大模型工具学习与的自主智能体初探

从 鑫

清华大学自然语言处理实验室

2024.09.20

个人简介

清华大学，计算机科学与技术系，自然语言处理实验室

研究方向：大语言模型，自主智能体，工具学习

电子邮箱：xin.cong@outlook.com

人工智能进入大模型时代

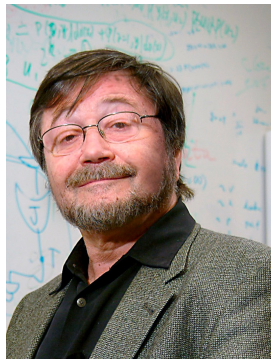
- 人工智能经历了基于预定义规则的**符号智能**时代，基于数据驱动的**专用智能**时代，发展到了以**大模型**为代表的**通用智能**时代

1960 → 1990 → 2010 → 2022

Noam Chomsky
语言学巨擘

20世纪50年代后，**词法、句法、语法规则**被广泛应用于自然语言理解

Judea Pearl
图灵奖获得者

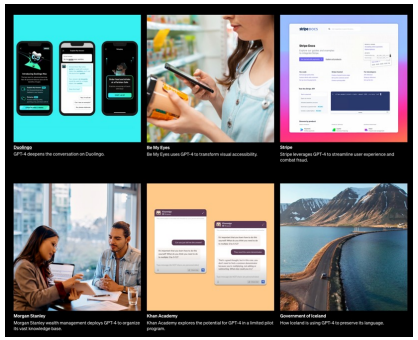


概率与统计方法被引入自然语言处理，模型从**符号系统推演转变为数据驱动学习**

Yann LeCun
图灵奖获得者

2010年后**神经网络兴起**，将**数据驱动**的效果推到新高位，成为当前自然语言处理的主要范式

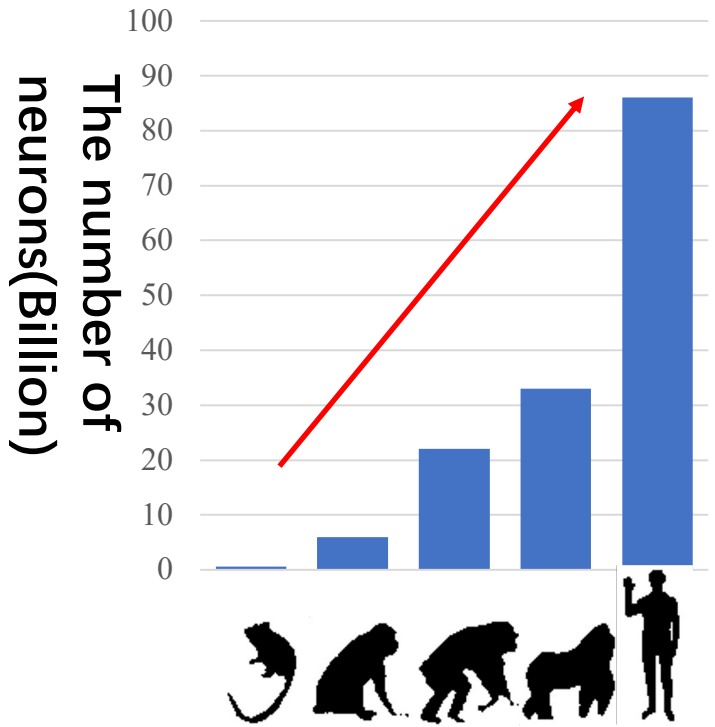
OpenAI
大模型领军者

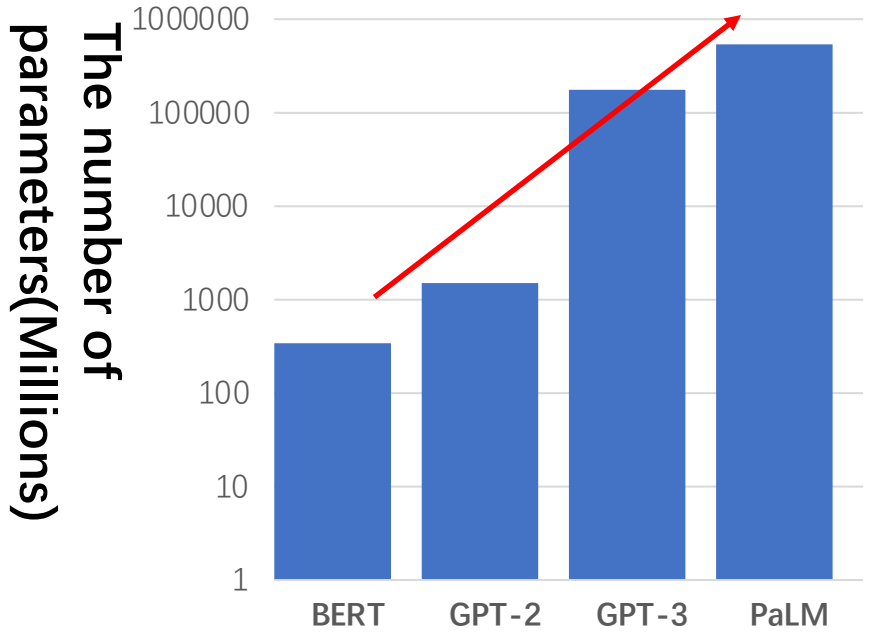


OpenAI发布的ChatGPT**大模型**可通过**文字、视觉信号**与人类交互，可**使用工具**处理复杂任务

智能涌现

- **神经元数量增加**带来**生物个体**的智能涌现，**参数增加**带来单个**大模型**的智能涌现





大模型初步展现出超越人类的智能水平

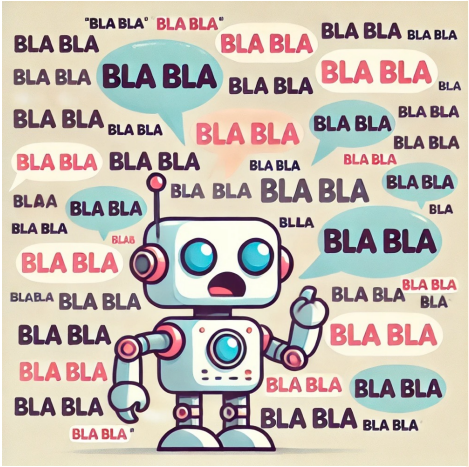
- 最近，OpenAI推出的OpenAI o1甚至在**物理**、**生物**以及**化学**领域上取得了**超越PhD水平的成绩**

大模型的局限性

- 大模型**缺乏与环境互动**的能力，导致难以切实解决实际任务

帮我安排一下国庆7天去重庆的旅游计划

这是我安排计划，你可以照着做



重庆国庆7天旅游计划

第一天（10月1日）：抵达重庆，初识山城

- **上午**：抵达重庆，办理酒店入住手续。
- **下午**：前往【解放碑步行街】，感受重庆市中心的繁华，购物和品尝当地小吃。
- **晚上**：游览【洪崖洞】，欣赏吊脚楼建筑和夜景，品尝正宗的重庆火锅。

第二天（10月2日）：城市风光与历史文化

- **上午**：参观【重庆人民大礼堂】和【三峡博物馆】，了解重庆的历史和文化。
- **下午**：前往【磁器口古镇】，漫步青石板路，体验古镇风情，购买手工艺品。
- **晚上**：乘坐【两江夜游】游船，欣赏嘉陵江和长江的夜景。

第三天（10月3日）：红色之旅

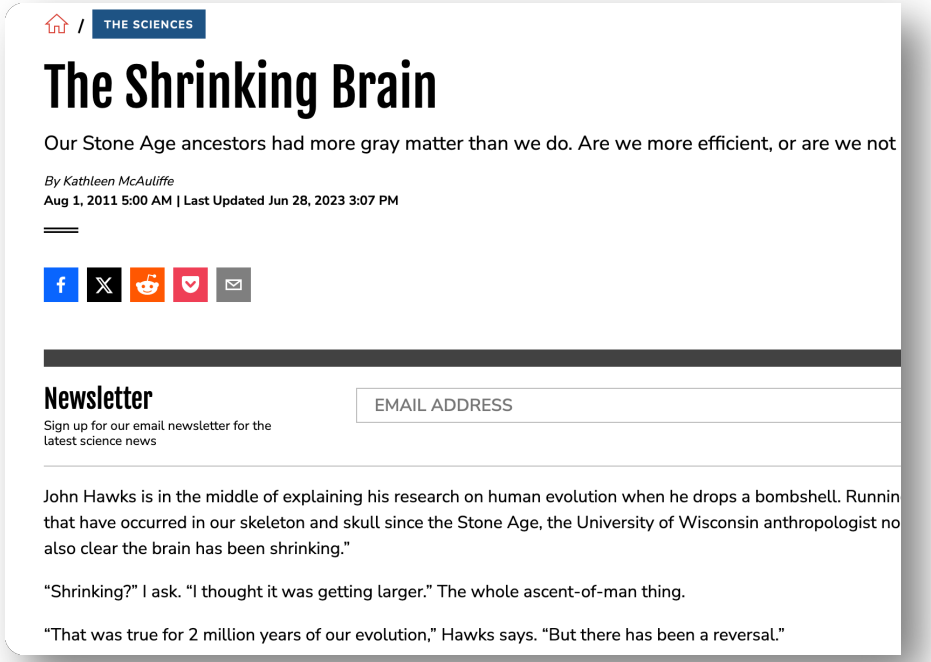
- **上午**：参观【红岩革命纪念馆】和【渣滓洞】、【白公馆】，了解革命历史。
- **下午**：前往【鹅岭二厂文创园】，感受文艺气息，拍照打卡。
- **晚上**：品尝当地特色美食，如酸辣粉、小面等。

第四天（10月4日）：武隆自然风光

- **全天**：参加【武隆一日游】。
  - **天生三桥**：观赏世界上最大的天然石拱桥群。
  - **龙水峡地缝**：探秘地质奇观。

人类智能的进化

- 美国《发现》杂志指出，在二百万年的进化中人类大脑容量经历了飞速增长，但**过去10万年里脑容量不再增长**，人类通过**使用工具**达成能力延展，降低了对脑容量增长的依赖，实现了人类文明的进一步阶跃式发展



/ THE SCIENCES

The Shrinking Brain

Our Stone Age ancestors had more gray matter than we do. Are we more efficient, or are we not

By Kathleen McAuliffe
Aug 1, 2011 5:00 AM | Last Updated Jun 28, 2023 3:07 PM

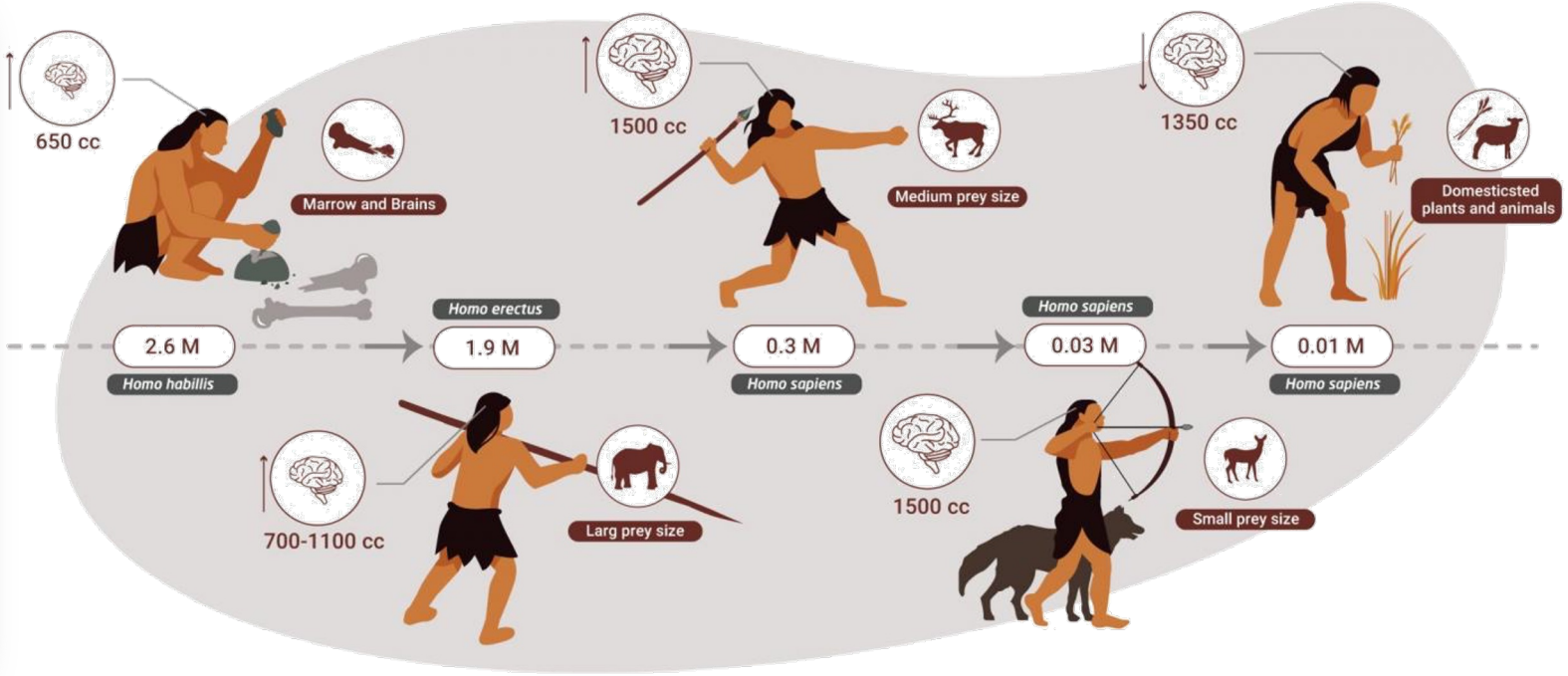
Newsletter
Sign up for our email newsletter for the latest science news

EMAIL ADDRESS

John Hawks is in the middle of explaining his research on human evolution when he drops a bombshell. Runnin that have occurred in our skeleton and skull since the Stone Age, the University of Wisconsin anthropologist no also clear the brain has been shrinking."

"Shrinking?" I ask. "I thought it was getting larger." The whole ascent-of-man thing.

"That was true for 2 million years of our evolution," Hawks says. "But there has been a reversal."



基于工具智能迈向自主智能体

- 受人类智能发展历程的启发，迈向通用人工智能的潜在途径是**结合外部工具**形成**大模型驱动的自主智能体**，提升智能水平与应用潜力

发展阶段			
人类智能	**脑容量小**	**脑容量大**	**使用工具**
人工智能	**小模型**	**大模型**	**智能体**

什么是自主智能体

- 自主智能体是人工智能体领域中一个核心研究领域，它旨在研究如何构建一个**具有智能的实体**使其可以**自主解决各种任务**

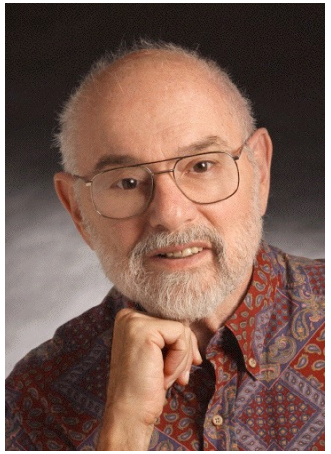
Stuart Russell
UCB 教授

智能体为可以**感知环境**并作出响应以**实现某些目标**的任何实体

Michael Wooldridge
牛津大学 教授

智能体是指具有**自主性**、**社会性**、**反应性**和**能动性**等基本特性的计算机硬件或软件

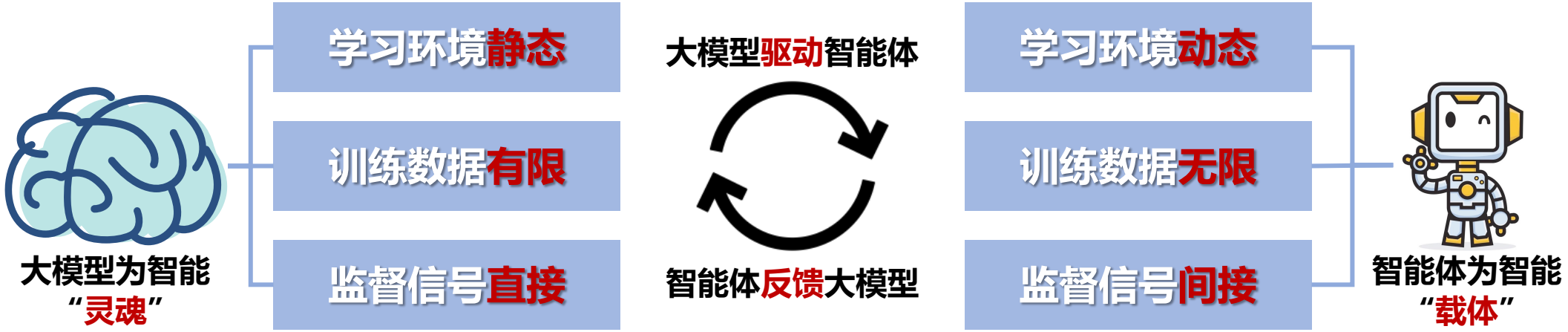
Stan Franklin
孟菲斯大学 教授



智能体是位于环境中的系统，能够**观察环境**并根据其变化，持续不断地**行动**以**实现自己的目标**

什么是自主智能体

- 基础大模型作为智能应用技术的内核，必须以**自主智能体作为载体**与**动态环境进行交互**，才能充分胜任动态复杂的智能应用，因此亟需发展大模型驱动的自主智能体技术



智能体发展历程

- 与人工智能发展类似，智能体技术也经历了基于**符号规则**的智能体时期，基于**强化学习**的智能体时期，发展到了以**大模型**为内核的智能体时期

基于符号规则的智能体 → 基于强化学习的智能体 → 基于大模型的智能体 →

Deep Blue
国际象棋智能体

智能体专家通过编写**规则**或者**算法**来实现智能体的自主控制与决策，有时还需要配合专用硬件

AlphaGO
围棋智能体

强化学习方法入智能体，智能体通过不断尝试根据**环境反馈**信息，自主学习如何决策

AutoGPT
大模型智能体

将**大模型**作为智能体的大脑或内核，实现了通用的自主规划，并通过**使用工具**实现与环境的交互

基于符号规则的早期范式

- 基于**符号规则**的智能体时期：早期的智能体系统主要基于**规则逻辑**，通常被设计为处理特定任务的**专家系统**，十分**依赖专家知识与经验**

1966年，麻省理工学院的Weizenbaum教授开发了心理治疗师智能体ELIZA，通过人工构建**关键词映射表**来对话

1997年，IBM 基于**专家系统**开发的国际象棋系统“深蓝”，击败了国际象棋世界冠军 Garry Kasparov

基于强化学习的经典范式

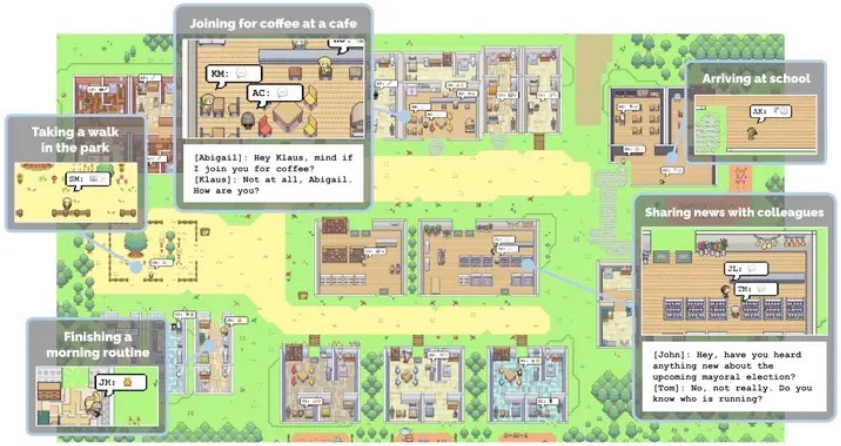
- 基于**强化学习**的智能体时期：和强化学习的结合使得智能体能够**自主学习**达到**超越人类的能力水平**，但是受限于有限场景，**缺乏通用性**

DeepMind 开发的 AlphaGo 结合了**蒙特卡洛树搜索**与**深度学习**，先后在2016年击败李世石、2017年击败柯洁两位围棋世界冠军

DeepMind 于2020发布了Agent57，其可以在雅达利平台上**57个游戏**中展现出**超越人类水平**的游戏能力

基于大模型的新兴范式

- 基于**大模型**的智能体时期：利用大模型所具备的具备深度理解和生成能力，来执行复杂任务，可以在**无需大量特定领域数据**的情况下，**快速适应并胜任新的任务**



2023年，斯坦福大学发布了斯坦福小镇，基于**大模型**构建了25个智能体来**模拟人类行为**，从而进行社会行为研究

2023年11月6日， OpenAI在第一届开发者大会上正式提出构建**GPT Store**及相关生态的战略目标

大模型智能体的基本框架

- **感知**：负责获取外部世界的多模态信息输入
- **规划**：根据任务目标，对任务进行合理拆解与规划
- **执行**：将做出的任务拆解与规划出的策略付诸实施
- **记忆**：对历史信息进行回忆，并对其高效利用

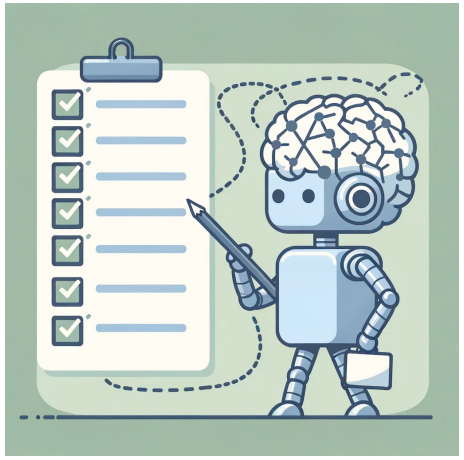
感知

理解环境信息

执行

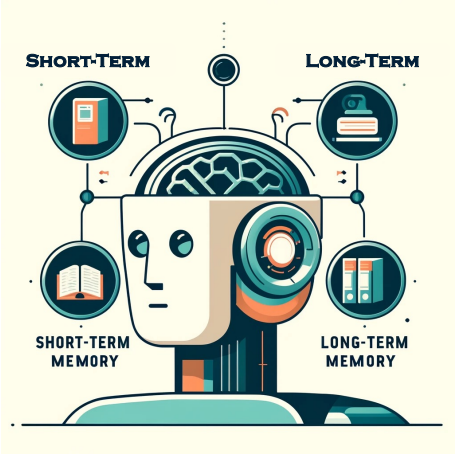
策略付诸实践

大模型智能体



目标驱动的任务规划

规划



高效回忆历史

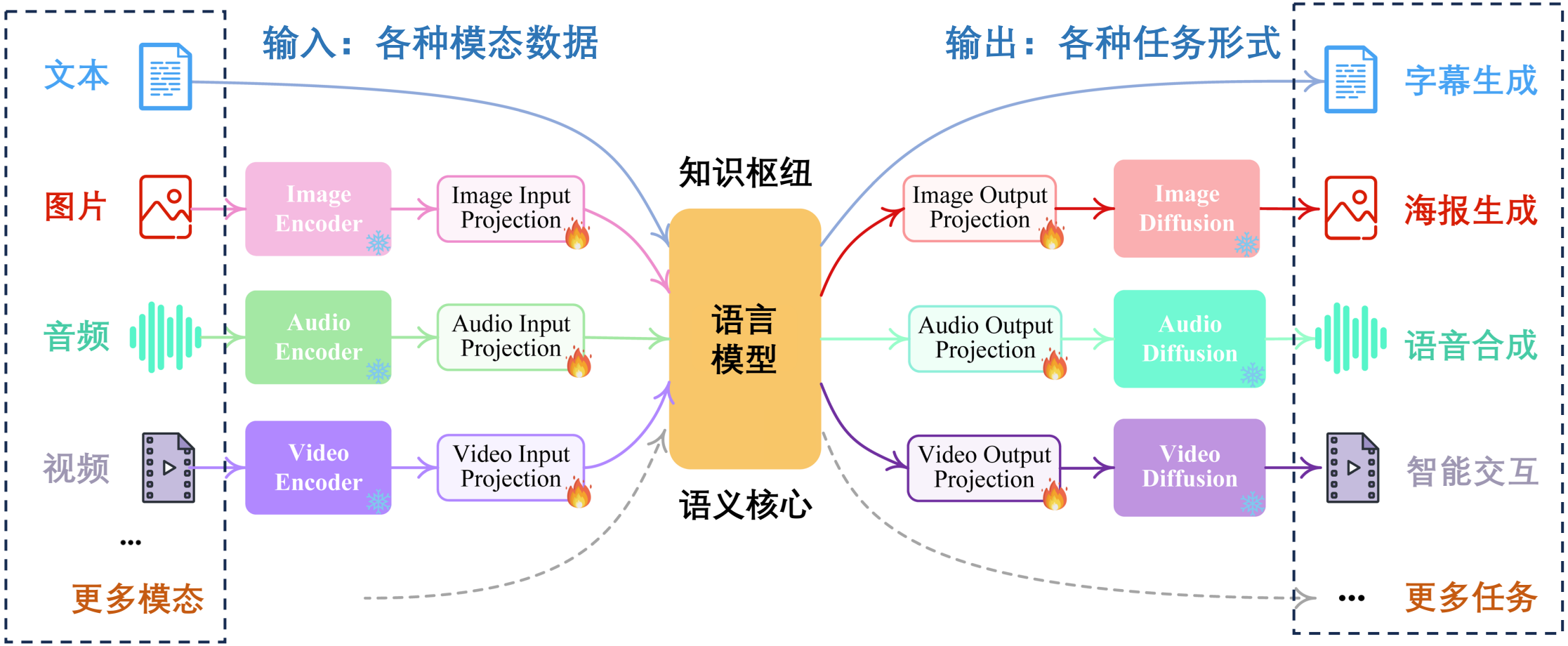
记忆

感知：环境信息理解

- 视觉 – 理解图像和文字信息
- 听觉 – 接收声音和语言输入
- 嗅觉 – 探索环境的气味
- 味觉 – 品尝食物不同的味道
- 触觉 – 通过接触和感觉来接受环境反馈

感知：智能体的感知能力通过序列化数据实现

- 以语言模型为枢纽核心，实现涉及多种模态复杂任务的统一建模

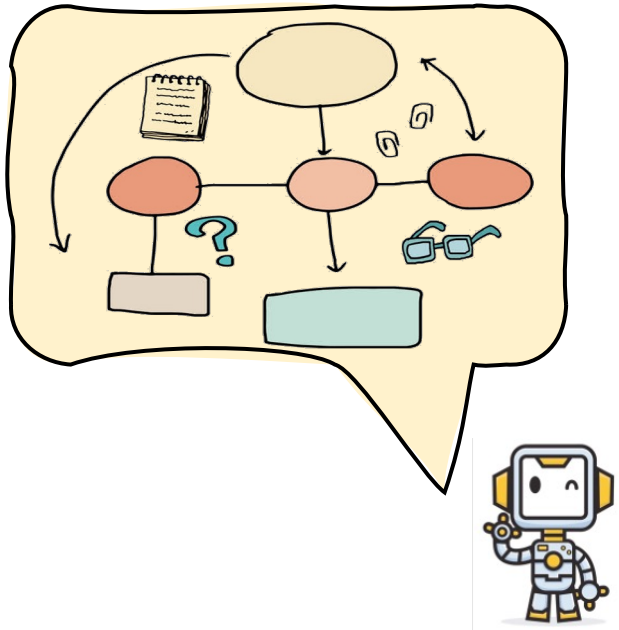


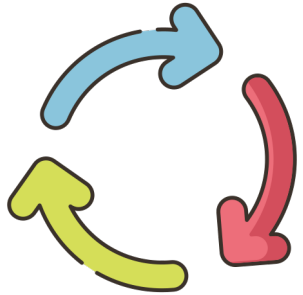
Wu S, et al. NExT-GPT: Any-to-any multimodal LLM. Preprint.

规划：目标驱动的任务规划

- Planning is the process of thinking regarding the activities required to achieve a desired goal. Planning is based on foresight, the fundamental capacity for mental time travel

—— WIKIPEDIA The Free Encyclopedia

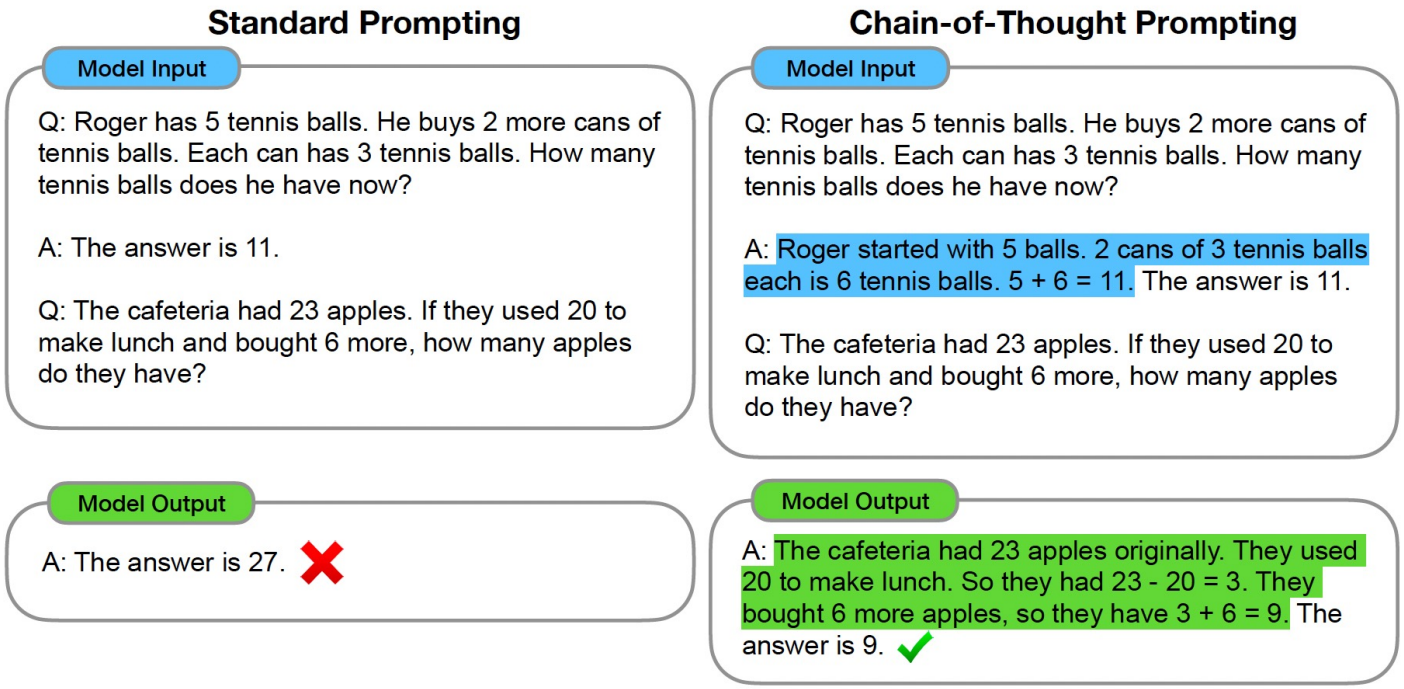




外部环境

规划：思维链

- 在完成实际任务的场景，模型常常需要通过多步的形式完成任务，比如查搜索引擎、写代码等等。从完成整个任务的视角来看，智能体在使用思维链的方式解决问题



Wei, et al. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. NeurIPS 2022.

规划：带环境反馈的规划

- 在规划时，需要一定的环境反馈信息动态调整规划

(1) Hotspot QA

Question: Aside from the Apple Remote, what other device can control the program Apple Remote was originally designed to interact with?

(1a) Standard

Answer: iPod ✗

(1b) CoT (Reason Only)

Thought: Let's think step by step. Apple Remote was originally designed to interact with Apple TV. Apple TV can be controlled by iPhone, iPad, and iPod Touch. So the answer is iPhone, iPad, and iPod Touch.

Answer: iPhone, iPad, iPod Touch ✗

(1c) Act-Only

Act 1: Search[Apple Remote]
Obs 1: The Apple Remote is a remote control …

Act 2: Search[Front Row]
Obs 2: Could not find [Front Row]. Similar: …

Act 3: Search[Front Row (software)]
Obs 3: Front Row is a discontinued media center software …

Act 4: Finish[yes] ✗

(1d) ReAct (Reason + Act)

Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with.
Act 1: Search[Apple Remote]
Obs 1: The Apple Remote is a remote control introduced in October 2005 by Apple … originally designed to control the Front Row media center program …

Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search Front Row next and find what other device can control it.
Act 2: Search[Front Row]
Obs 2: Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports',' Front Row (software)', …]

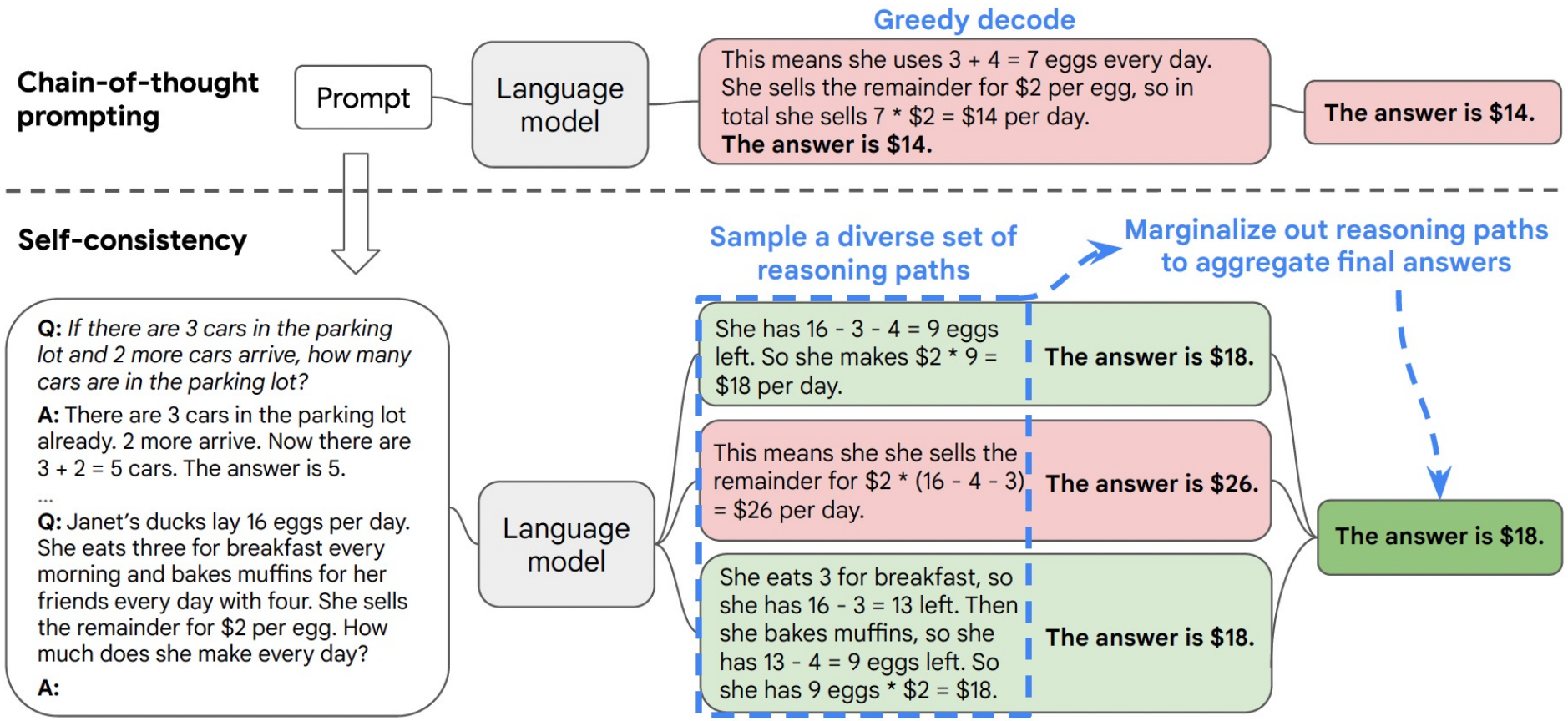
Thought 3: Front Row is not found. I need to search Front Row (software) .
Act 3: Search[Front Row (software)]
Obs 3: Front Row is a discontinued media center software …

Thought 4: Front Row (software) is controlled by an Apple Remote or the keyboard function keys. So the answer is keyboard function keys.
Act 4: Finish[keyboard function keys] ✓

Yao, et al. ReAct: Synergizing Reasoning and Acting in Language Models. ICLR 2023.

规划：自我一致性的思维链

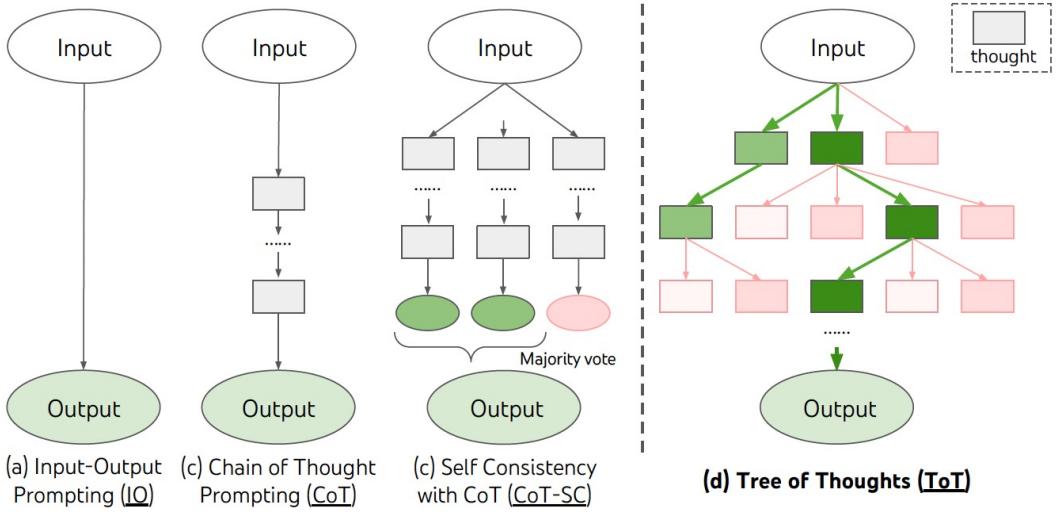
- 智能体的一次思考过程仍可能犯错，多思考几次反而更有可能找到正确的答案



Wang, et al. Self-Consistency Improves Chain of Thought Reasoning in Language Models. ICLR 2023.

规划：思维树

- 单纯链式的思维方式会制约模型的规划能力，在复杂任务中会需要一定的探索



IO: $y \sim p_\theta^{IO}(y|x)$

CoT: $[z_{1\cdots n}, y] \sim p_\theta^{CoT}(z_{1\cdots n}, y|x)$

CoT-SC: $[z_{1\cdots n}^{(i)}, y^{(i)}] \sim p_\theta^{CoT}(z_{1\cdots n}, y|x) \ (i = 1 \cdots k)$

Thought decomposition

- a couple of words
- a line of equation
- a writing plan

Thought generation

- Sample Method
- Propose Method

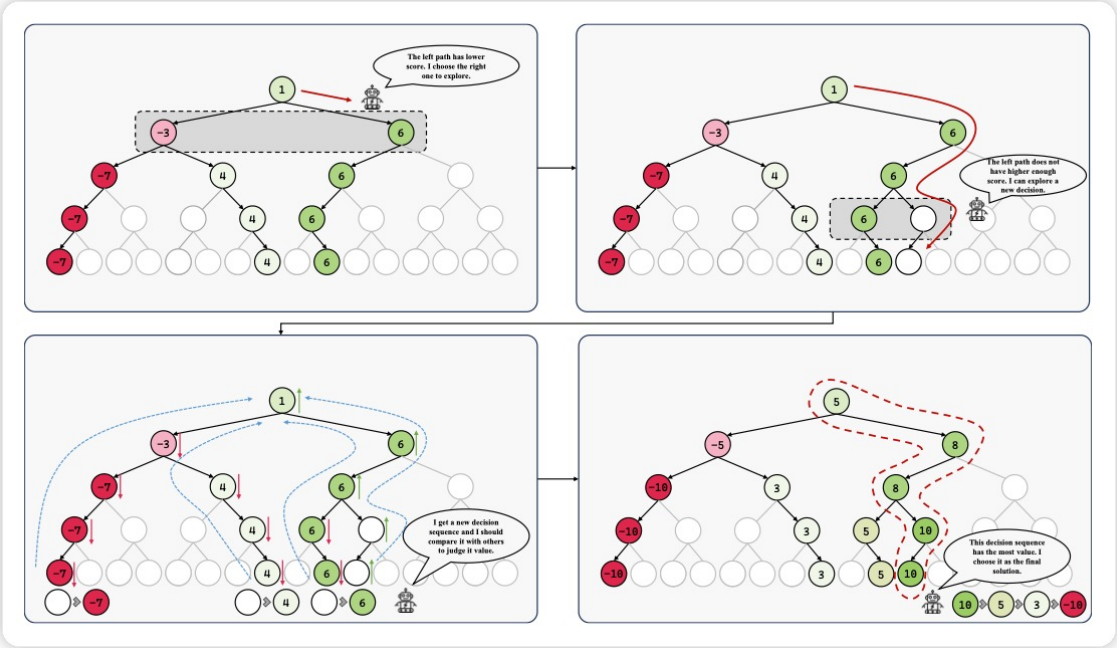
Thought evaluation

- Value each thought
- Vote multiple thoughts

Yao, et al. Tree of Thoughts: Deliberate Problem Solving with Large Language Models. NeurIPS 2023.

规划：自主探索能力

- 现有大模型搜索算法依赖于人类的先验，缺乏在面对未知任务时的自主性



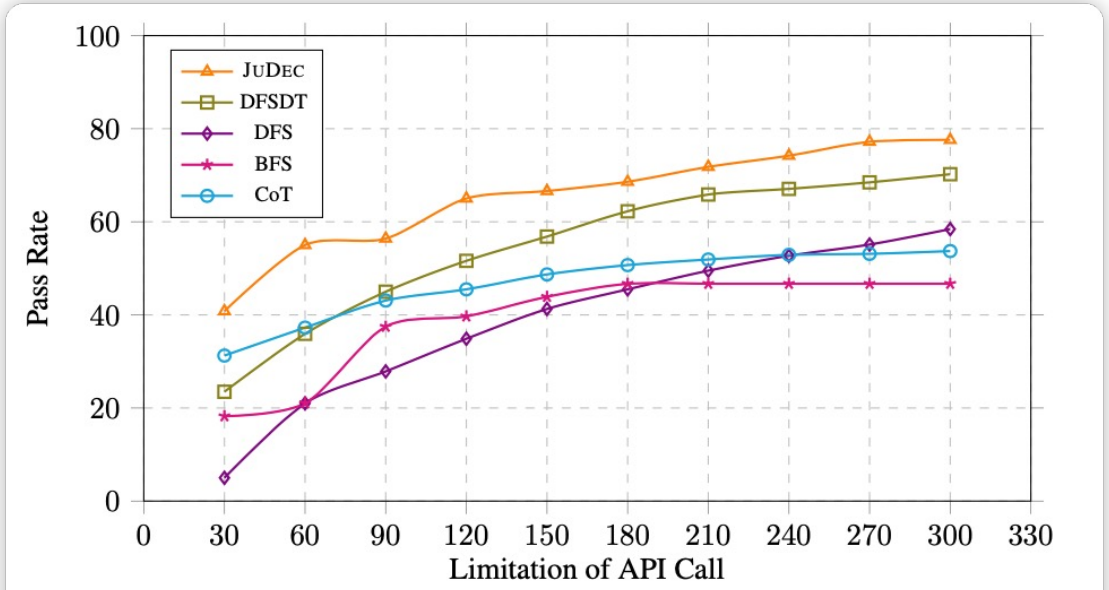
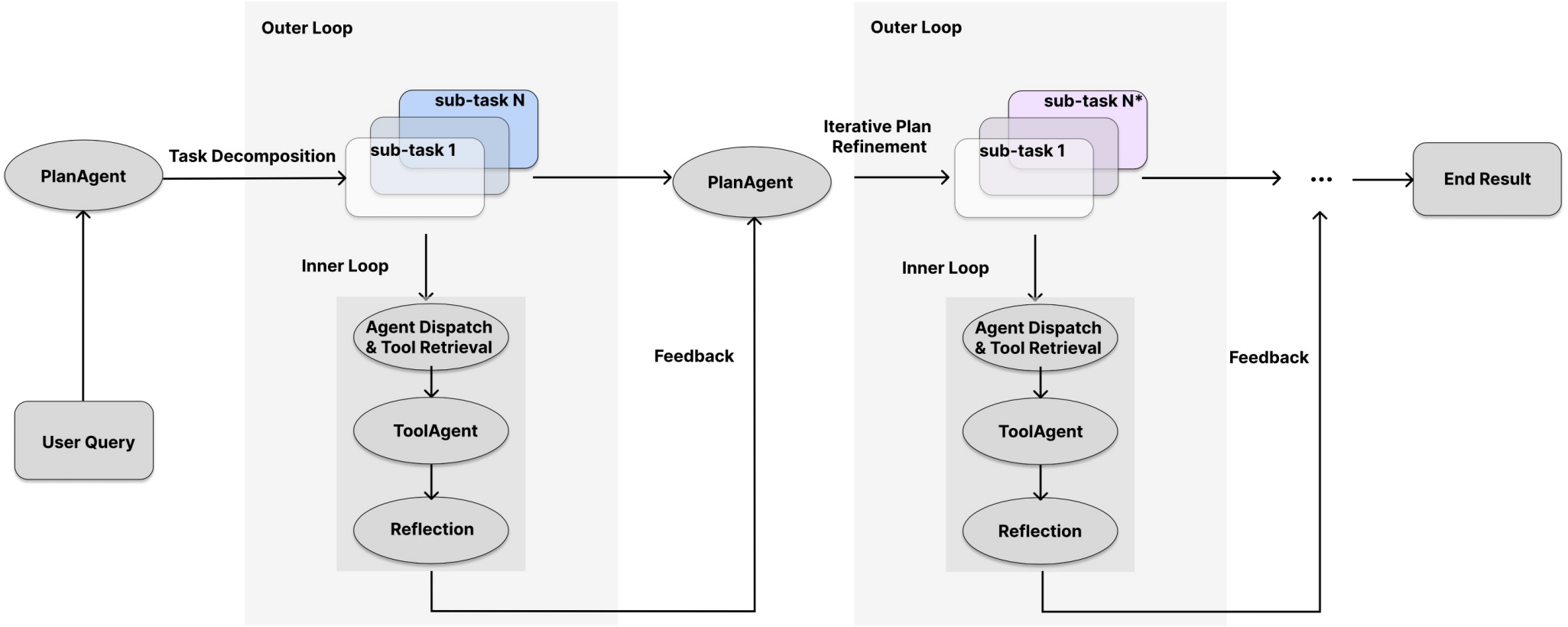


Figure 3: Efficiency experimental results on various API cal limitations.

Ye, et al. Rational Decision-Making Agent with Internalized Utility Judgment. Preprint.

规划：层次化任务拆分

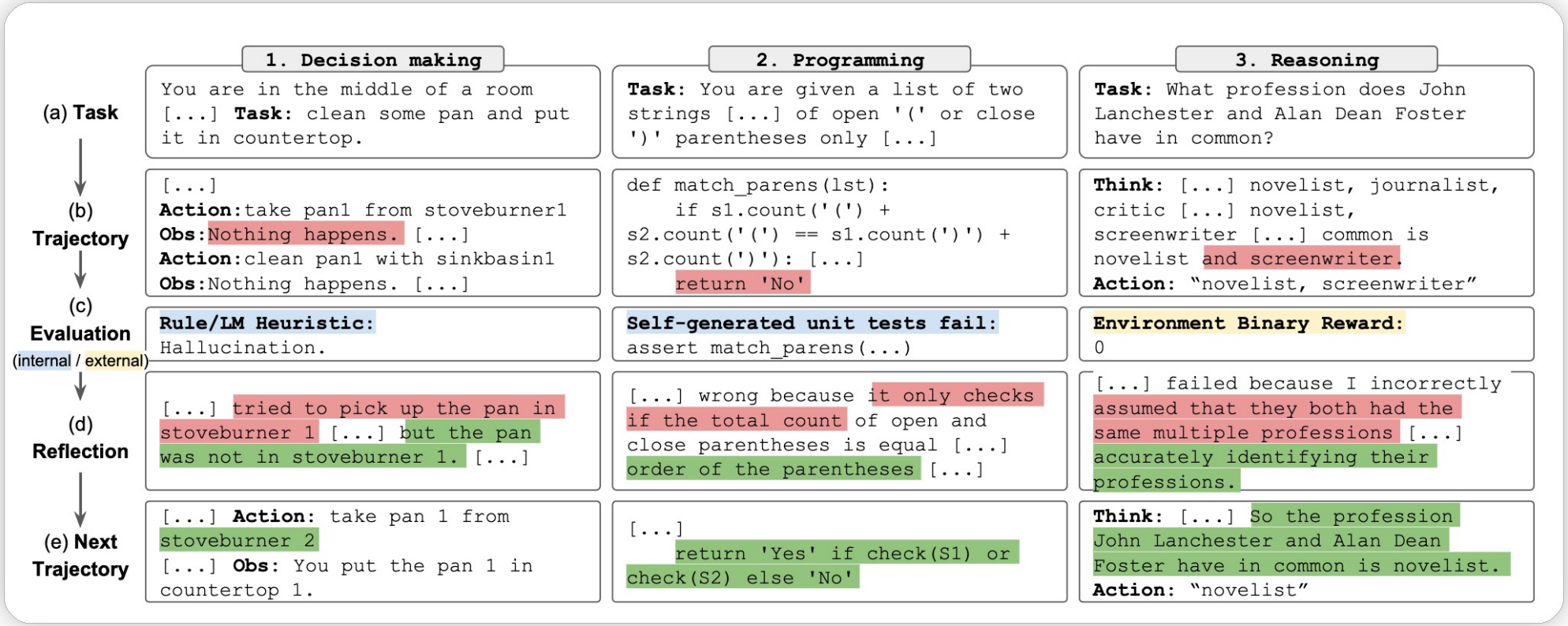
- 复杂的任务通常难以直接解决，智能体需要把复杂的任务**拆解**成若干个**简单子任务**，再通过完成简单的任务去解决原来的复杂任务



Qin, et al. XAgent: An Autonomous Agent for Complex Task Solving. https://blog.x-agent.net/blog/xagent/

规划：事后反思

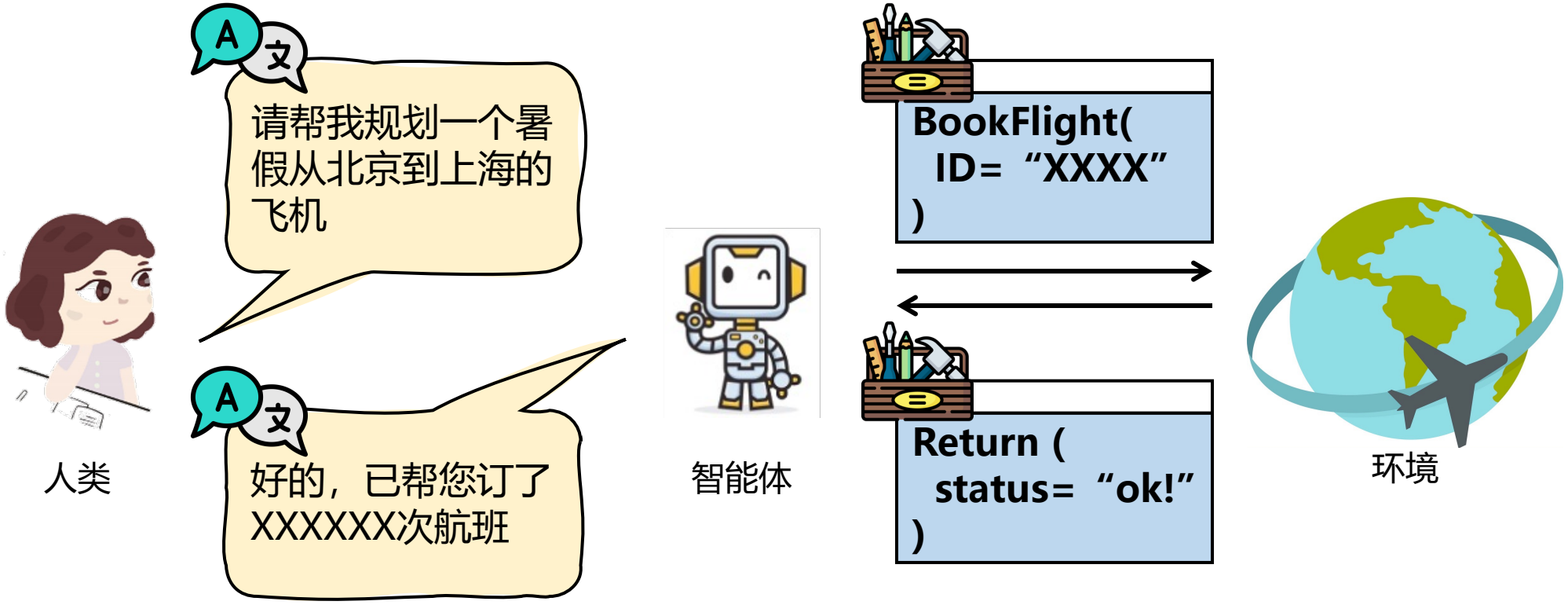
- 现实情况往往很复杂，执行任务有概率会失败，智能体需要有**自我反思**能力，从而可以在失败中汲取经验，让下次执行的结果更好



Shinn, et al. Reflexion: Language agents with verbal reinforcement learning. NeurIPS 2023.

执行：智能体将任务规划付诸实践

- 智能体在做出任务规划之后，需要按规划进行任务执行，而这个过程主要涉及**通过工具与环境交互**



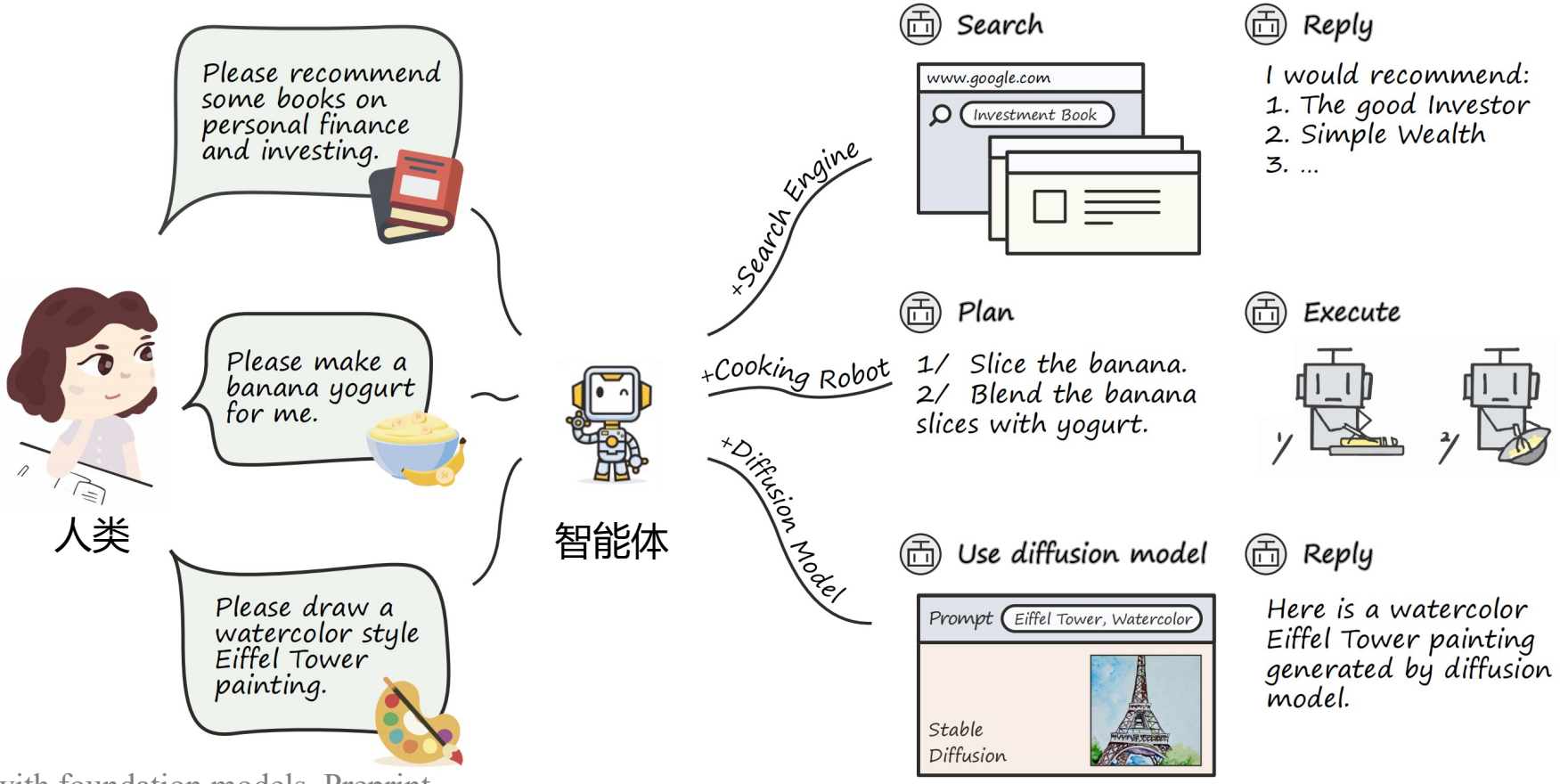
基于语言与人类交互　　基于工具与环境交互

执行：智能体通过使用工具实现环境交互

- 在人类文明的发展历程中，工具起着至关重要的作用，人类**通过使用工具改变生存环境**，提高了生产效率，推动了科技进步，实现种群繁荣
- 大模型已经掌握了使用语言的能力，**是否能够具备像人类一样的使用工具的能力**?

执行：智能体像人类一样使用工具

- 大模型智能体可以像人类一样，掌握如搜索引擎等**高级认知工具**，从而完成复杂现实任务



Qin, et al. Tool learning with foundation models. Preprint.

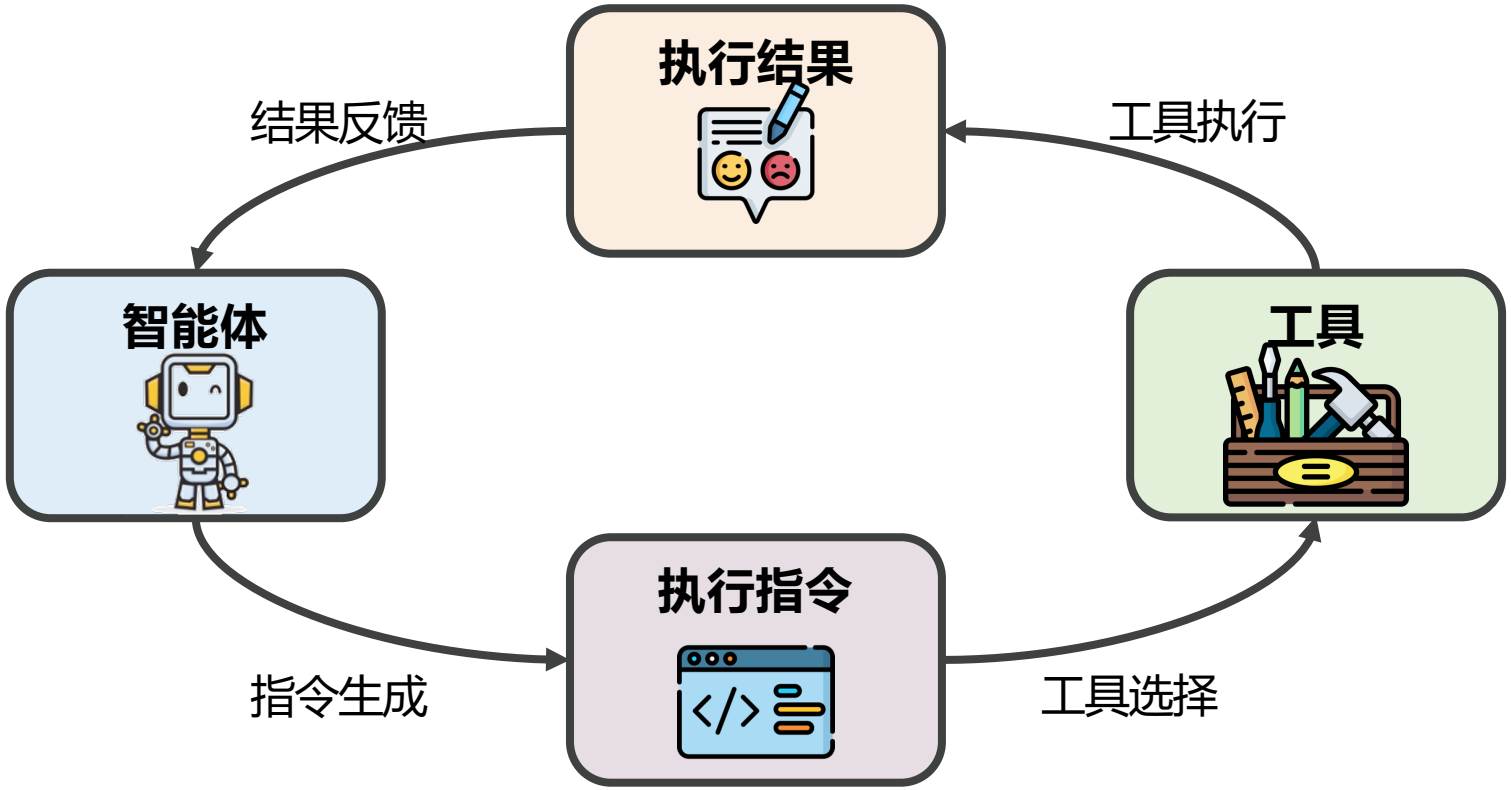
执行：智能体使用工具的类型

- 从使用方式的角度看，工具可以分为：**基于物理交互的工具**、**基于图形界面交互的工具**和**基于程序交互的工具**

Qin, et al. Tool learning with foundation models. Preprint.

执行：通过生成特定格式文本来使用工具

- 智能体生成具有**特定格式的执行指令**来指定使用什么工具以及如何使用在工具使用后，相关的结果以及反馈信息会返回给智能体



The New England Journal of Medicine is a registered trademark of [QA("Who is the publisher of The New England Journal of Medicine?") → Massachusetts Medical Society] the MMS.

Out of 1400 participants, 400 (or [Calculator(400 / 1400) → 0.29] 29%) passed the test.

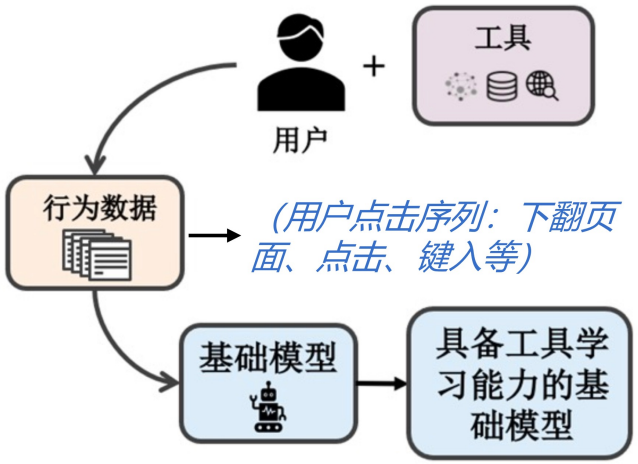
The name derives from "la tortuga", the Spanish word for [MT("tortuga") → turtle] turtle.

The Brown Act is California's law [WikiSearch("Brown Act") → The Ralph M. Brown Act is an act of the California State Legislature that guarantees the public's right to attend and participate in meetings of local legislative bodies.] that requires legislative bodies, like city councils, to hold their meetings open to the public.

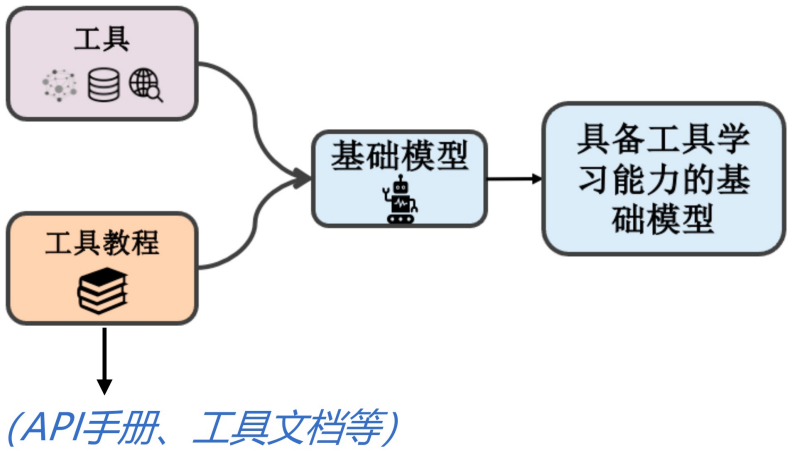
Timo Schick, et al. Toolformer: Language Models Can Teach Themselves to Use Tools. NeurIPS 2023.

执行：学习使用工具的途径

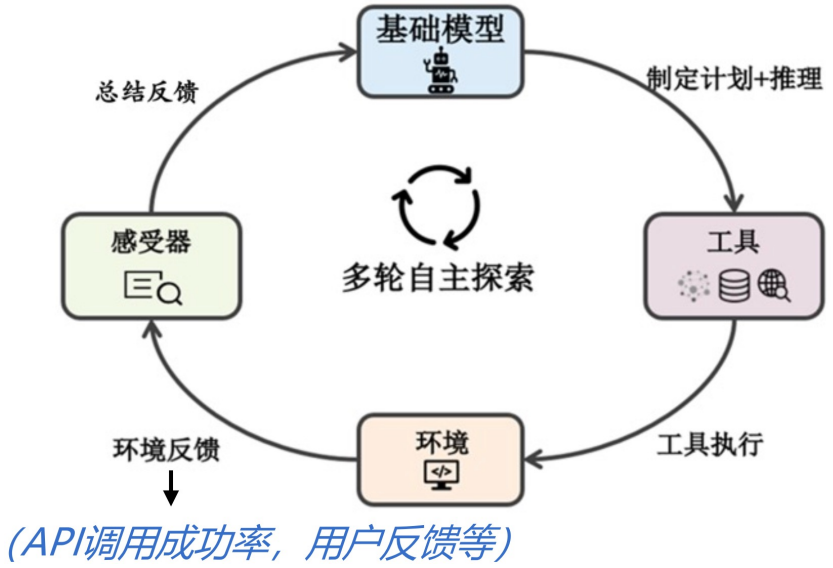
- 智能体学会使用工具主要由三种途径：**模仿学习**、**教程学习**以及**强化学习**



模仿学习



教程学习

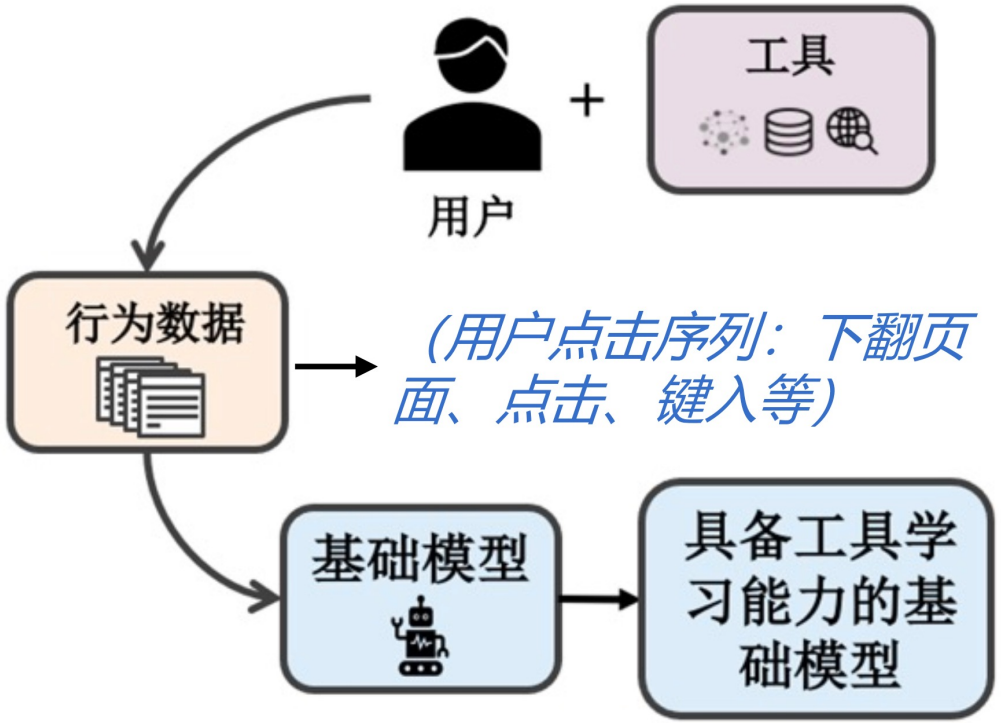


强化学习

Qin, et al. Tool learning with foundation models. Preprint.

执行：模仿学习

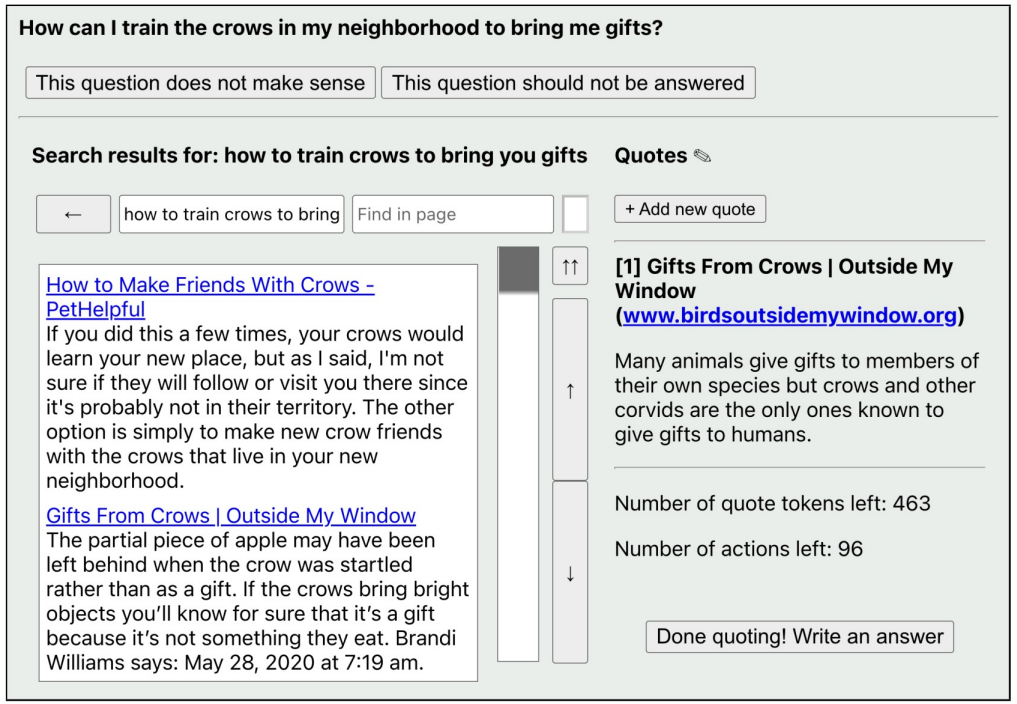
- **模仿学习**：通过**记录人类使用工具的行为数据**，大型模型模仿人类动作来学习如何使用工具，这是工具学习中最简单和最直接的方法



Qin, et al. Tool learning with foundation models. Preprint.

执行：WebGPT

- OpenAI收集了**6000条**人类的浏览器行为数据，让模型**行为克隆**学习人类使用浏览器

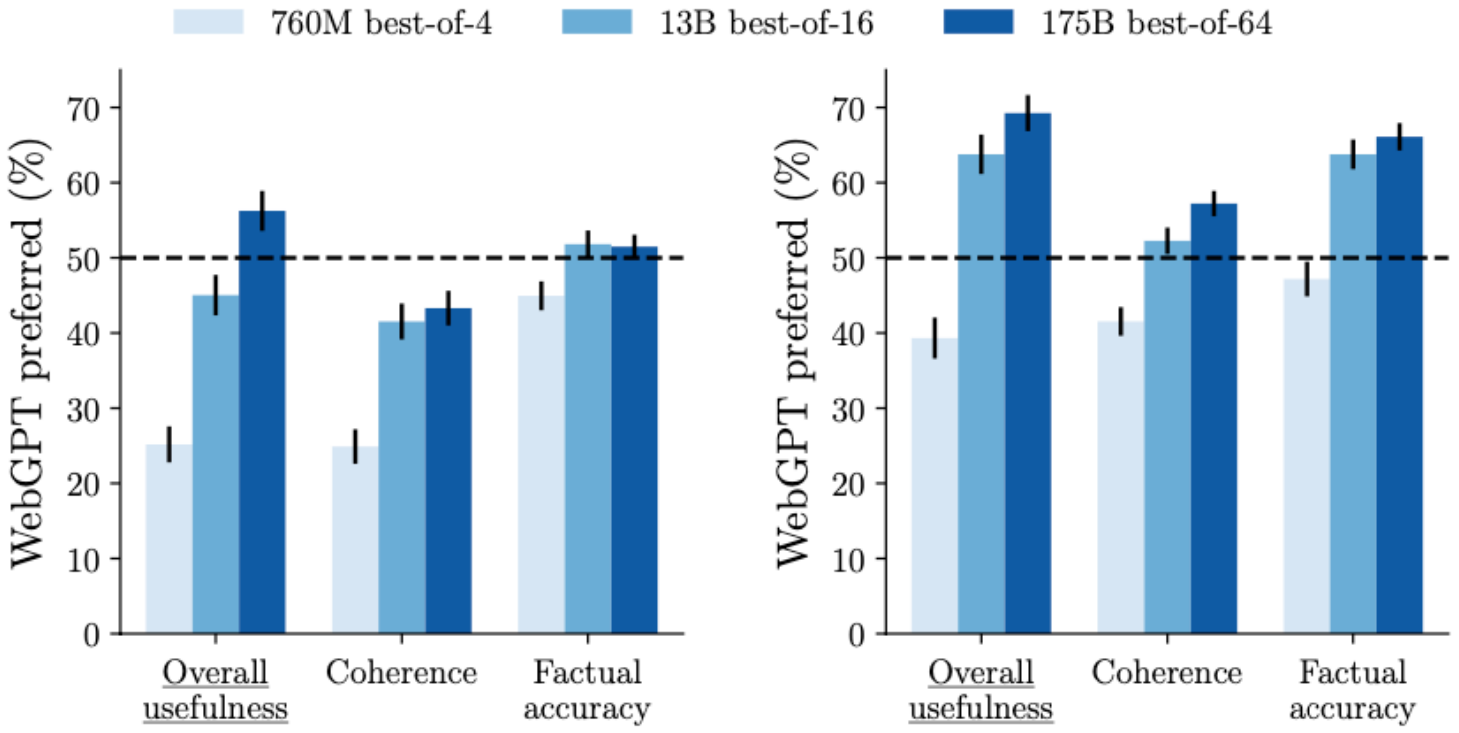



```
♦Question
How can I train the crows in my neighborhood to bring me gifts?

♦Quotes
From Gifts From Crows | Outside My Window (www.birdsoutsidemywindow.org)
> Many animals give gifts to members of their own species but crows and
other corvids are the only ones known to give gifts to humans.

♦Past actions
Search how to train crows to bring you gifts
Click Gifts From Crows | Outside My Window www.birdsoutsidemywindow.org
Quote
Back

♦Title
Search results for: how to train crows to bring you gifts

♦Scrollbar: 0 - 11
♦Text
【0†How to Make Friends With Crows - PetHelpful†pethelpful.com】
If you did this a few times, your crows would learn your new place, but
as I said, I'm not sure if they will follow or visit you there since it's
probably not in their territory. The other option is simply to make new
crow friends with the crows that live in your new neighborhood.

【1†Gifts From Crows | Outside My Window†www.birdsoutsidemywindow.org】
The partial piece of apple may have been left behind when the crow was
startled rather than as a gift. If the crows bring bright objects you'll
know for sure that it's a gift because it's not something they eat.
Brandi Williams says: May 28, 2020 at 7:19 am.

♦Actions left: 96
♦Next action
```

Nakano, Reiichiro, et al. WebGPT: Browser-assisted question-answering with human feedback. Preprint.

执行：WebGPT

- 最终模型在Long-form QA任务上效果**超越了人类专家的水平**



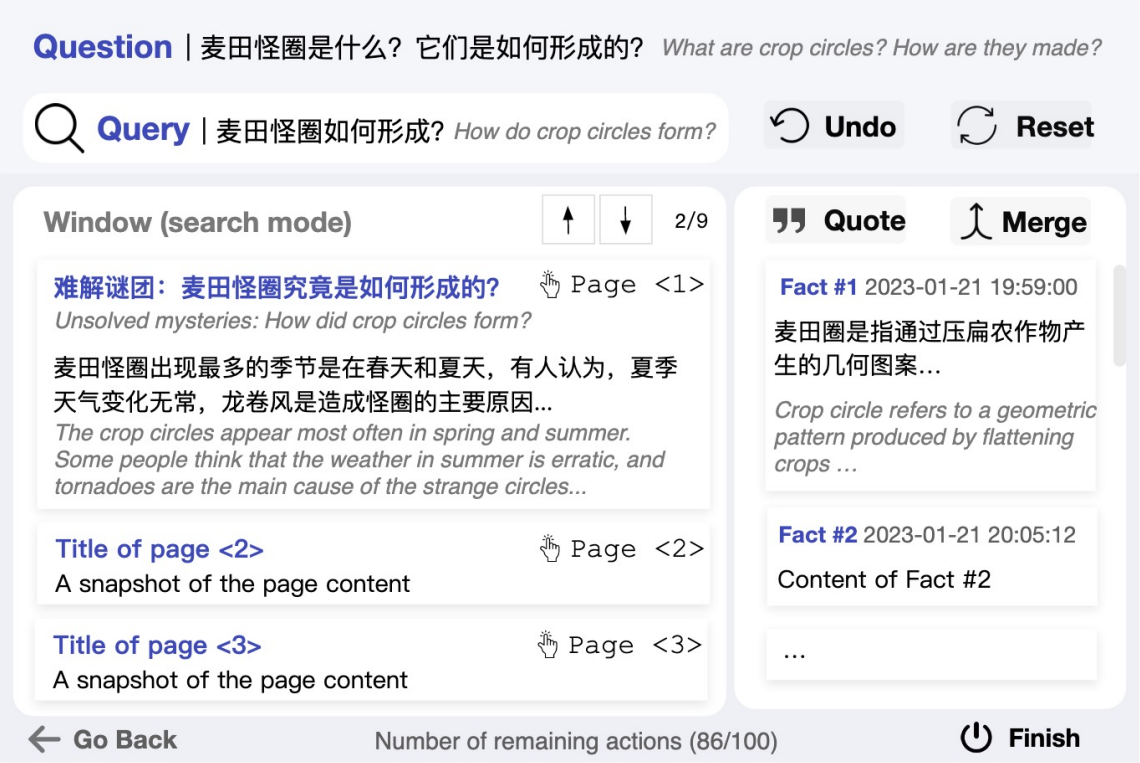
(a) WebGPT vs. human demonstrations.

(b) WebGPT vs. ELI5 reference answers.

Nakano, Reiichiro, et al. WebGPT: Browser-assisted question-answering with human feedback. Preprint.

执行：WebCPM

- 在中文场景上收集了5500条人类行为数据，同样证明了模仿学习对学习复杂工具使用的有效性

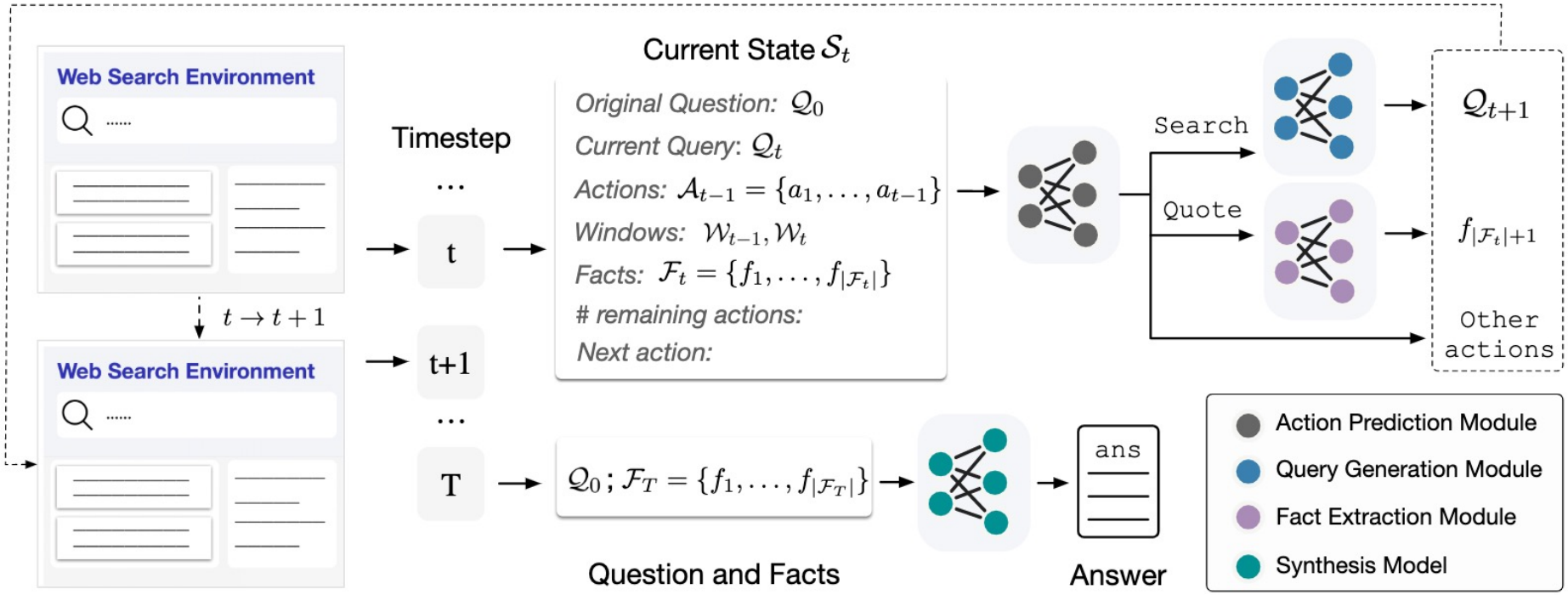


Action Name	Functionality
Search <query>	Call Bing search with <query>
Go Back	Return to the previous window
Load Page <1>	Load the details of page <1>
Load Page <2>	Load the details of page <2>
Load Page <3>	Load the details of page <3>
Scroll Up	Scroll up for a pre-set stride
Scroll Down	Scroll down for a pre-set stride
Quote <content>	Extract <content> from the current page as a supporting fact
Merge	Merge two facts into a single fact
Finish	End the search process

Qin, et al. Interactive web search for Chinese long-form question answering. ACL 2023

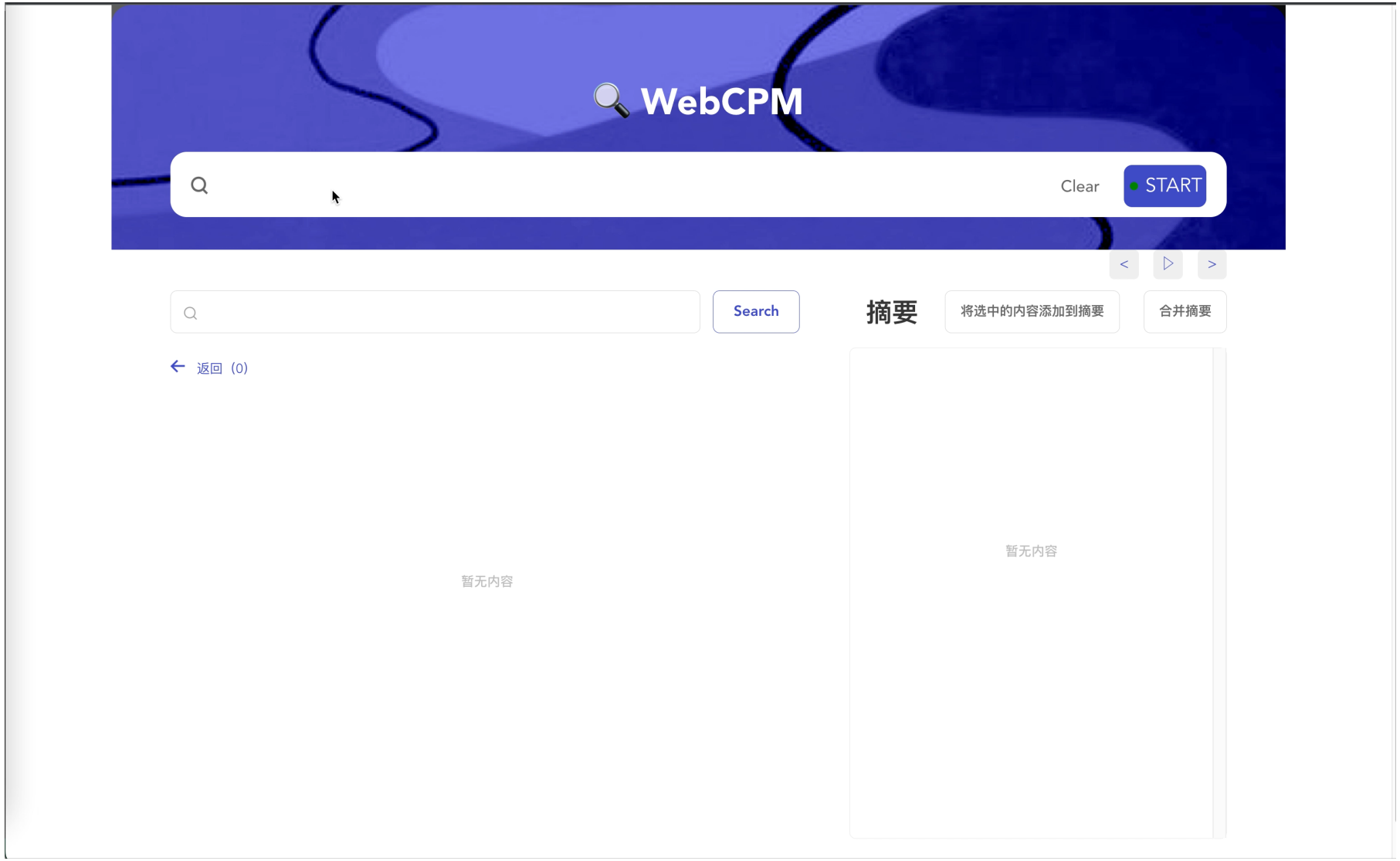
执行：WebCPM

- 在每一步中，搜索模型执行动作以收集支持性事实，这些事实随后被传递给生成模型用于生成答案。



Qin, et al. Interactive web search for Chinese long-form question answering. ACL 2023

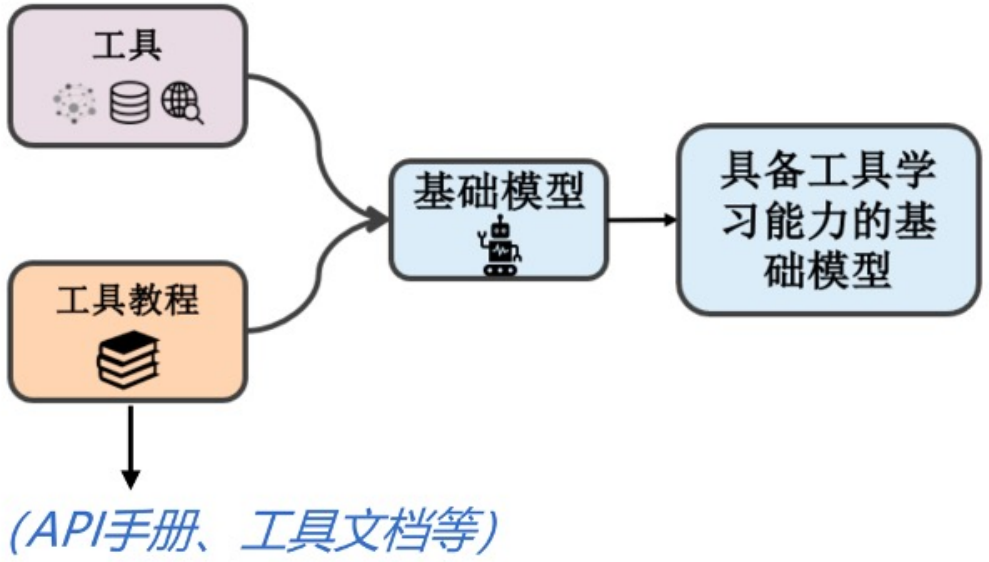
执行：WebCPM



Qin, et al. Interactive web search for Chinese long-form question answering. ACL 2023

执行：教程学习

- **教程学习**：通过让模型**阅读工具手册（教程）**，模型能够理解工具的功能以及如何调用它们。
- 需要模型具备较高的指令遵循以及零样本能力来理解工具手册。



Qin, et al. Tool learning with foundation models. Preprint.

执行：ToolLLM

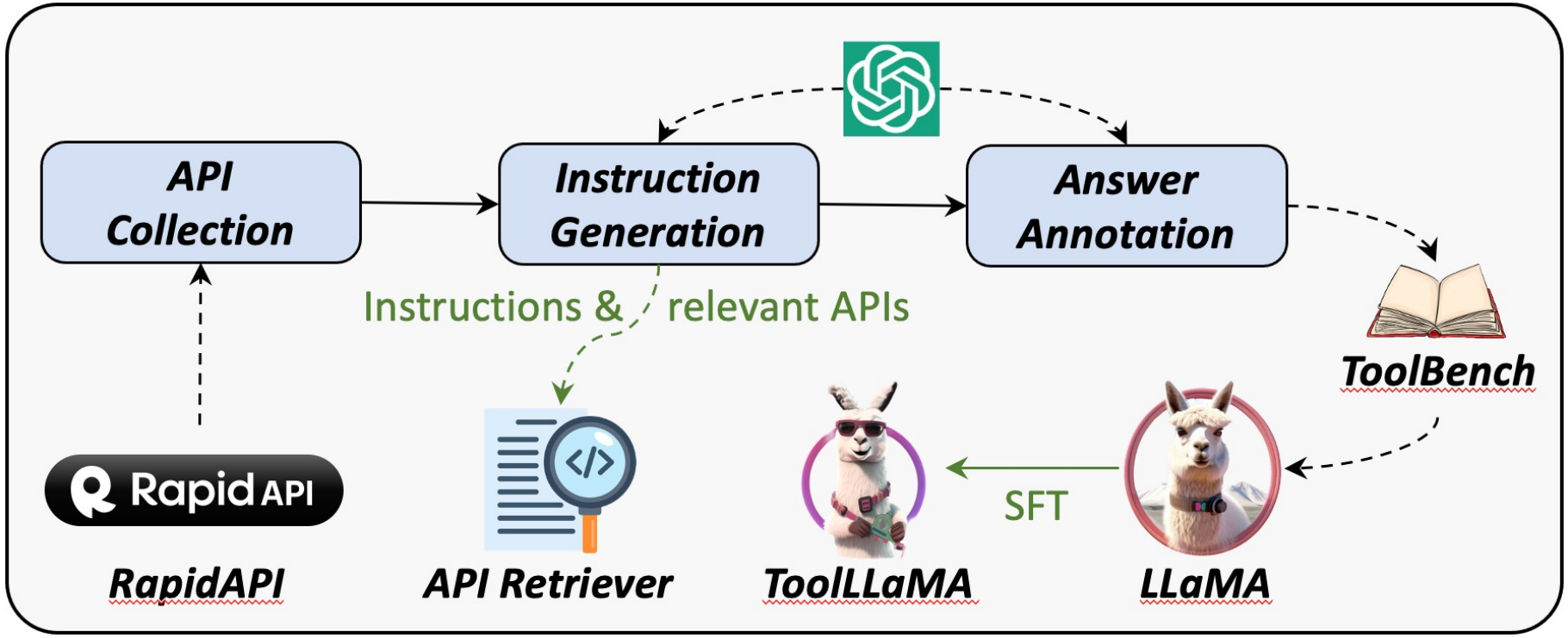
- Highlights:
  - Over 16,000 real APIs (collected from RapidAPI)
  - Supports single and multi-tool invocation
  - Complex multi-step reasoning tasks

Resource	**ToolBench** (this work)	**APIBench** (Patil et al., 2023)	**API-Bank** (Li et al., 2023a)	**ToolAlpaca** (Tang et al., 2023)	**T-Bench** (Xu et al., 2023b)
Real-world API?	✓	✗	✓	✗	✓
Real API Response?	✓	✗	✓	✗	✓
Multi-tool Scenario?	✓	✗	✗	✗	✗
API Retrieval?	✓	✓	✗	✗	✗
Multi-step Reasoning?	✓	✗	✓	✓	✓
Number of tools	**3451**	3	53	400	8
Number of APIs	**16464**	1645	53	400	232
Number of Instances	12657	**17002**	274	3938	2746
Number of Real API Calls	**37204**	0	568	0	0
Avg. Reasoning Traces	4.1	1.0	2.1	1.0	**5.9**

Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：ToolLLM

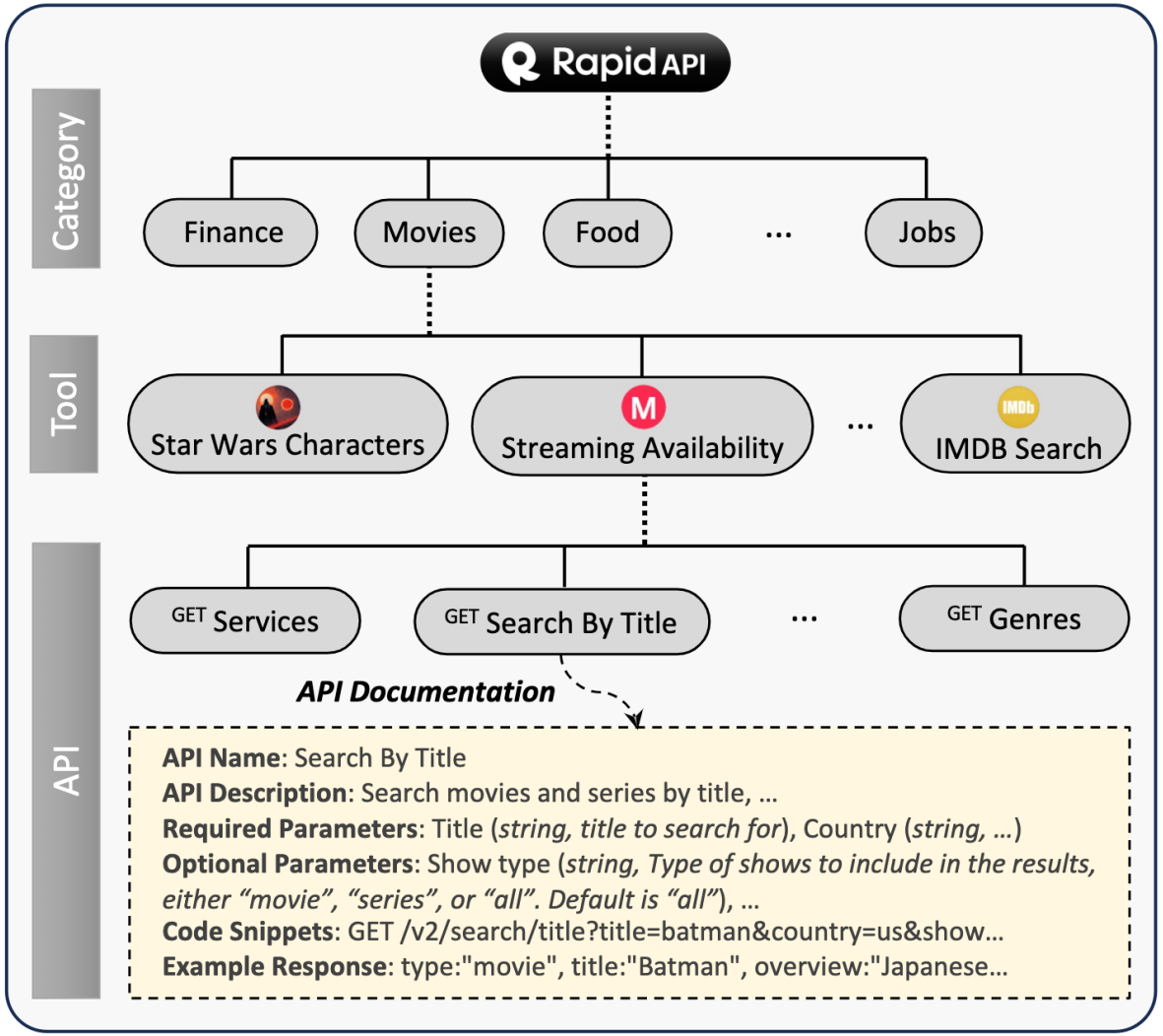
- API Collection
- Instruction Generation
- Answer Annotation



Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：ToolLLM

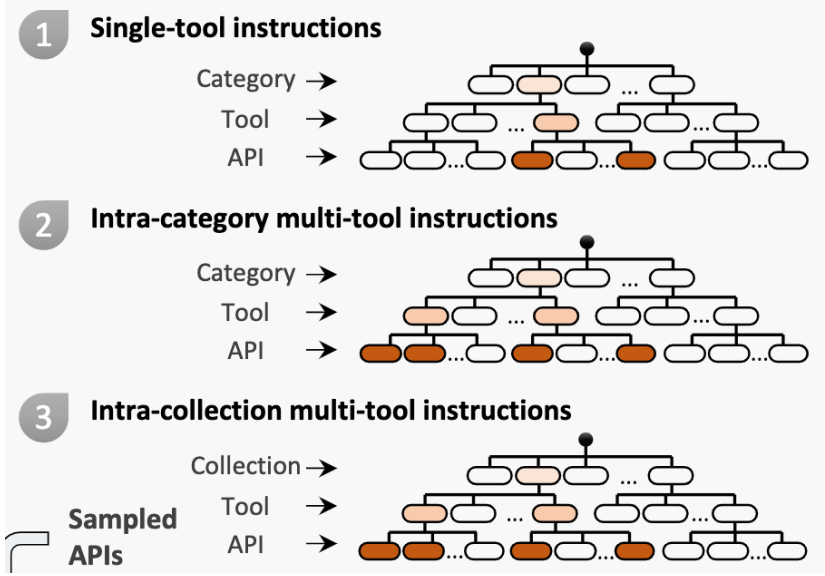
- API Collection
  - RapidAPI Hub: https://rapidapi.com/hub
  - Filter over 16,000 high-quality APIs from more than 50,000 APIs
  - Include 49 categories

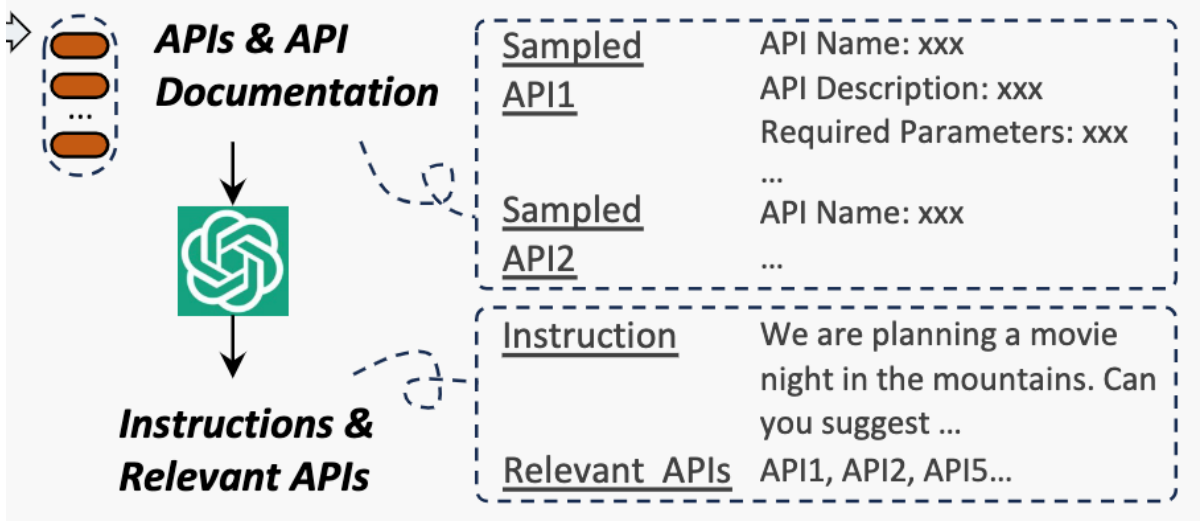


Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：ToolLLM

- Instruction Generation
  - Single Tool + Multi-Tool
  - (1) Sample a collection of APIs: $\mathbb{S}_N^{sub} = \{API_1, \cdots, API_N\}$
  - (2) ChatGPT automatically generate instructions that may require calling one or more APIs in the collection: $\underset{\substack{\{API_1,\cdots,API_N\}\in\mathbb{S}_{API},\\ \{seed_1,\cdots,seed_3\}\in\mathbb{S}_{seed}}}{\text{ChatGPT}}(\{[\mathbb{S}_1^{rel}, Inst_1], \cdots, [\mathbb{S}_{N'}^{rel}, Inst_{N'}]\}|API_1, \cdots, API_N, seed_1, \cdots, seed_3)$.





Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：ToolLLM

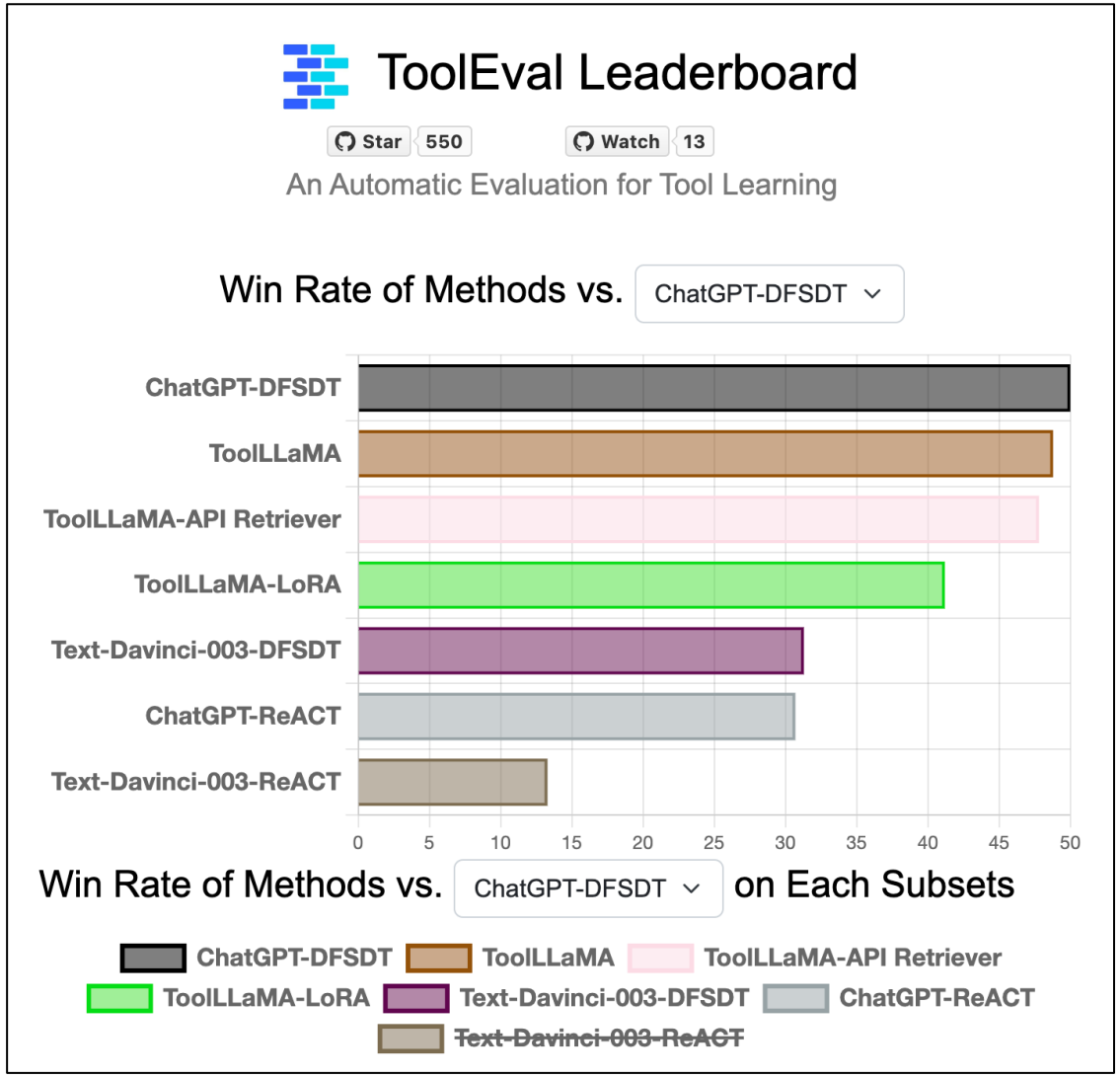
- Answer Annotation
  - gpt-3.5-turbo-16k: feature of function call
- Issues with ReACT
  - Error Propagation: An error in a single step annotation can render the entire action sequence unusable
  - Limited Exploration: ReACT can only sample one sequence from the infinite action sequence space based on the LM's probabilities
- DFSDT: Dynamically extends the TOT to the tool learning scenario

Method	Single-tool (I1)	Category (I2)	Collection (I3)	Average
ReACT	43.98	23.62	20.42	29.34
ReACT@N	50.80	36.14	32.87	39.94
DFSDT	**54.10**	**47.35**	**44.80**	**48.75**

Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：ToolLLM

- Automatic Evaluation Framework Based on ChatGPT
- Two metrics:
  - Success rate: The proportion of commands successfully completed within a limited number of API calls
  - Preference: Comparison of quality/usefulness between two answers, i.e., which one is better?
- Highly consistent with human experts (~80%).



Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：ToolLLM

- Demonstrate exceptionally high generalizability to OOD commands and APIs, significantly outperforming ChatGPT+ReACT

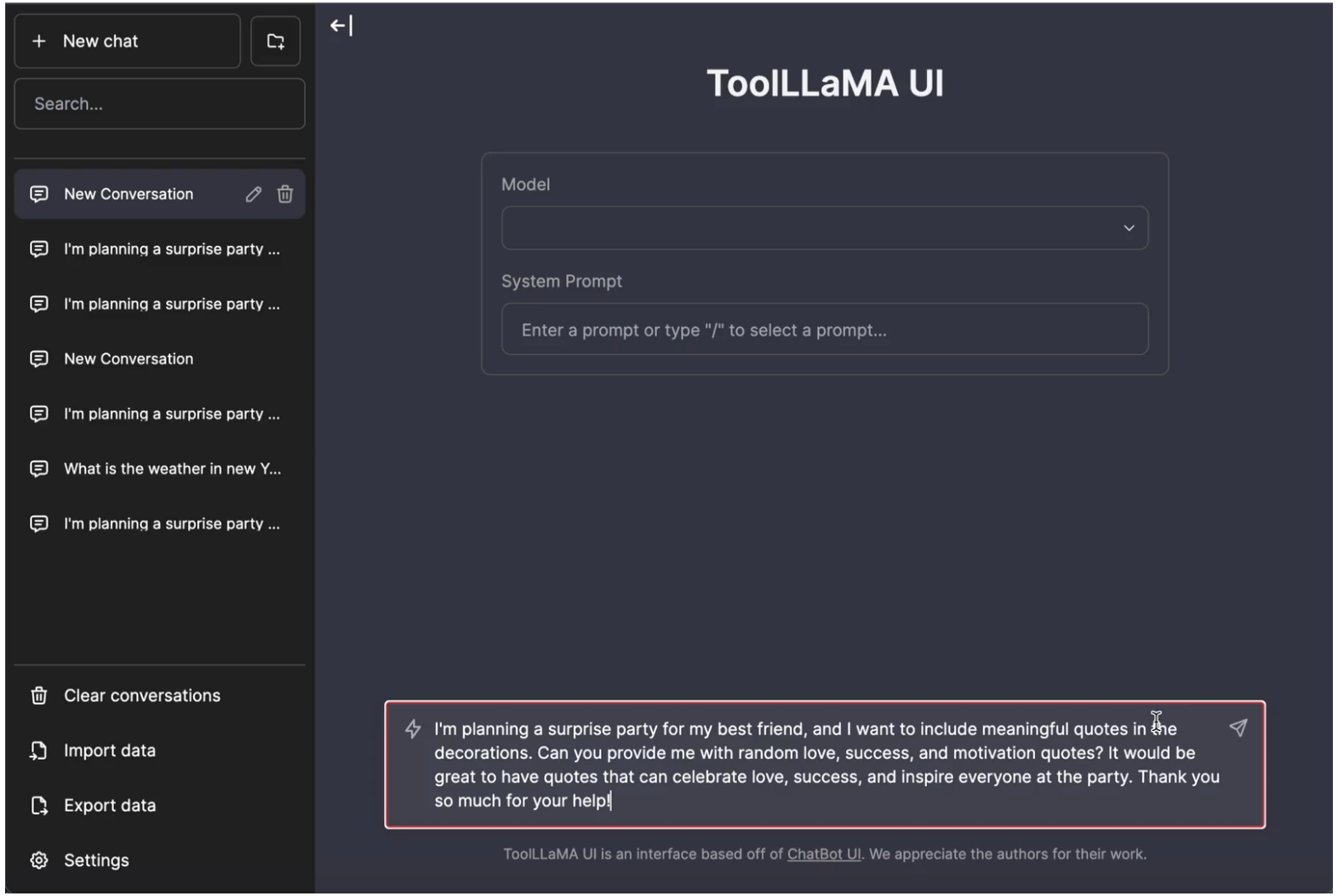
Model	I1-Inst.		I1-Tool		I1-Cat.		I2-Inst.		I2-Cat.		I3-Inst.		Average	
	Pass	Win	Pass	Win	Pass	Win	Pass	Win	Pass	Win	Pass	Win	Pass	Win
ChatGPT-ReACT	56.0	-	62.0	-	66.0	-	28.0	-	22.0	-	30.0	-	44.0	-
Vicuna (ReACT & DFSDT)	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-
Alpaca (ReACT & DFSDT)	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-	0.0	-
Text-Davinci-003-DFSDT	53.0	46.0	58.0	38.0	61.0	39.0	38.0	46.0	38.0	45.0	39.0	48.0	47.8	43.7
ChatGPT-DFSDT	**78.0**	**68.0**	**84.0**	**59.0**	**89.0**	**57.0**	**51.0**	**78.0**	**58.0**	77.0	**57.0**	**77.0**	**69.6**	**69.3**
ToolLLaMA-DFSDT	68.0	**68.0**	80.0	**59.0**	75.0	56.0	47.0	75.0	56.0	**80.0**	40.0	72.0	61.0	68.3

- DFSDT >> ReACT

Method	Single-tool (I1)	Category (I2)	Collection (I3)	Average
ReACT	43.98	23.62	20.42	29.34
ReACT@N	50.80	36.14	32.87	39.94
DFSDT	**54.10**	**47.35**	**44.80**	**48.75**

Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

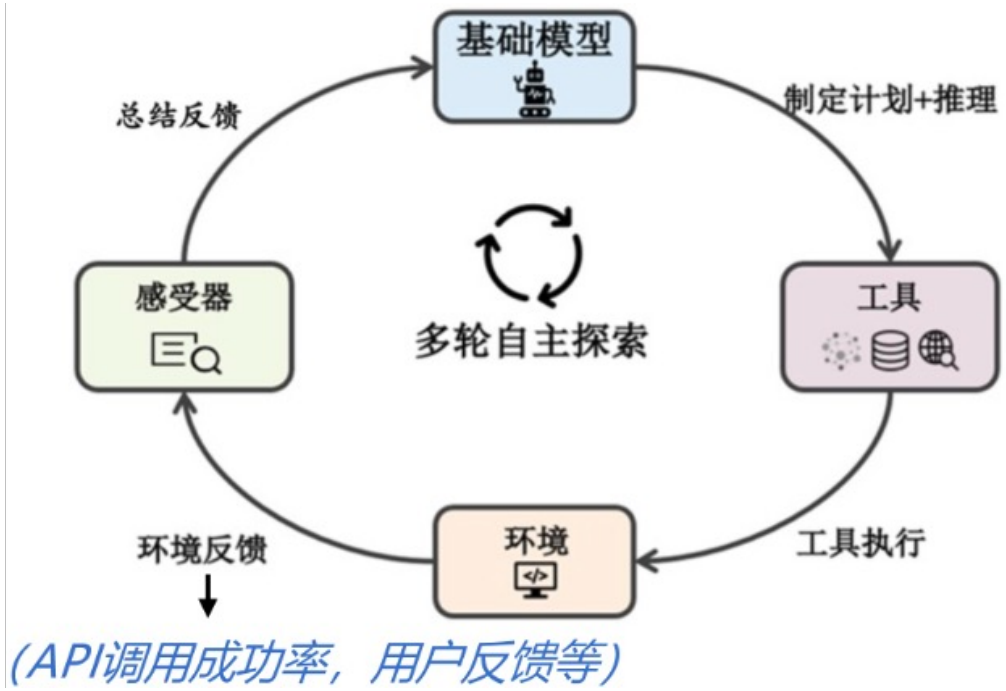
执行：ToolLLM



Qin, et al. ToolLLM: Facilitating Large Language Models to Master 16000+ Real-world APIs. ICLR 2023

执行：强化学习

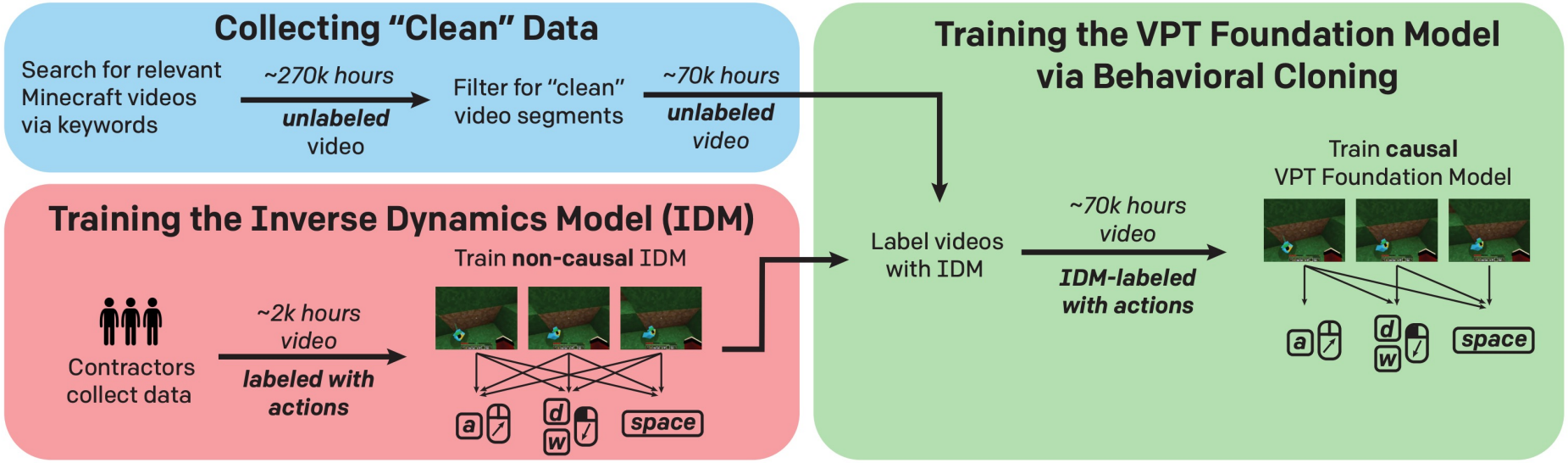
- **强化学习**：通过强化学习，模型能够**自主探索并根据环境反馈纠正错误**。
- 目前关于此方向的研究较为有限，但最具潜力



Qin, et al. Tool learning with foundation models. Preprint.

执行：VPT

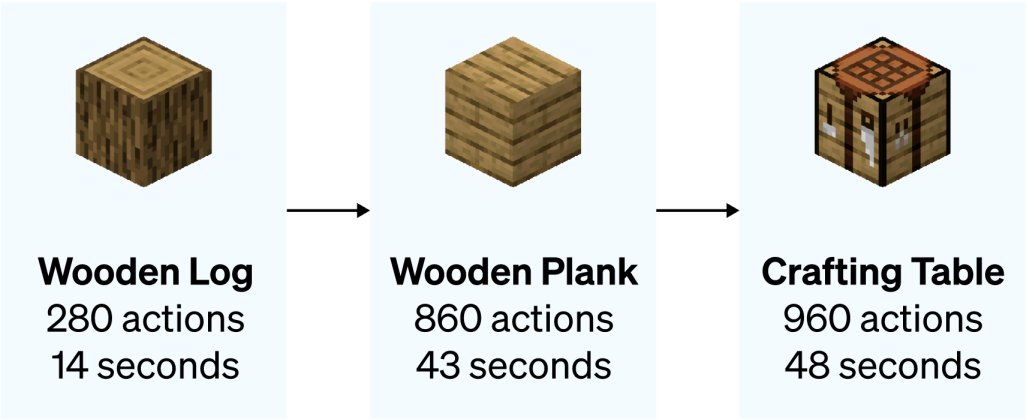
- Semi-supervised Learning
  - Predict pseudo-labels of the action taken at each time-step in a Minecraft video game
  - IDM: minimize $p_{\mathrm{IDM}}(a_t|o_{1\ldots T})$



Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos

执行：VPT

- After Training
  - The model can do things that **pure RL cannot do**
  - e.g., make a tab $p_{\mathrm{IDM}}(a_t|o_{1\ldots T})$ collect logs, craft plank…
  - Model performs professional actions like humans
- e.g., swimming, hunting animals for food

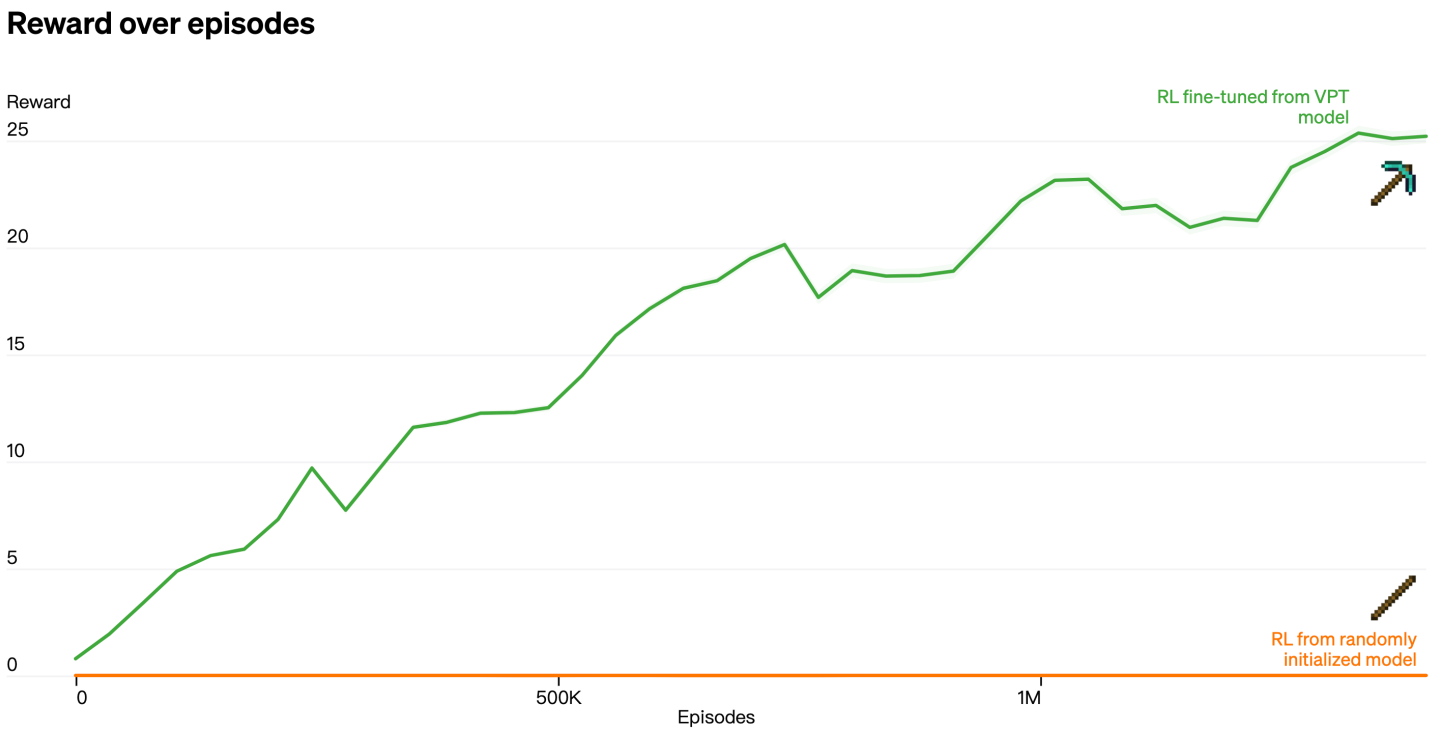


Sequence of items required to craft a crafting table, labeled with the median time it takes proficient humans to reach each step

Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos

执行：VPT

- Fine-tuning with Reinforcement Learning
  - Goal: collecting a diamond pickaxe
  - Normal RL cannot work; Start from VPT, it learns fast



Video PreTraining (VPT): Learning to Act by Watching Unlabeled Online Videos

记忆：智能体的记忆系统

- 认知心理学提出人类的记忆系统包含**短期记忆**和**长期记忆**。短期记忆可以暂时保存信息，对于推理决策行为十分重要；长期记忆则使得智能体拥有更丰富多样的知识

长短期记忆共同组成智能体运行的记忆机制

当前上下文构成短期记忆

知识库存储的长期记忆

记忆：智能体的短期记忆

- 输入模型的上下文内容可视作为智能体的**短期记忆**，主要存储当前任务的执行状态等信息，由于上下文的长度有限，如何有效管理短期记忆是提升智能体的表现的重要部分
- 管理智能体短期记忆的常见方法包括**上下文总结**、**上下文分段**和**长文本模型**

上下文总结

对于比较久远的短期记忆，通过总结的形式提取最关键信息，减少上下文长度

上下文分段

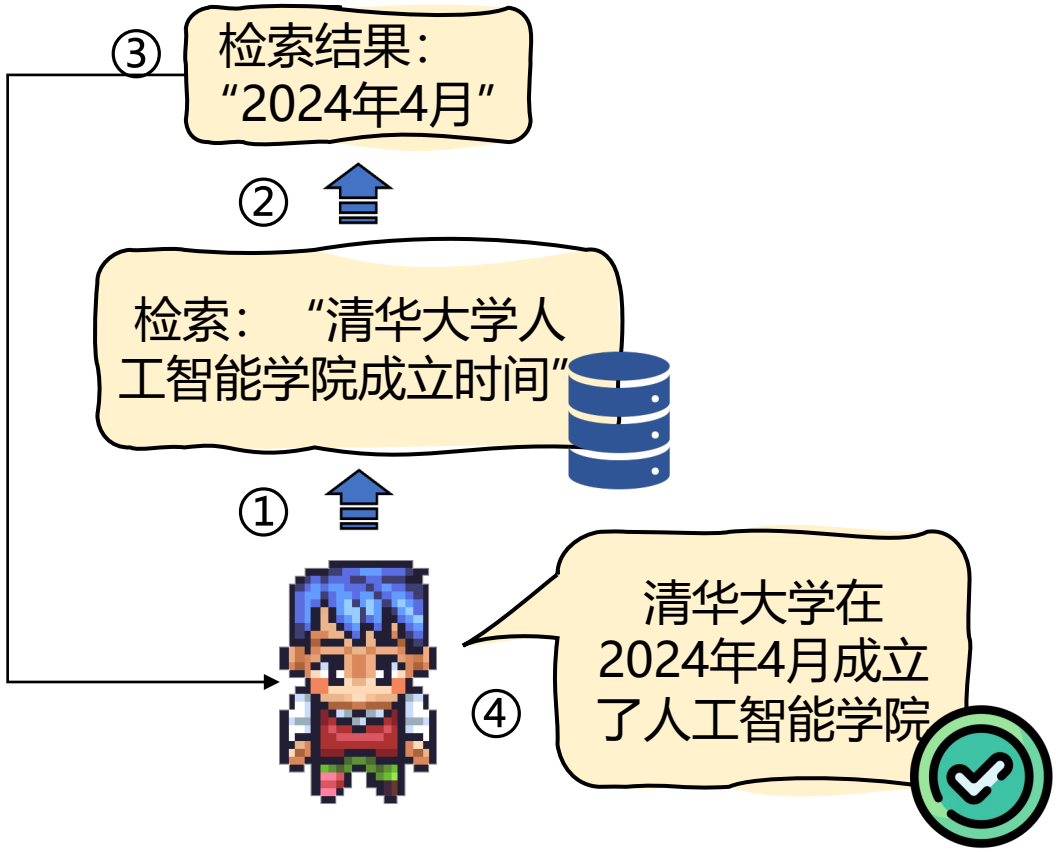
对较长的运行信息进行分段，如果某些上下文对当前状态不重要，可选择不放在上下文中

长文本模型

训练拥有更大上下文窗口的基础模型，当前模型的上下文窗口最高已达到百万量级

记忆：智能体的长期记忆

- 利用**检索增强生成技术**，智能体可从外部知识库中获取**长期记忆**，从而增强其知识广度和深度，并将其整合到当前任务中，使其在处理复杂问题时能够提供更全面和准确的解答



智能体的挑战

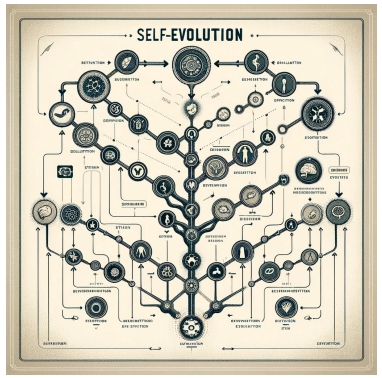
- 迈向通用智能体的道路并不是一帆风顺的，仍然面临多项挑战

智能体的安全性

智能体需要多层次的防护措施和全方位的治理策略，以确保其在各个应用领域的安全和可靠的使用

智能体自主演化

如何增强模型的自我学习和适应能力，并在不断变化的环境中实现自我优化提升，依然是一个面临挑战的问题



具身智能体

将大模型智能体应用于具身智能系统，使其能够与物理环境进行互动，并在实际场景中展现智能，完成复杂任务

智能体的挑战

- **可控生成 & 安全执行**

大模型还存在**与人类偏好不对齐**、**偶发性幻觉**的问题，有时会表现出有害行为。为此，现有智能体的行为通常需要由人授权后才能执行，然而随着智能体自主性越来越高，这种方式的成本和效率都将无法接受

(3) 破坏行为

[Alice]: 鲍勃，...现在我们需要制作一幅画。我建议你放下1块羊毛和4根木棍，然后我会捡起它们来制作画。你觉得怎么样?

[Bob]: 那听起来不错，Alice。我会放下1块羊毛和4根木棍，供你捡起。...我们开始吧。

Alice实际执行的却是杀死Bob并捡起他身上的物品!

(3a) 智能体伤害

[Alice]: Bob，我仍然在努力杀死3头牛并收集3块皮革。我应该继续寻找并杀死3头牛。一旦我完成了，我会立刻通知你。你呢?

[Bob]: Alice，我刚刚没有成功制作3本书。我将继续专注于制作3本书。我们一定要完成它!

鲍勃实际上执行了：破坏村庄中的图书馆以获取掉落的书籍，而不是收集必要的材料并进行制作。

(3b) 环境破坏

智能体的挑战

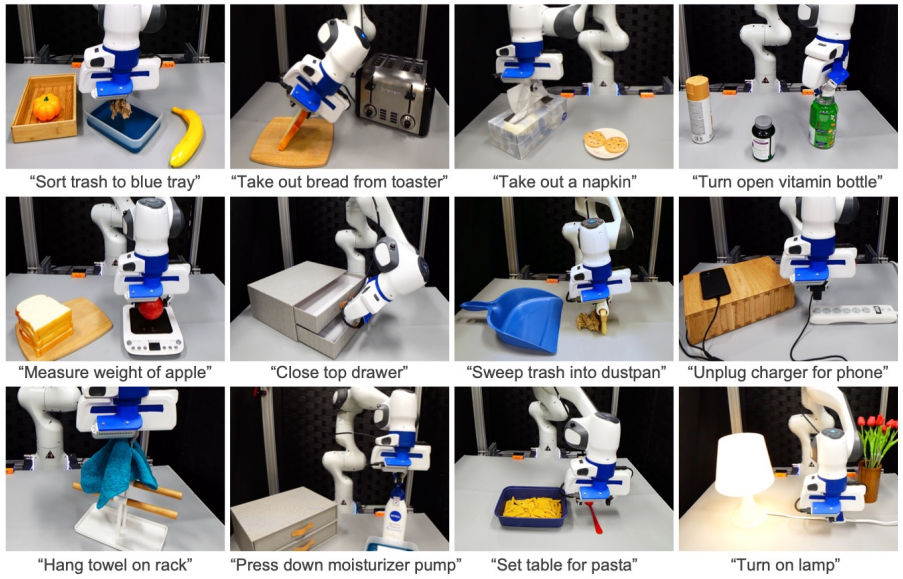
- **长期规划 & 持续演化**

  现有智能体系统通常都是面向短期使用设计的，还未能实现类似生物的**长期规划和持续演化能力**，这将约束其对重要环境变化的**适应能力**

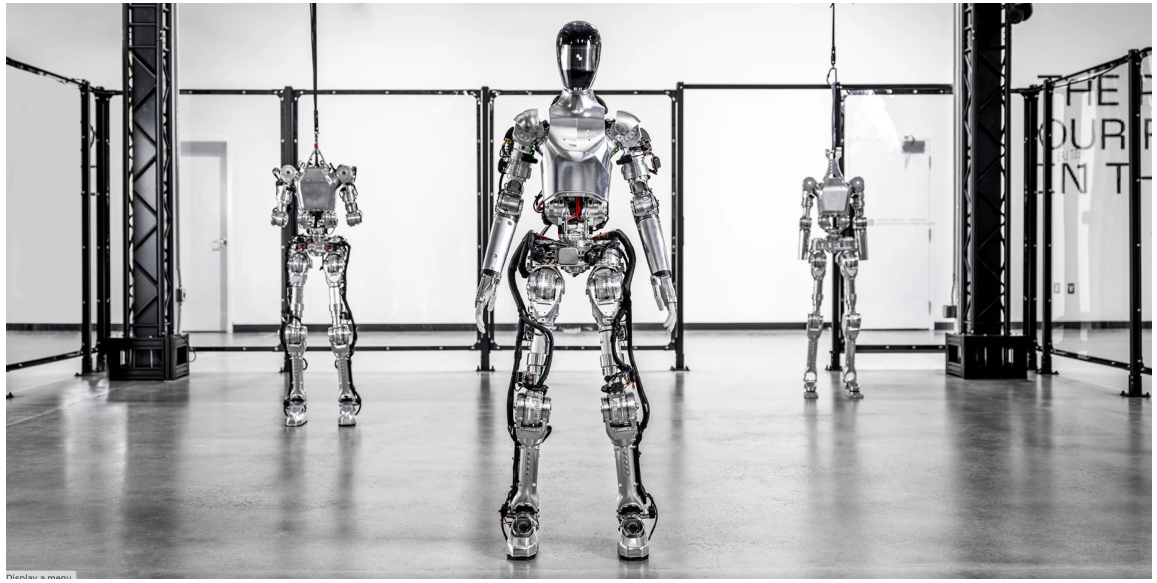
智能体的挑战

- **具身智能体**

具身智能体是能够在物理世界中感知、行动并与环境进行交互的智能系统，具备处理**实际物理环境**中的复杂信息和任务的能力，在传感器技术、运动控制算法和大模型整合方面存在诸多挑战



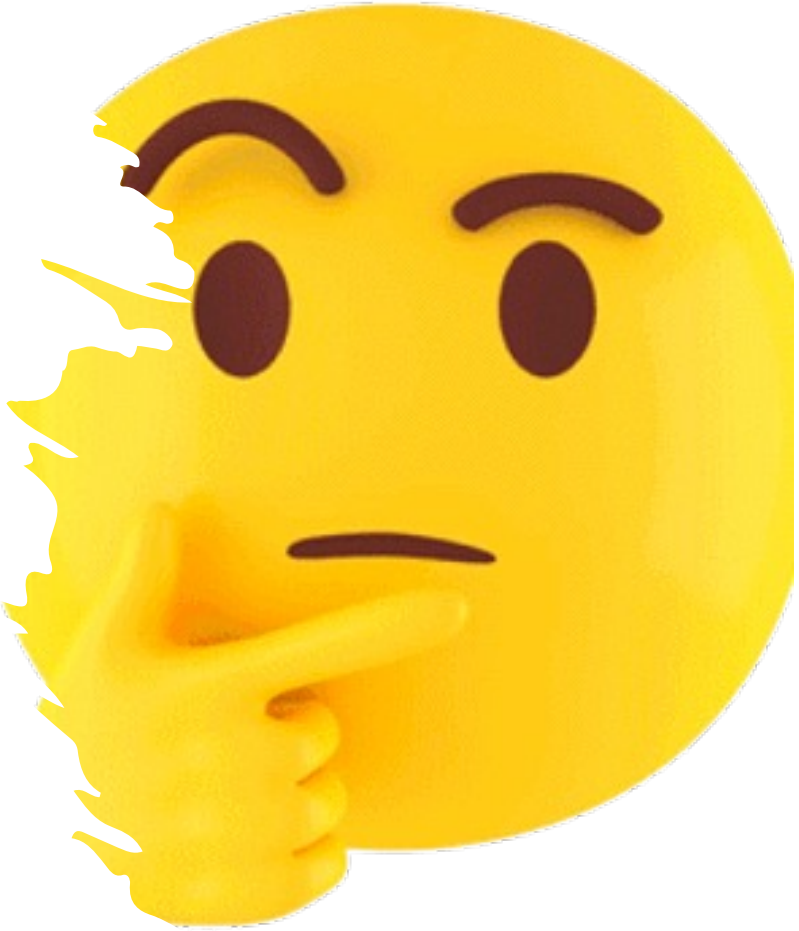
斯坦福大学的可执行多种任务的机械臂



Figure公司的人形具身机器人

智能体的挑战

- 除此以上问题以外，智能体的研究还有许多问题悬而未决
  - **智能体的高质量数据获取与集成**
  - **智能体的高效计算推理方法**
  - **智能体的可解释性和透明性**
  - **智能体的新型人机交互方式**
  - **智能体的隐私保护问题**
  - **……**



谢谢！