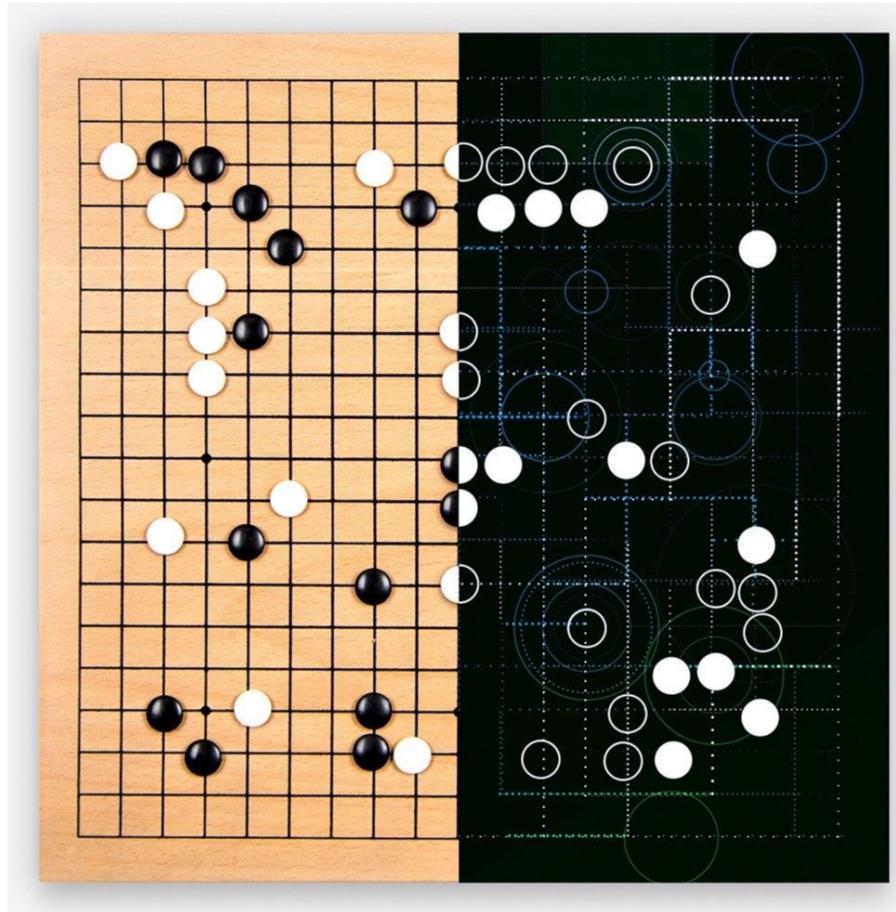


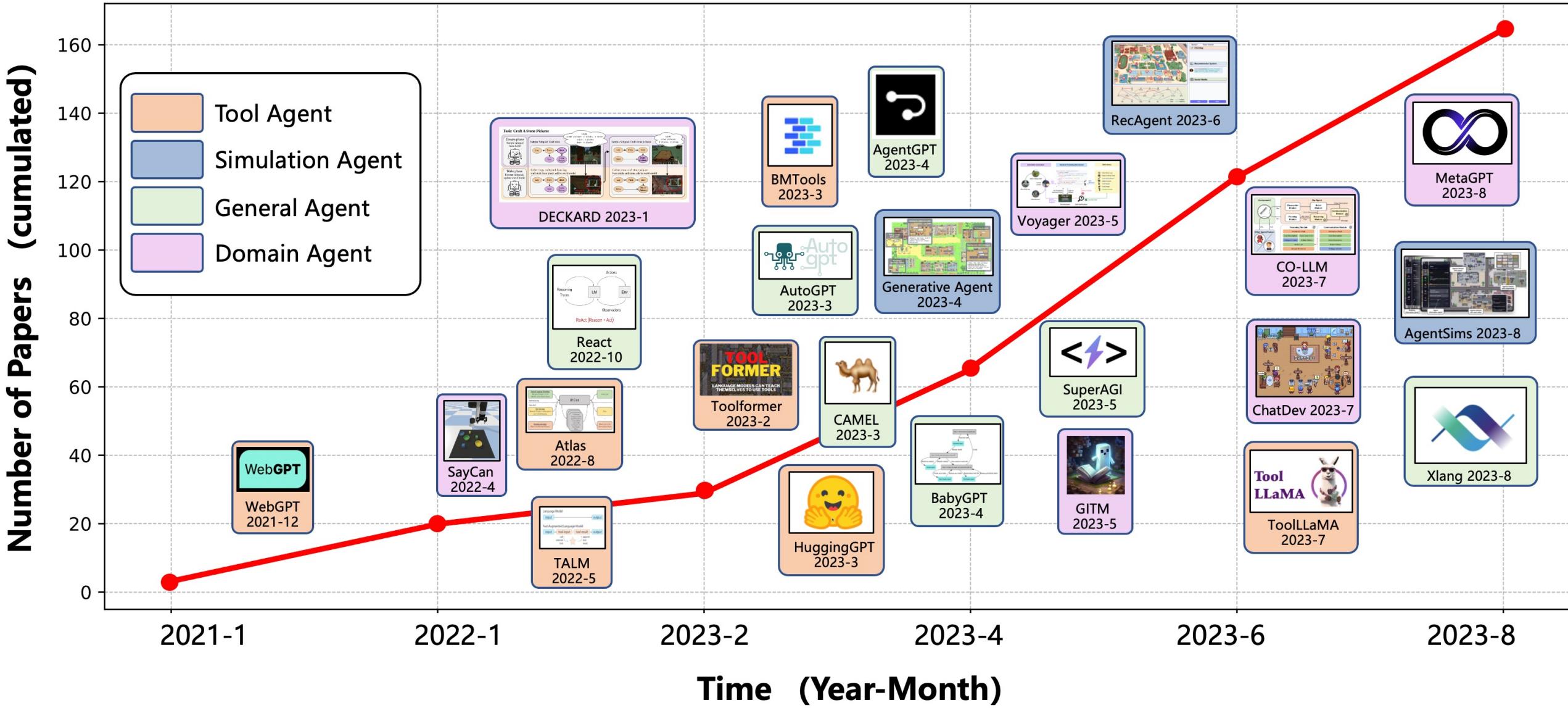
# LLM智能体

Scott

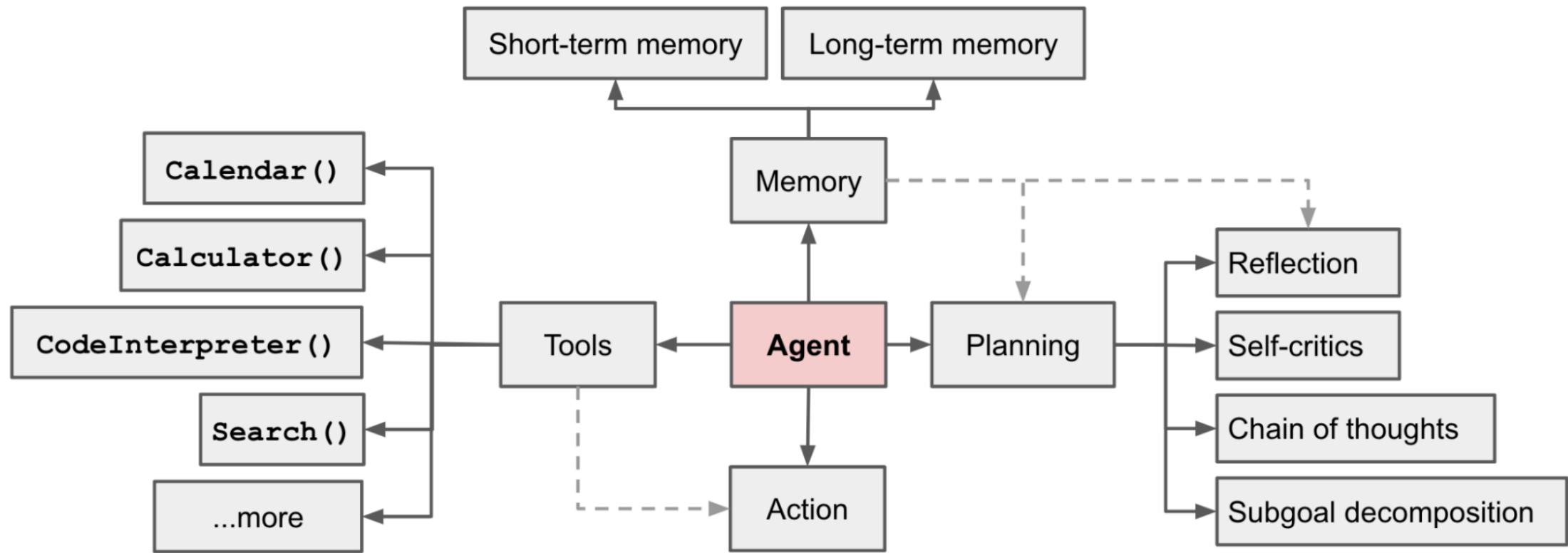
# Agent in RL



# LLM-Based Agent



# 核心



# Planning

任务规划

# 任务分解 Task Decomposition

- Chain of thought <https://lilianweng.github.io/posts/2023-03-15-prompt-engineering/#chain-of-thought-cot>
- Tree of thought <https://arxiv.org/abs/2305.10601>
- LLM+P <https://arxiv.org/abs/2304.11477>
- [Calibrating trust of multi-hop question answering systems with decompositional probes](#)
- [Interleaving retrieval with chain-of-thought reasoning for knowledge-intensive multi-step questions](#)

# 自我反思 Self-Reflection

- ReAct <https://arxiv.org/abs/2210.03629>

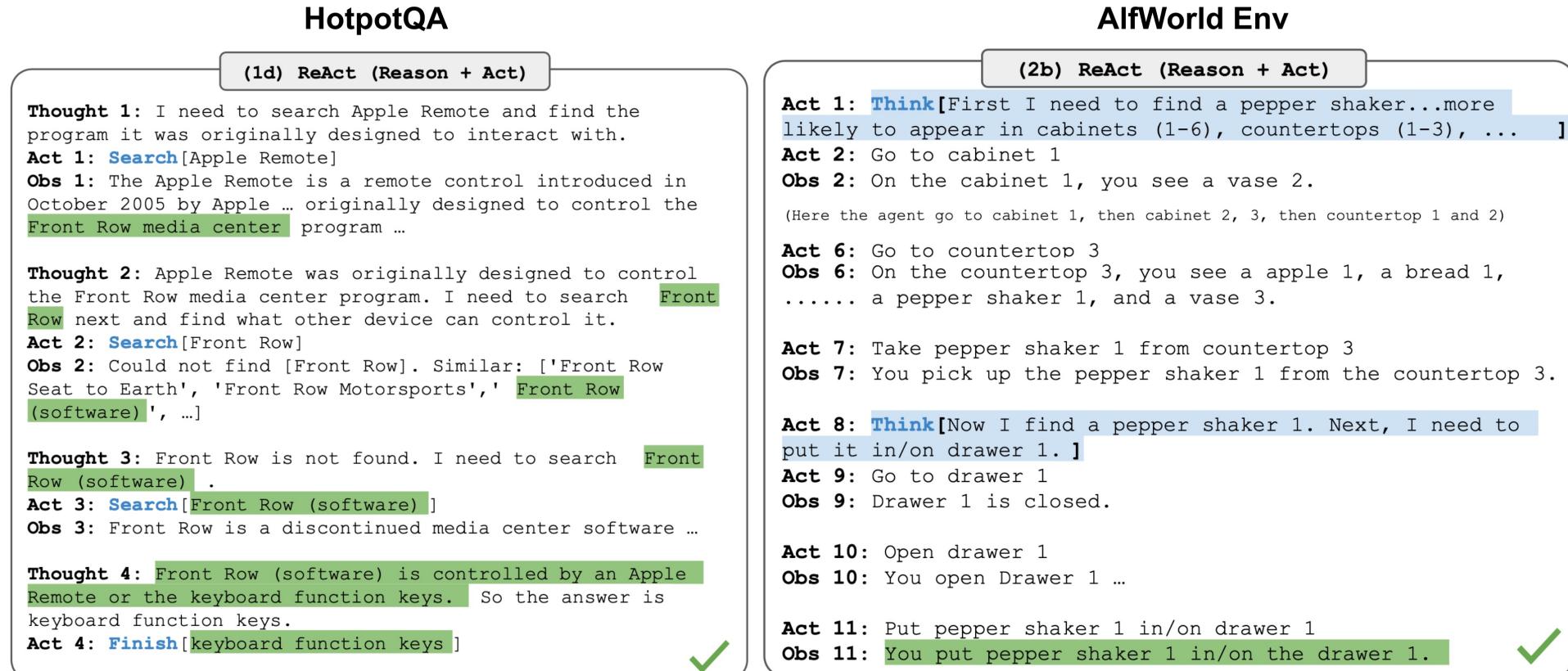
Thought: ...

Action: ...

Observation: ...

... (Repeat)

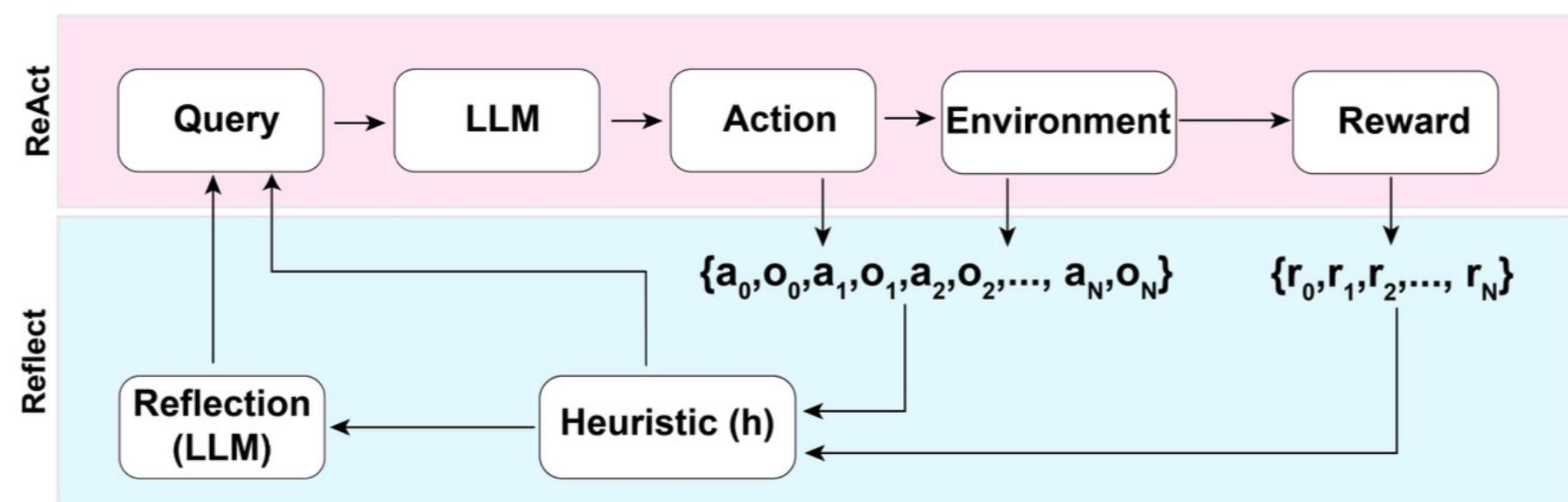
通過改進過去的  
行動決策和糾正  
錯誤來實現迭代  
改進



# 自我反思 Self-Reflection

- Reflexion <https://arxiv.org/abs/2303.11366>

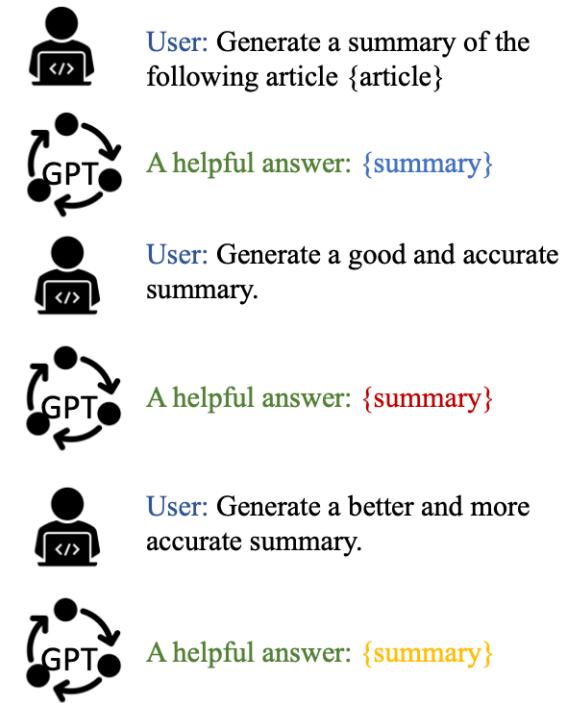
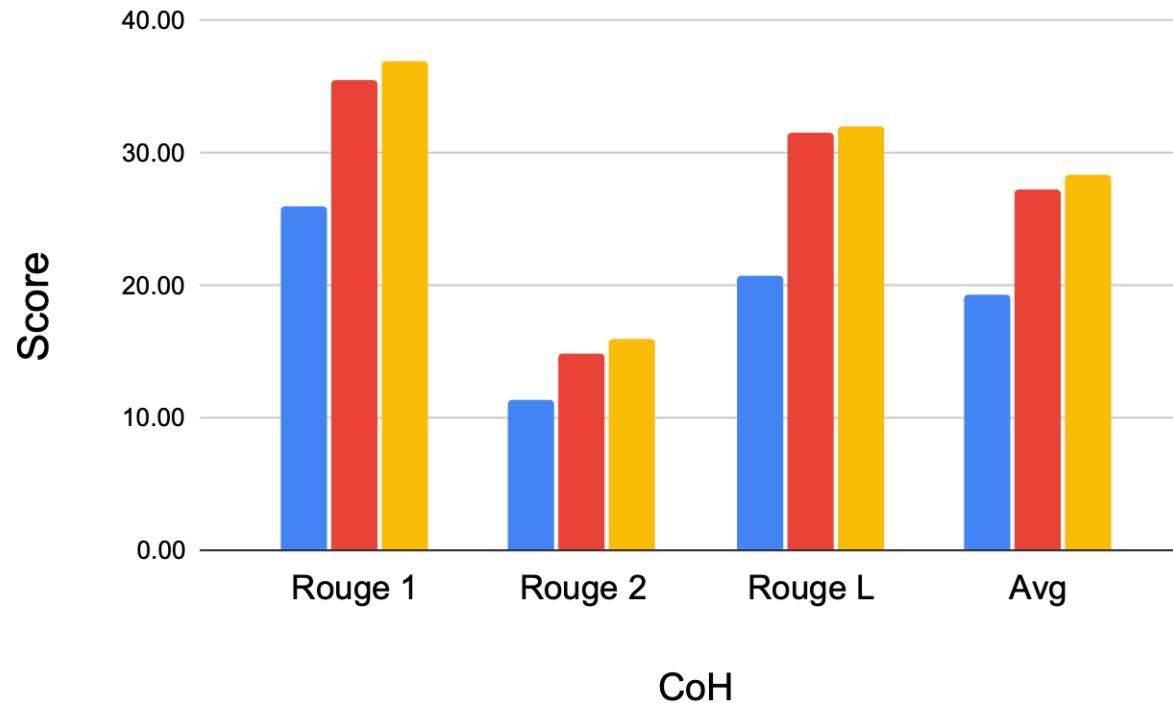
动态记忆和  
自我反思能力



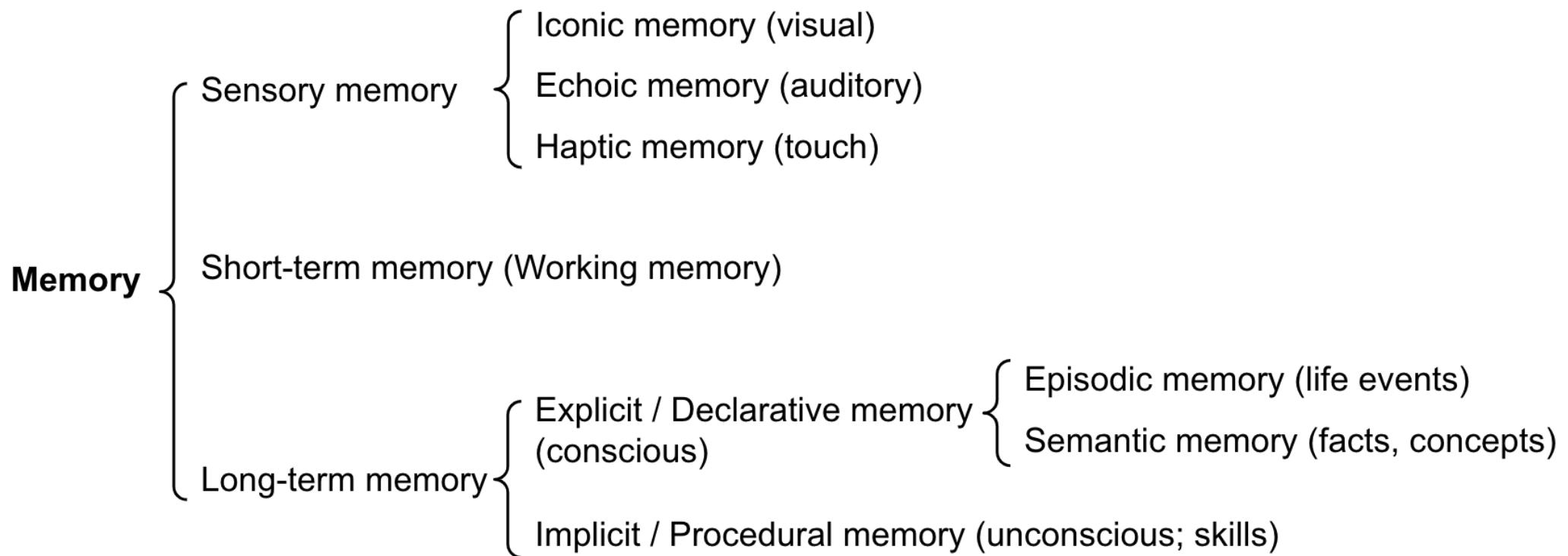
# 自我反思 Self-Reflection

- 回溯链Chain of Hindsight <https://arxiv.org/abs/2302.02676>

不断给予模型过去的表  
现反馈，包  
括评分和建  
议  
向模型展示  
一系列不斷  
進步的答案，  
然後訓練模  
型學習這種  
進步的趨勢，  
從而得到更  
好的答案



# Memory



# Tool Use

- MRKL <https://arxiv.org/abs/2205.00445>

Modular Reasoning, knoledge and language

外部符号工具

- TALM <https://arxiv.org/abs/2205.12255>

Tool Augmented Language Models

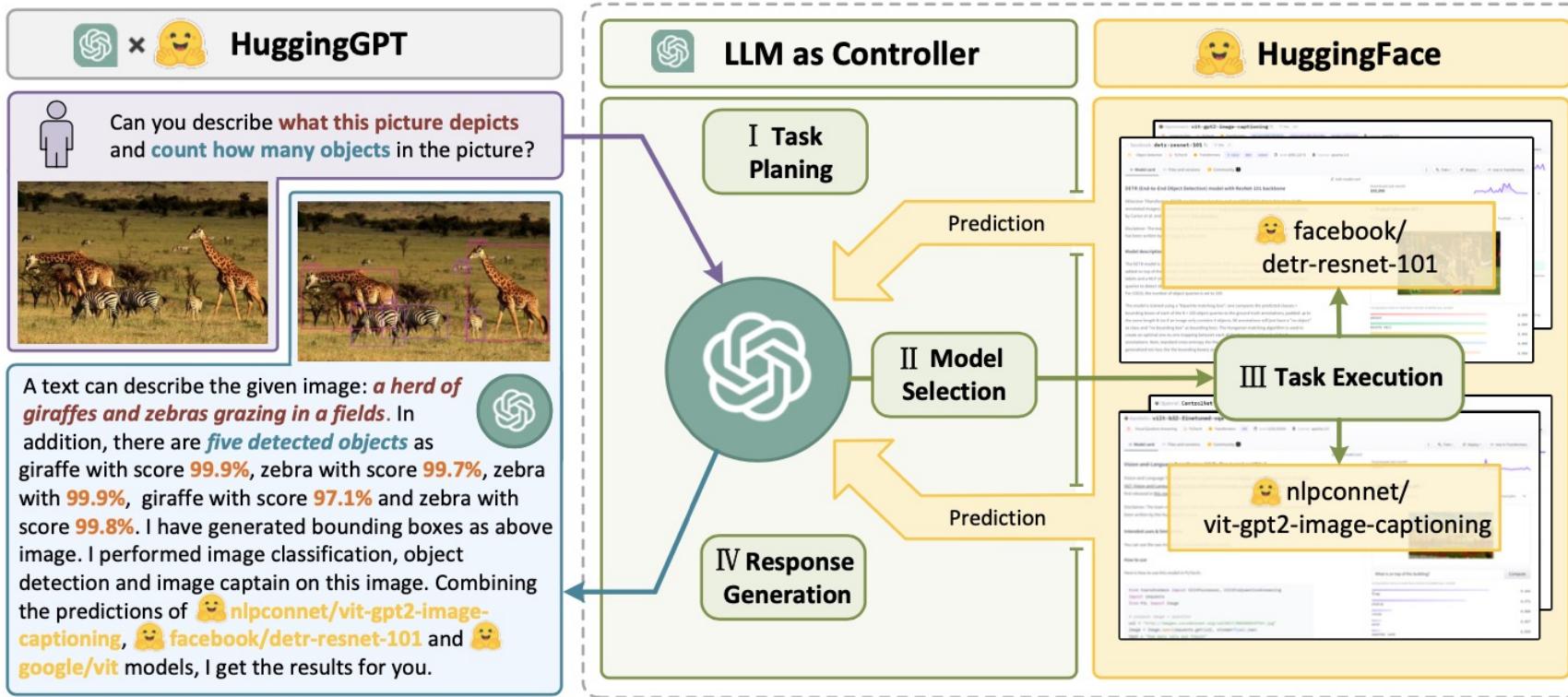
- Toolformer <https://arxiv.org/abs/2302.04761>

微调语言模型学习外部工具API

- WebGPT <https://www.microsoft.com/en-us/bing/apis/bing-web-search-api>

# HuggingGPT

- <https://arxiv.org/abs/2303.17580>



- 任务计划 Task Planning
- 模型选择 Model Selection
- 任务执行 Task Execution
- 回应生成 Response Generation

# AI Society

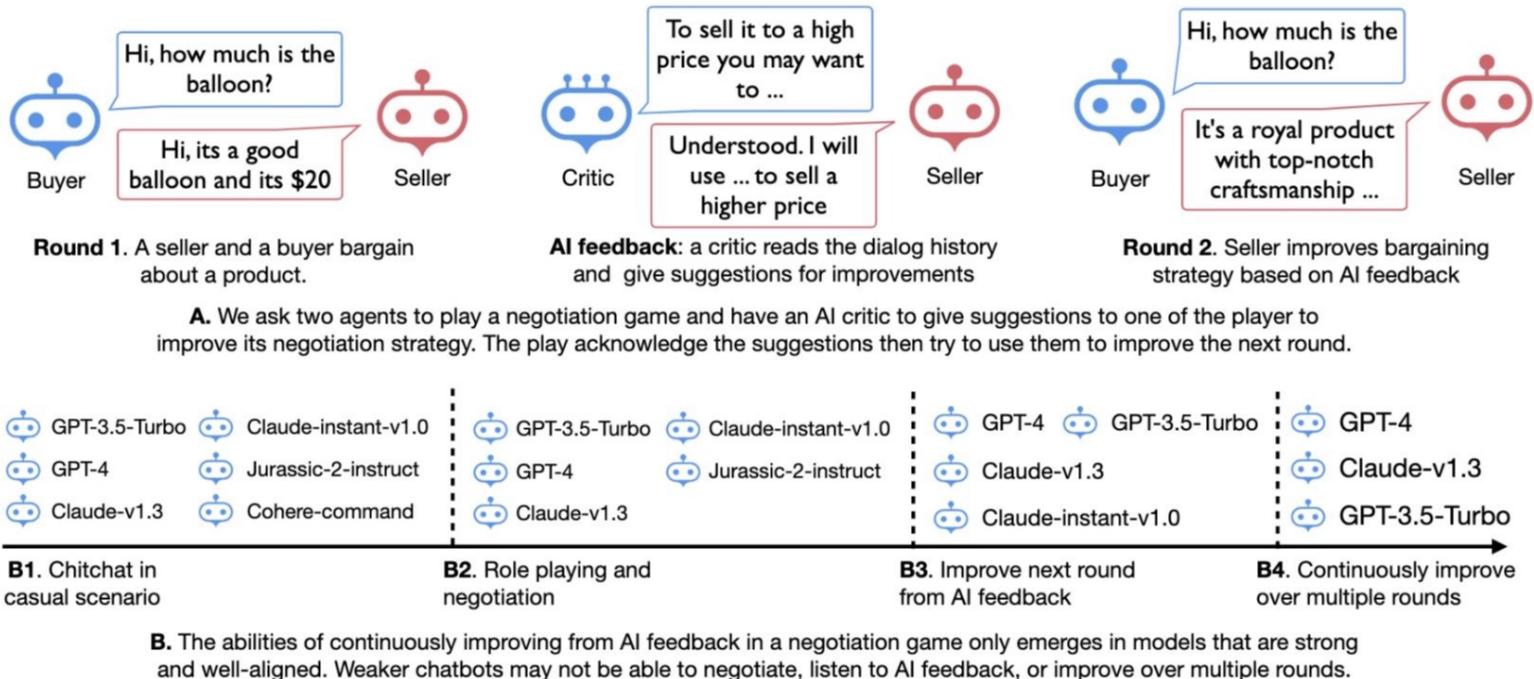




- 25 Agents**
- The Smallville sandbox world
  - Emergent Social Behaviors

# GPT-bargaining

- Negotiation
- Buyer, Seller
- Critic



Improving Language Model Negotiation with  
Self-Play and In-Context Learning from AI Feedback  
<https://arxiv.org/pdf/2305.10142.pdf>

# Multiagent Debate

- Enhances mathematical and strategic reasoning
- Reducing fallacious answers and hallucinations

**User Input:** A treasure hunter found a buried treasure chest filled with gems. There were 175 diamonds, 35 fewer rubies than diamonds, and twice the number of emeralds than the rubies. How many of the gems were there in the chest?

**Round 1**

**Agent 1:** Let  $x$  be the number of rubies. Then the number of diamonds is 175. And the number of emeralds is  $2(x-35)$ , since there are twice as many emeralds as rubies and there are 35 fewer rubies than diamonds.

To find the total number of gems, we add the number of diamonds, rubies, and emeralds: Total gems =  $175 + x + 2(x-35) = 3x + 105$ .

We don't know the exact value of  $x$ , but we do know that it represents the number of rubies, which is less than 175 (since there are 35 fewer rubies than diamonds). Let's try some values of  $x$ : If  $x = 50$ , then the total gems =  $3(50) + 105 = 255$ . If  $x = 100$ , then the total gems =  $3(100) + 105 = 405$ .

Since we know there were fewer rubies than diamonds, the answer must be closer to 255 than to 405. After some trial and error, we can find that  $x = 40$  gives us: Total gems =  $3(40) + 105 = 225$ .

**Agent 2:** If there were 175 diamonds, then the number of rubies is 35 fewer, or  $175 - 35 = 140$  rubies.

And since there were twice as many emeralds as rubies, we can represent the number of emeralds as  $2 * 140 = 280$  emeralds.

To find the total number of gems in the chest, we can add up the number of diamonds, rubies, and emeralds:  $175 + 140 + 280 = 595$ . Therefore, there were 595 gems in the chest.

X✓

**User Input:** Using the solutions from other agents as additional information, can you give an updated response....  
(Copy Agent 2's reply) (Copy Agent 1's reply)

**Round 2**

**Agent 1:** Given the information provided in other agents' solutions, we have two answers: One agent found that the number of rubies is 40 and the total number of gems is 225. Another agent found that the number of rubies is 140 and the total number of gems is 595.

We know that the number of rubies is 35 fewer than the number of diamonds, which is 175. So the number of rubies should be  $175 - 35 = 140$ , as the second agent found. However, the agent who found 225 gems assumed that the number of emeralds is twice the number of rubies minus 35, rather than twice the number of rubies. If we use this assumption, we get: Number of rubies = 140. Number of emeralds =  $2 * 140 - 35 = 245$ . Total number of gems =  $175 + 140 + 245 = 560$ . Our final answer is 560.

**Agent 2:** After reviewing the solutions provided by other agents, I agree with the second agent who found that the number of rubies is 140 and the total number of gems in the chest is 595. Therefore, my answer to the math problem is 595.

X✓

**User Input:** Using the solutions from other agents as additional information, can you give an updated response....  
(Copy Agent 2's reply) (Copy Agent 1's reply)

**Round 3**

**Agent 1:** After reviewing the solutions provided by other agents, I agree with the second agent who found that the number of rubies is 140 and the total number of gems in the chest is 595. Therefore, my answer to the math problem is 595.

**Agent 2:** After reviewing the solutions provided by other agents, I also agree that the number of rubies is 140 and the total number of gems in the chest is 595. Therefore, my answer to the math problem is also 595.

✓✓

Figure 2: **Illustration of Debate.** Illustration of the debate procedure.

# Agent项目

- Auto-GPT <https://github.com/Significant-Gravitas/Auto-GPT>
  - LangChain <https://github.com/langchain-ai/langchain>
  - AutoGen <https://github.com/microsoft/autogen>
  - GPT Engineer <https://github.com/AntonOsika/gpt-engineer>
  - BabyAGI <https://github.com/yoheinakajima/babyagi>
- 
- AI Town <https://github.com/a16z-infra/ai-town>
  - GPTeam <https://github.com/101dotxyz/GPTeam>
  - ChatArena <https://github.com/chatarena/chatarena>
  - AgentVerse <https://github.com/OpenBMB/AgentVerse>

# 数据集基线

| Motivation       | Domain               | Datasets and Benchmarks   | Used by   | Data Link  |
|------------------|----------------------|---|---|--|
| Problem Solving  | Software Development | HumanEval<br>MBPP<br>SoftwareDev  | [Hong <i>et al.</i> , 2023]<br>[Hong <i>et al.</i> , 2023]<br>[Hong <i>et al.</i> , 2023]   | <a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a>   |
|                  | Embodied AI          | RoCoBench<br>Communicative Watch-And-Help (C-WAH)<br>ThreeDWorld Multi-Agent Transport (TDW-MAT)<br>HM3D v0.2 | [Mandi <i>et al.</i> , 2023]<br>[Zhang <i>et al.</i> , 2023c]<br>[Zhang <i>et al.</i> , 2023c]<br>[Yu <i>et al.</i> , 2023]   | <a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a>   |
|                  | Science Debate       | MMLU<br>MedQA<br>PubMedQA<br>GSM8K<br>StrategyQA<br>Chess Move Validity                                       | [Tang <i>et al.</i> , 2023]<br>[Tang <i>et al.</i> , 2023]<br>[Tang <i>et al.</i> , 2023]<br>[Du <i>et al.</i> , 2023]<br>[Xiong <i>et al.</i> , 2023]<br>[Du <i>et al.</i> , 2023]       | <a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a> |
| World Simulation | Society              | SOTOPIA<br>Gender Discrimination<br>Nuclear Energy  | [Zhou <i>et al.</i> , 2023b]<br>[Gao <i>et al.</i> , 2023a]<br>[Gao <i>et al.</i> , 2023a]  | /  |
|                  | Gaming               | Werewolf<br>Avalon<br>Welfare Diplomacy<br>Layout in the Overcooked-AI environment<br>Chameleon<br>Undercover | [Xu <i>et al.</i> , 2023b]<br>[Light <i>et al.</i> , 2023b]<br>[Mukobi <i>et al.</i> , 2023]<br>[Agashe <i>et al.</i> , 2023]<br>[Xu <i>et al.</i> , 2023a]<br>[Xu <i>et al.</i> , 2023a] | /  |
|                  | Psychology           | Ultimatum Game TE<br>Garden Path TE<br>Wisdom of Crowds TE  | [Aher <i>et al.</i> , 2023]<br>[Aher <i>et al.</i> , 2023]<br>[Aher <i>et al.</i> , 2023]   | <a href="#">Link</a><br><a href="#">Link</a><br><a href="#">Link</a>   |
|                  | Recommender System   | MovieLens-1M<br>Amazon review dataset   | [Zhang <i>et al.</i> , 2023a]<br>[Zhang <i>et al.</i> , 2023e]  | <a href="#">Link</a><br>/  |
|                  | Policy Making        | Board Connectivity Evaluation   | [Hua <i>et al.</i> , 2023]  | <a href="#">Link</a>   |

Table 2: Datasets and Benchmarks commonly used in LLM-MA studies. “ / ” denotes the unavailability of data link.