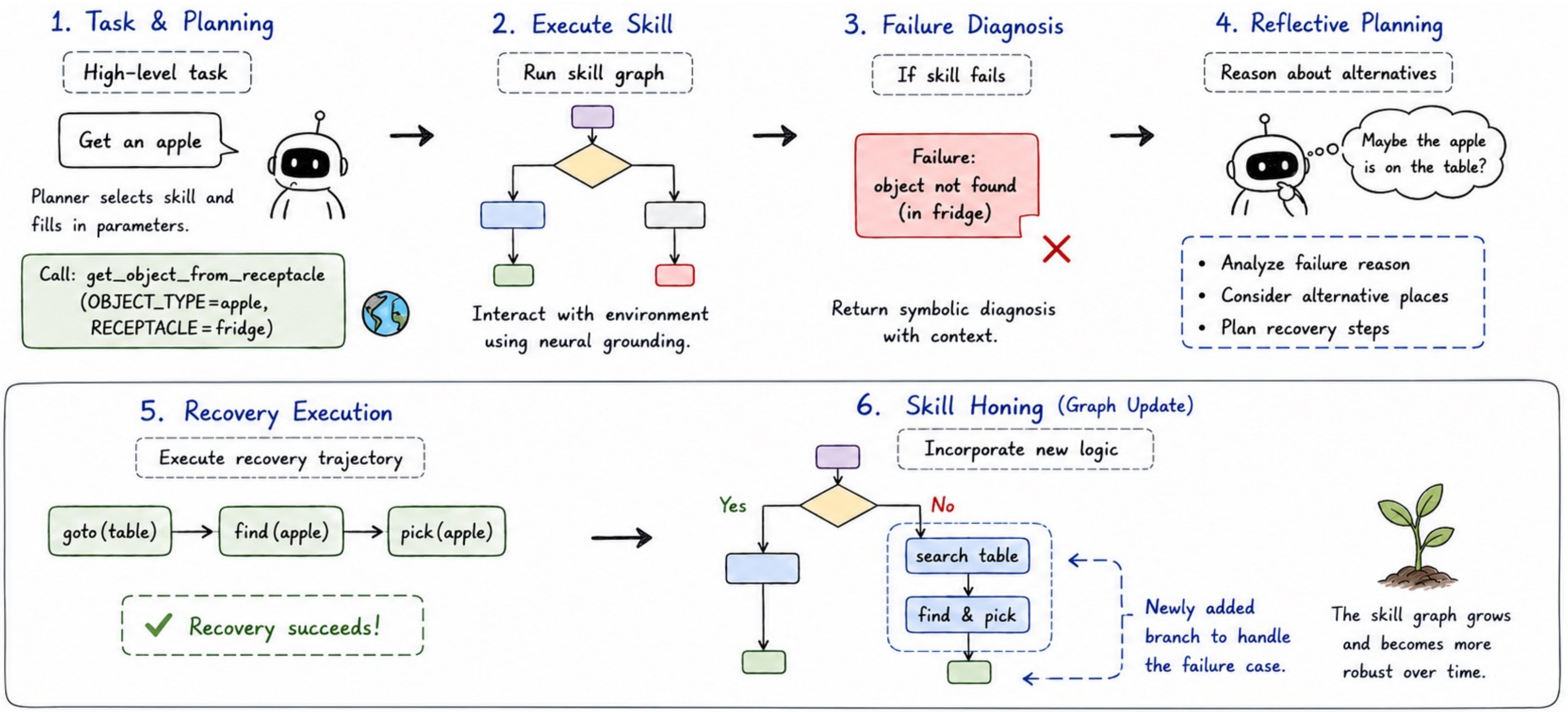

## Neuro-Symbolic Skill Learns When, Why, and How to act



# Offline Skill Induction from Successful Trajectories

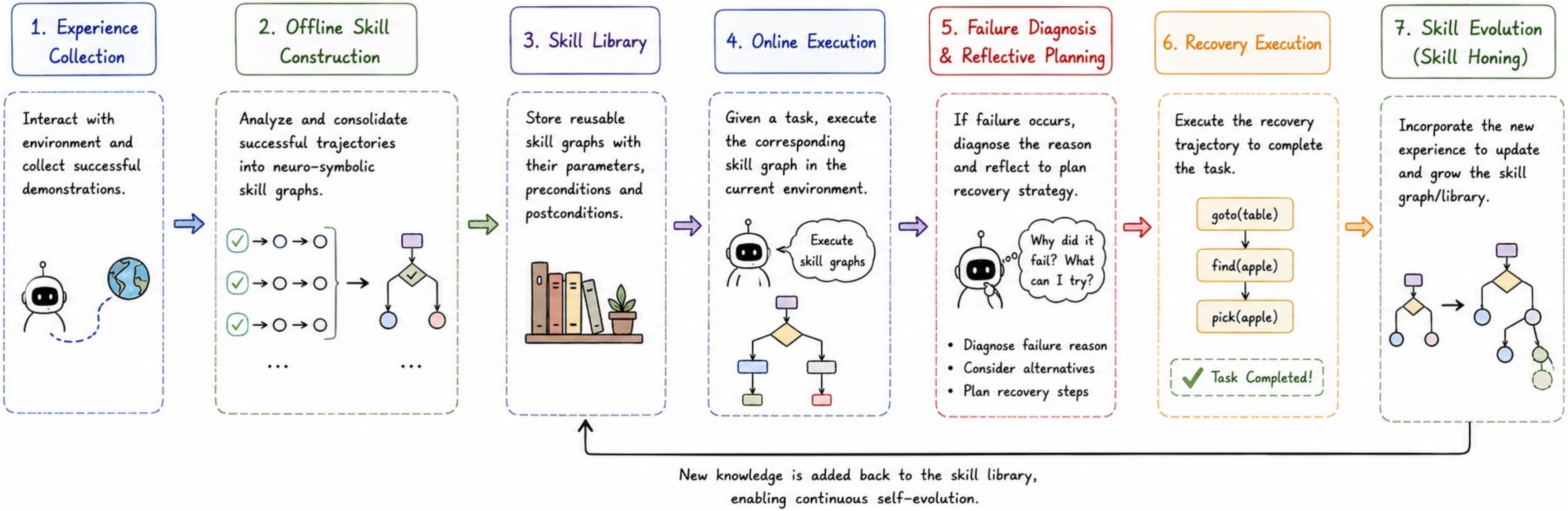


# Online Skill Evolution



# Neuro-Symbolic Skill

将智能体的成功交互轨迹从状态盲的动作脚本，提升为包含神经感知Grounding、符号变量绑定、显式控制流和失败修复机制的逻辑驱动技能图

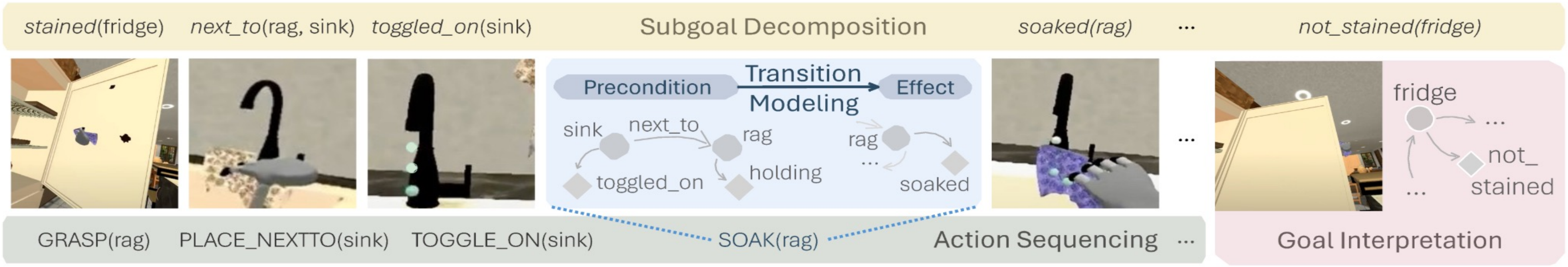


# Experimental Results

在Embodied Agent Interface Benchmark， 包含Vitural Home与Behavior两个场景，

均验证了所提 “失败-反思-规则提取-重规划” 机制的性能

Task: use the rag to clean the refrigerator



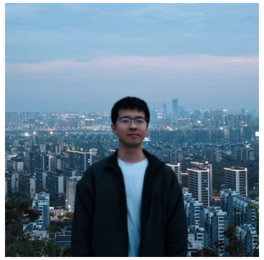

| Model | Overall Perf. | Average Perf. Module SR | | Goal Interpretation $F_1$ | | Action Sequencing Task SR | | Action Sequencing Execution SR | | Subgoal Decomposition Task SR | | Subgoal Decomposition Execution SR | | Transition Modeling $F_1$ | | Transition Modeling Planner SR | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | V | B | V | B | V | B | V | B | V | B | V | B | V | B | V | B |
| GPT-4o | 52.25 | 49.41 | 55.09 | 24.9 | 77.4 | 68.3 | 40.0 | 83.6 | 50.0 | 68.1 | 45.0 | 85.1 | 51.0 | 43.1 | 62.9 | 29.6 | 53.0 |
| GPT-o3 | 64.43 | 61.75 | 67.10 | 35.3 | 79.1 | 68.9 | 64.6 | 83.2 | 70.8 | 74.3 | 52.0 | 87.1 | 57.3 | 44.6 | 52.4 | 92.4 | 93.0 |
| GPT-5 | 68.32 | 68.09 | 68.55 | 36.3 | 79.2 | 72.3 | 71.0 | 84.9 | 74.0 | 73.2 | 54.0 | 87.2 | 60.0 | 36.4 | 43.0 | 86.3 | 97.0 |
| Re$^1$ Prompt | 74.54 | 71.58 | 77.50 | 49.0 | 84.9 | 71.2 | 73.0 | 84.9 | 78.0 | 74.7 | 80.0 | **88.0** | 86.6 | 54.8 | 43.0 | 91.4 | 97.0 |
| **Re$^2$ Agent** | **81.36** | **76.36** | **86.35** | **57.9** | **85.0** | **73.5** | **81.0** | **84.9** | **91.0** | **74.7** | **80.0** | 86.6 | **88.0** | **98.8** | **99.8** | **99.9** | **99.0** |

# Experimental Results

在ALFWorld、WebShop、TextCraft中的实验验证所提方法的有效性，
Online Evolution可进一步显著提升性能

| Method | ALFWorld Success Rate (%) | WebShop Score(%) | WebShop Success Rate (%) | TextCraft Success Rate (%) |
|---|---|---|---|---|
| ReAct (Yao et al., 2022b) | 85.8 | 44.0 | 20.0 | 62.0 |
| Reflexion (Shinn et al., 2023) | 84.3 | 40.8 | 23.0 | 59.0 |
| ADaPT (Prasad et al., 2024) | 67.9 | 45.8 | 29.0 | 72.5 |
| StateAct (Rozanov & Rei, 2025) | 84.3 | 40.6 | 8.00 | 83.0 |
| $\text{AWM}_{\text{offline}}$ (Wang et al., 2025c) | 88.6±1.1 | 51.8±1.7 | 32.2±1.3 | 66.3±2.0 |
| AWM (Wang et al., 2025c) | 91.3±0.8 | 49.2±1.9 | 30.0±2.0 | 92.5±3.6 |
| ASI (Wang et al., 2025b) | 70.6±1.9 | 7.7±1.7 | 7.50±3.0 | 77.8±1.8 |
| NSI w.o. online honing | 93.5±1.9 | 58.8±1.8 | 30.5±1.5 | 78.5±2.5 |
| NSI (Ours) | **98.0±1.2** | **76.5±1.2** | **44.5±1.5** | **95.2±0.8** |

# Summary

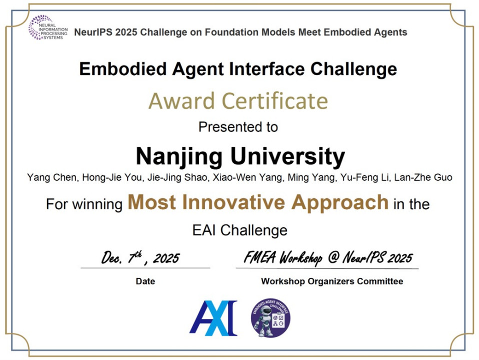


Jie-Jing Shao

Yang Chen

- 具身规划的输出必须满足**环境状态与物理约束**，能在环境中被真实执行
- 符号可以作为 LLM 规划与环境执行间的**verification layer**，**负责检查错误、诊断原因、指导修复**

**获斯坦福Fei-Fei Li教授主办的CVPR 2026 EmbodiedBench Challenge 冠军、NeurIPS 2025 Embodied Agent Challenge 最具创新算法奖**

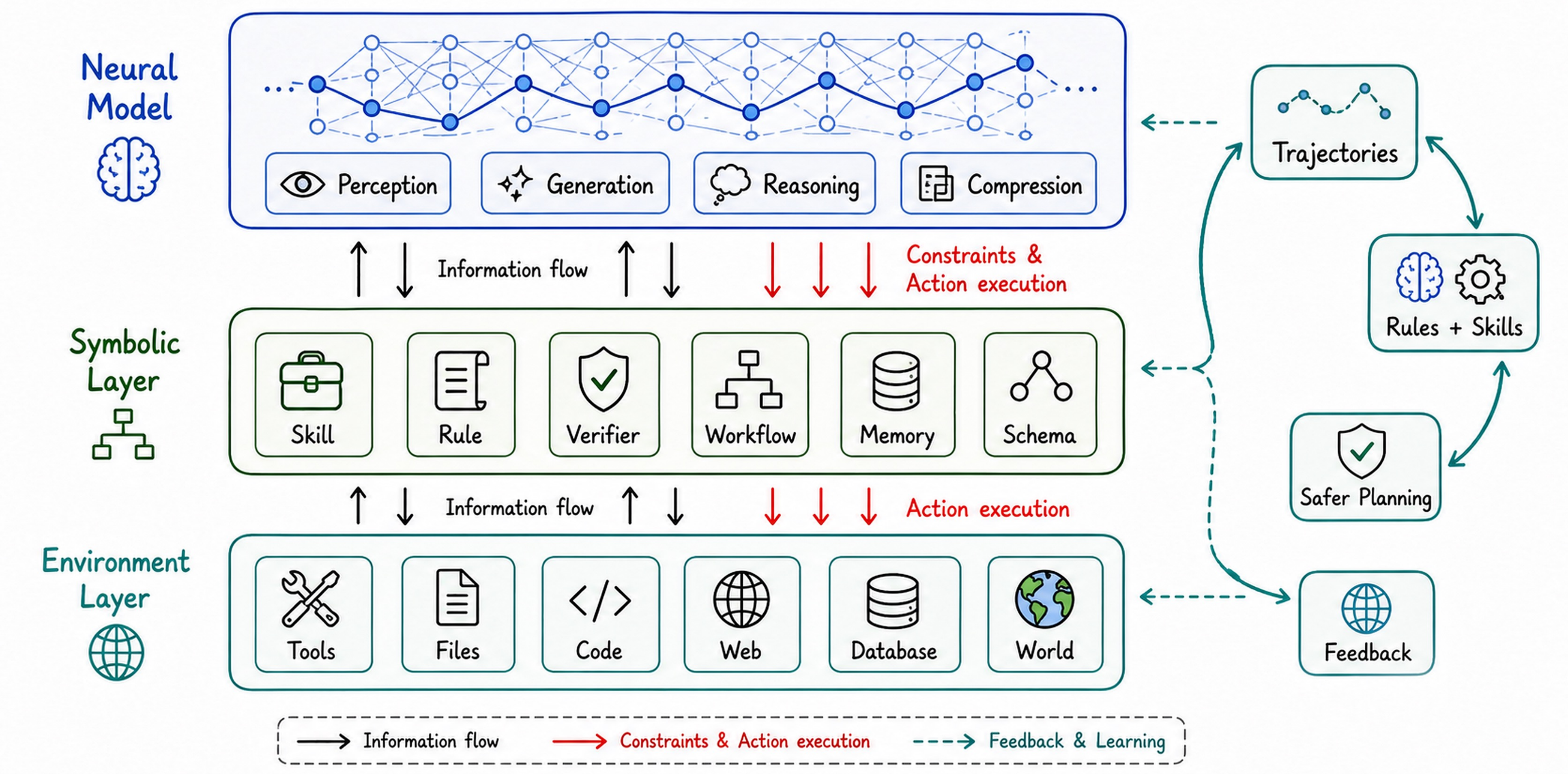

相关文献


- Jie-Jing Shao, Haiyan Yin, Yueming Lyu, Xingrui Yu, Lan-Zhe Guo, Ivor W. Tsang, James T. Kwok, Yu-Feng Li. **Lifting Traces to Logic: Programmatic Skill Induction with Neuro-Symbolic Learning for Long-Horizon Agentic Tasks**. **ICML 2026**.
- Yang Chen, Hong-Jie You, Jie-Jing Shao, Xiao-Wen Yang, Ming Yang, Yu-Feng Li, Lan-Zhe Guo. **Re2 Agent: Reflective Rule Induction and Rule-Guided Refinement for Embodied Planning**. **CVPR 2026 Foundation Model Meet Embodied Agents Workshop.**

# Takeaway

神经网络把经验压缩进权重；符号学习把经验压缩进可操作结构

**Agent 既需要大模型感知与推理能力，也需要维护可验证、可泛化的符号结构**



# Thank you!

郭兰哲

南京大学

guolz@nju.edu.cn www.lamda.nju.edu.cn/guolz