AI Agent

李宏毅

免責聲明：AI Agent 是一個被廣泛使用的詞彙，故本課程中所講的 AI Agent 不一定跟其他地方一樣

今天使用 AI 的方式

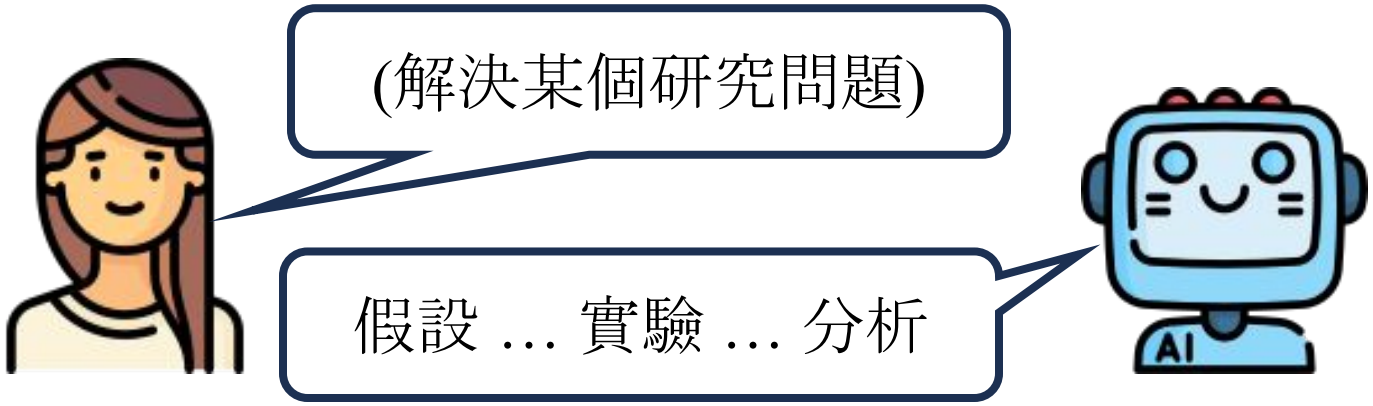
人類給予明確指令

AI 一個口令
一個動作

AI Agent

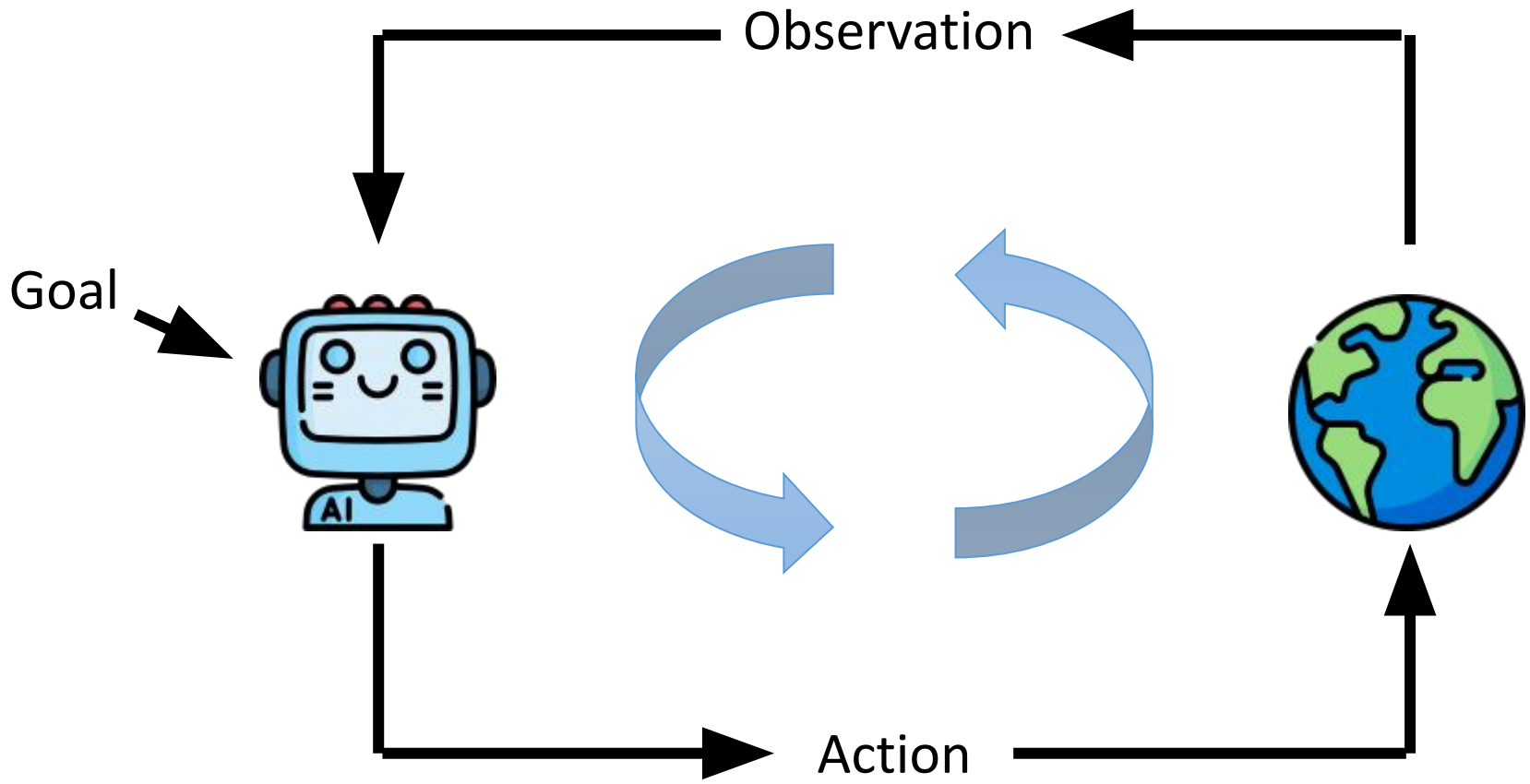
人類給予目標

AI 自己想辦法達成

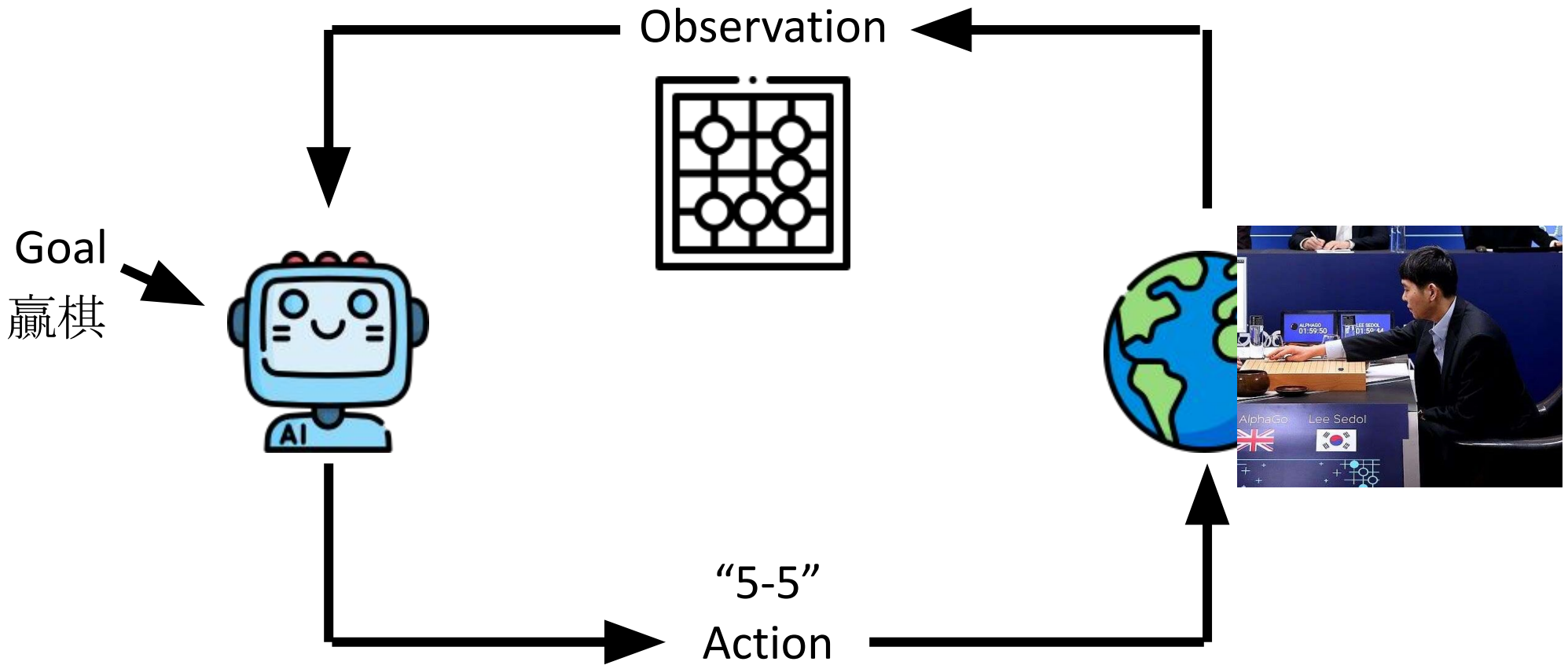


需要多步驟、靈活調整計畫

AI Agent



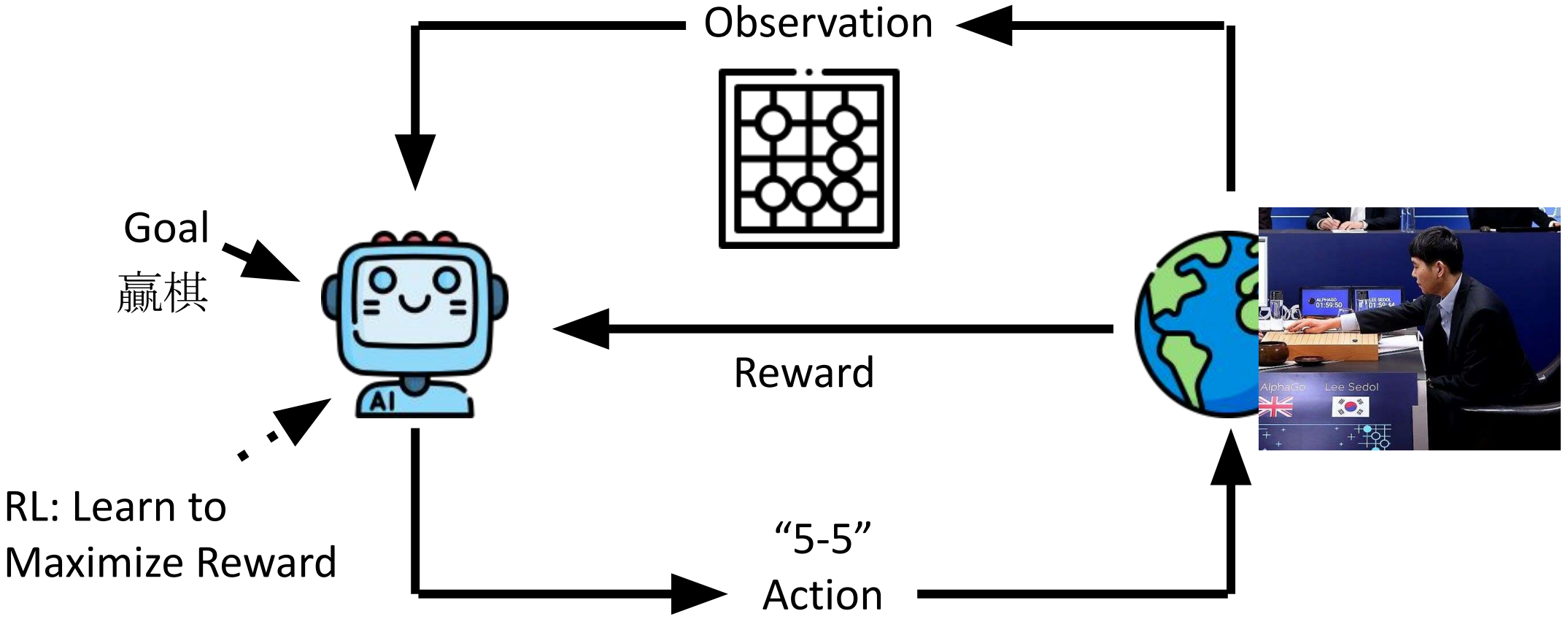
AI Agent (AlphaGo)



好像在那裡聽過這個段落？這是 Reinforcement Learning (RL) 常見開場

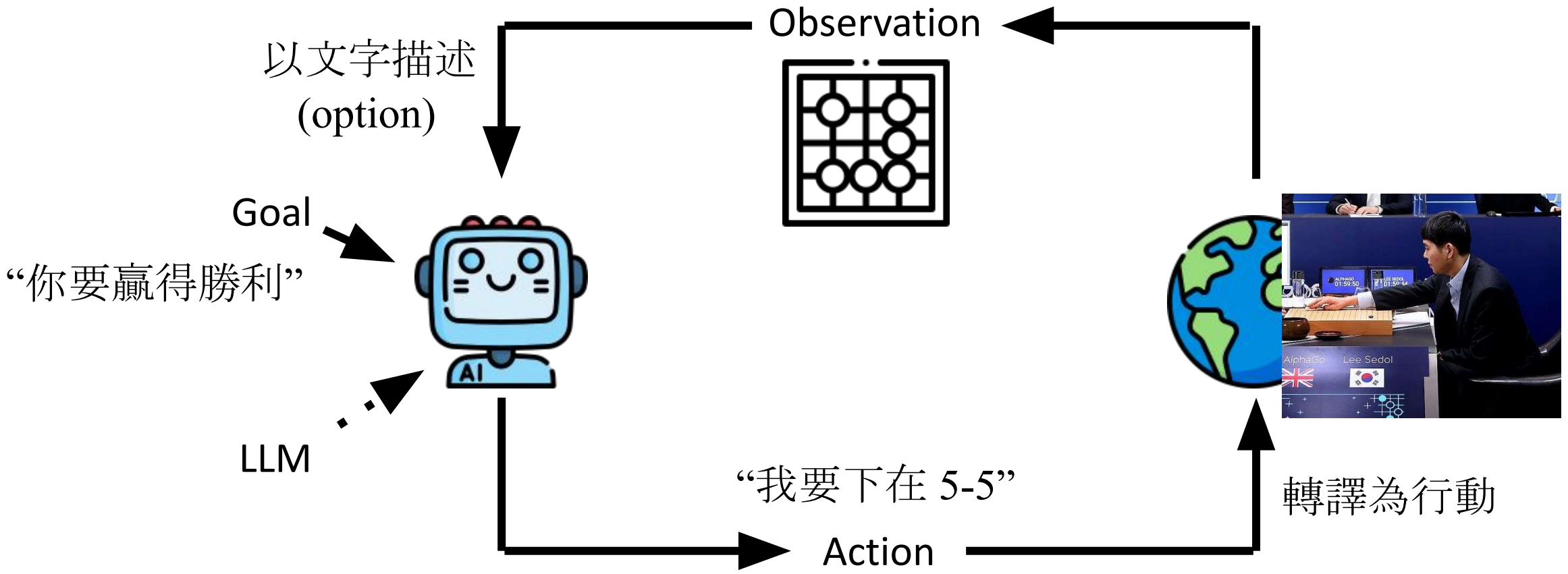
(RL: Reinforcement Learning)

如何打造 AI Agent? RL?



侷限:需要為了每一個任務以 RL 訓練模型

如何打造 AI Agent? 直接用 LLM！

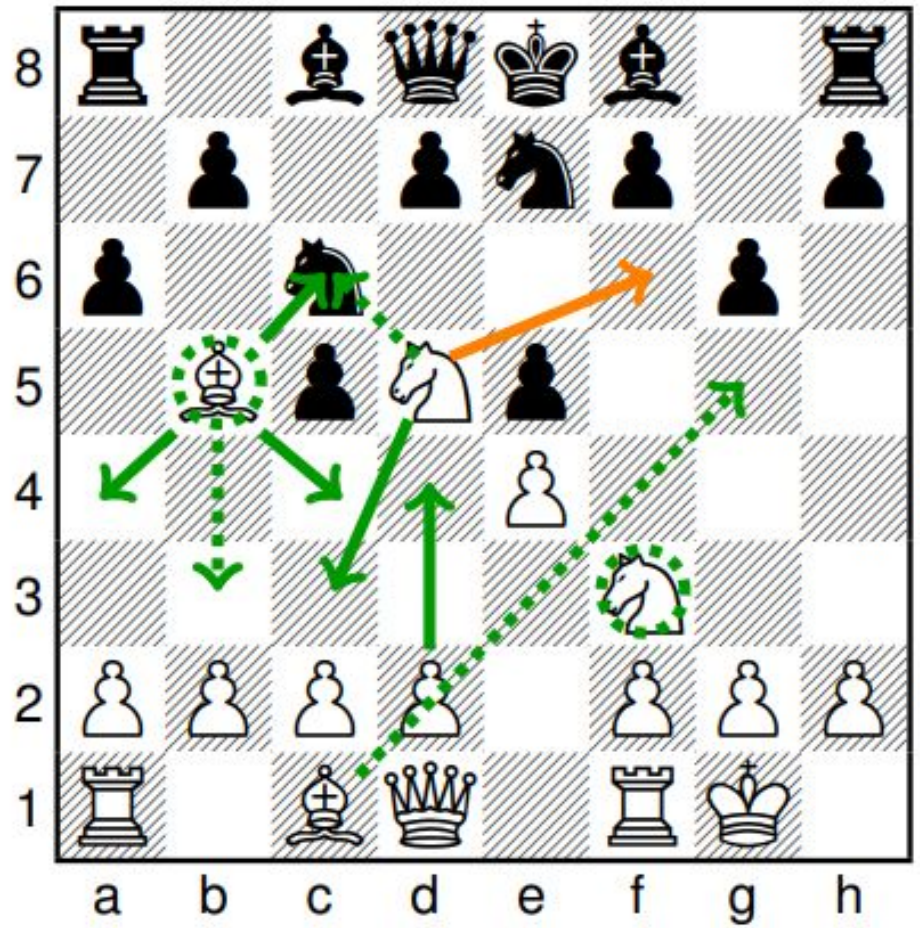


以 LLM 直接實踐人類對於擁有 Agent 的渴望

LLM 能不能下棋？

- BIG-bench

https://arxiv.org/abs/2206.04615



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

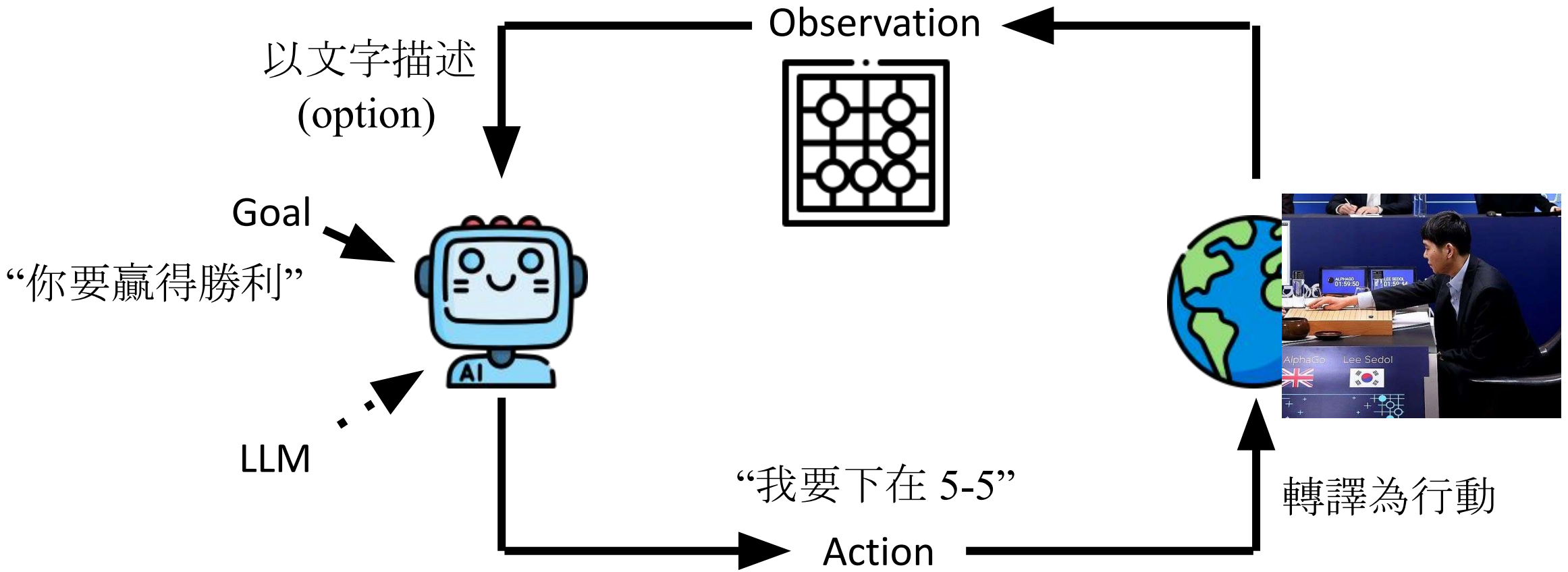
1. e4 c5 2. Nf3 e5 3. Nc3 Nc6 4. Bb5 Nge7 5. O-O g6 6. Nd5 a6 7.

LLM 能不能下棋？

https://youtu.be/JHq4EKMg7fI?si=izKsH-GCVnZkooq_

ChatGPT vs DeepSeek: CRAZY Chess

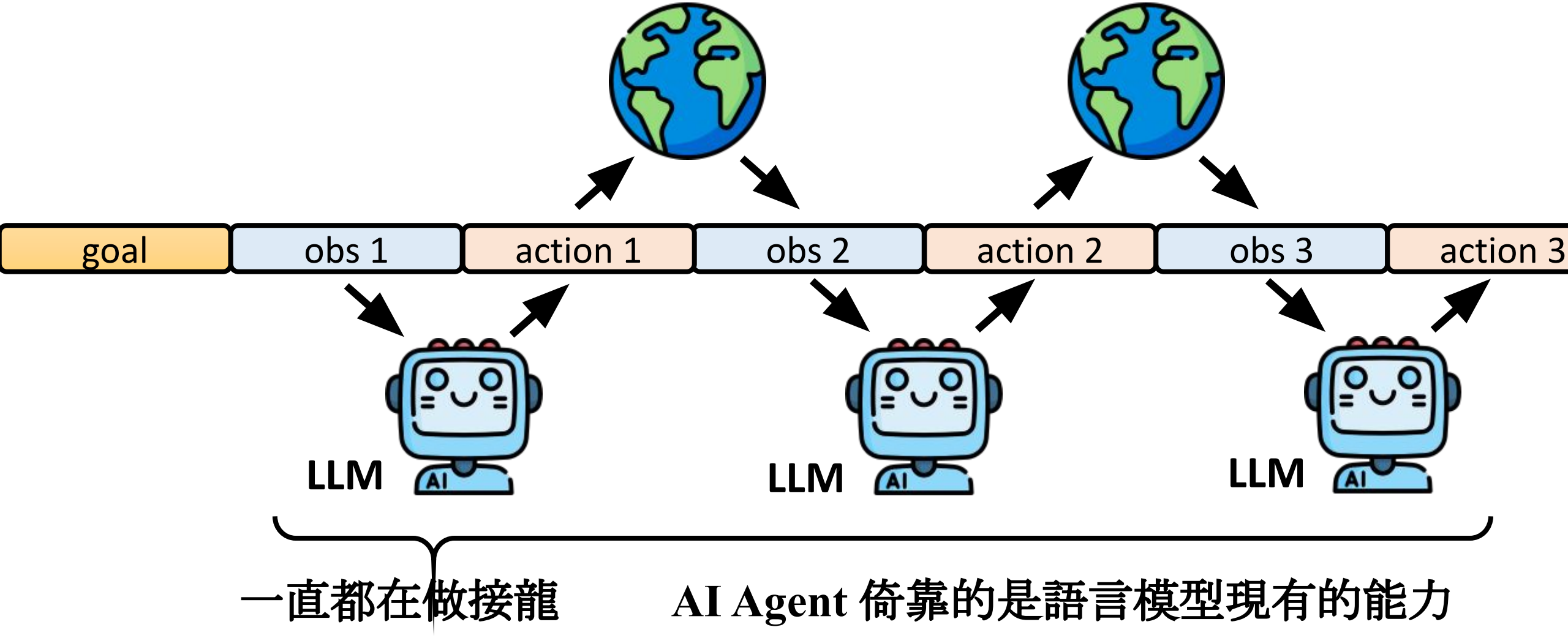
如何打造 AI Agent? 直接用 LLM！



以 LLM 直接實踐人類對於擁有 Agent 的渴望

還有多遠？
還可以多做什麼？

從 LLM 的角度來看 Agent 要解的問題



請注意在這堂課中
沒有任何模型被訓練

AI Agent 不是最近才熱門

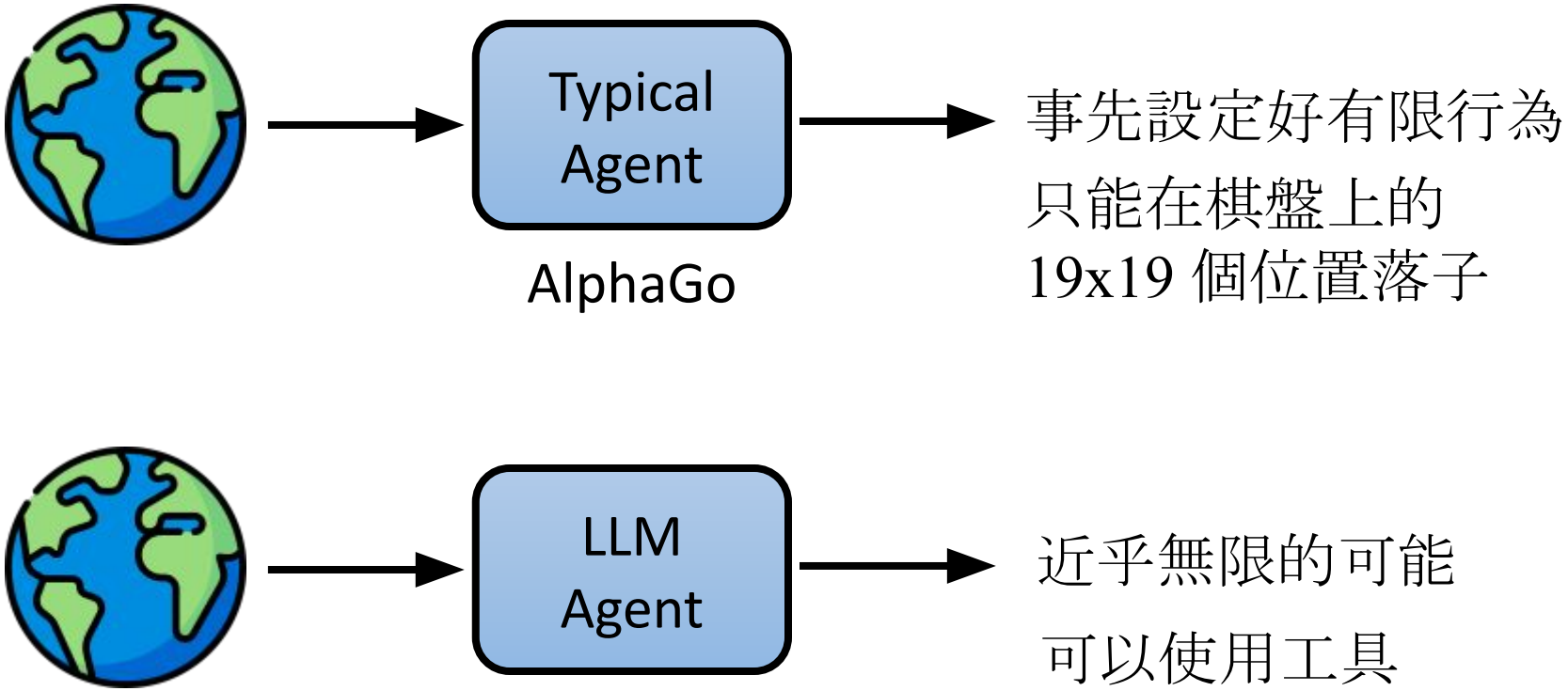
- 2023 年春天曾經爆紅過一次

AutoGPT, AgentGPT,
BabyAGI, Godmode …

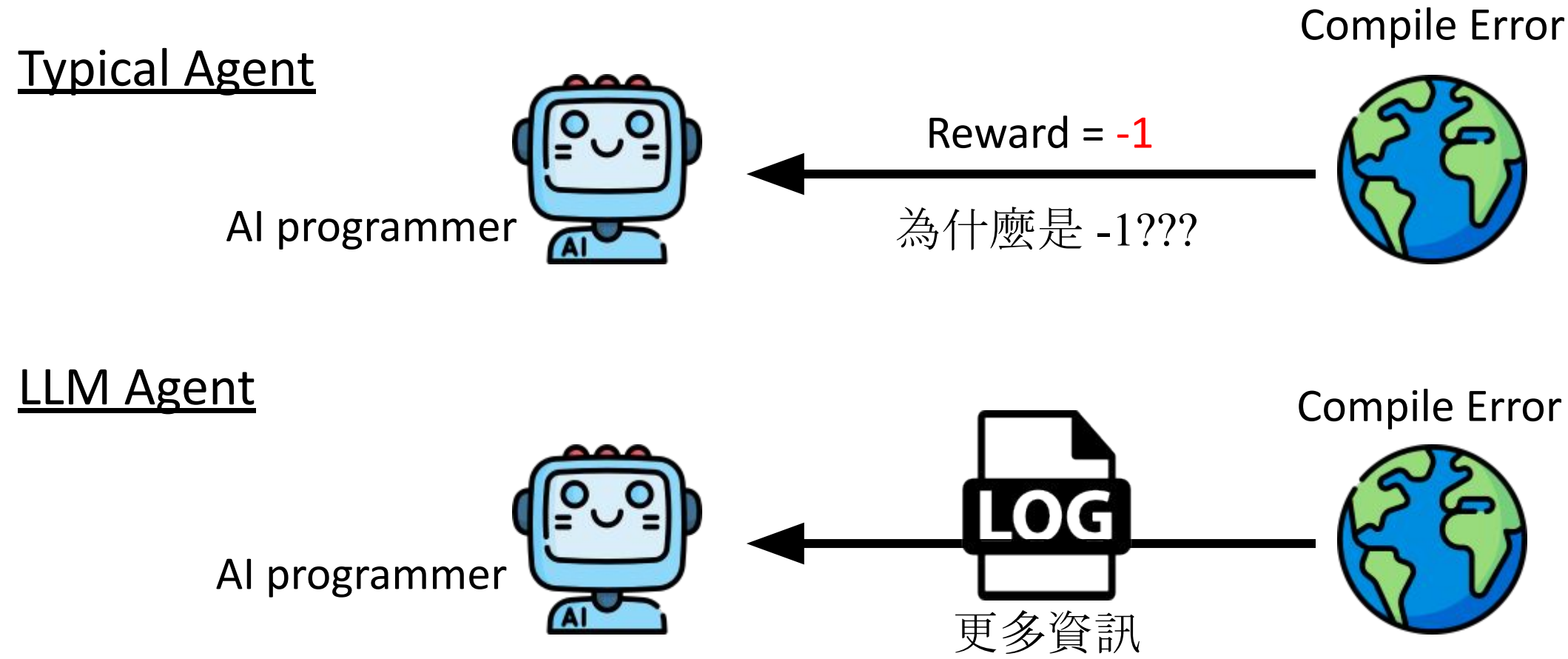
https://youtu.be/eQNADlR0jSs?si=4yGZEluAUzKK2VD0

【生成式AI 2023】讓 AI 做計劃然後自己運行自己

以 LLM 運行 AI Agent 的優勢



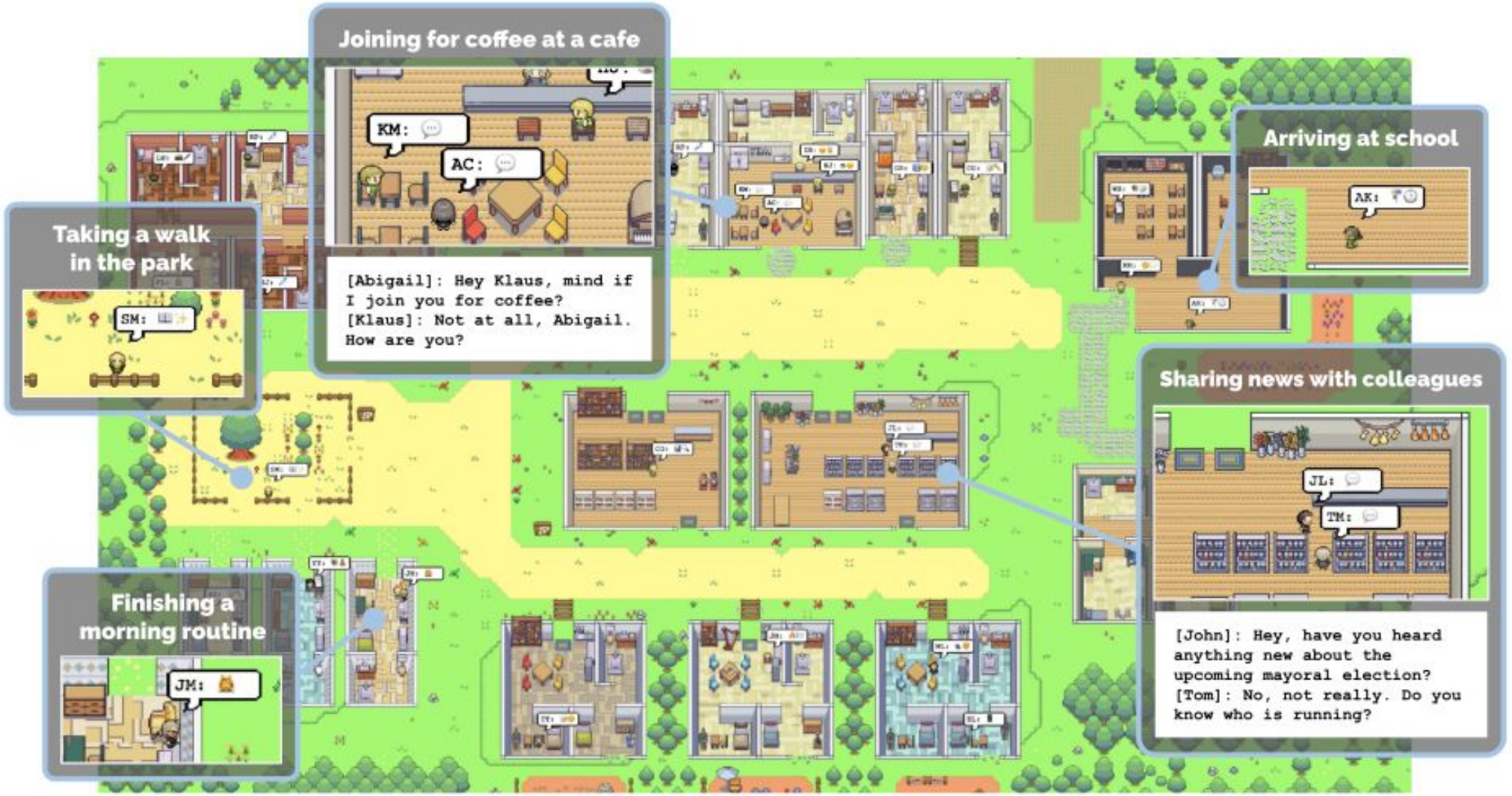
以 LLM 運行 AI Agent 的優勢



https://youtu.be/G44Lkj7XDsA?si=cMbKG3tqPbIgnnBq

https://arxiv.org/abs/2304.03442

AI Agent 舉例：AI 村民組成的虛擬村莊



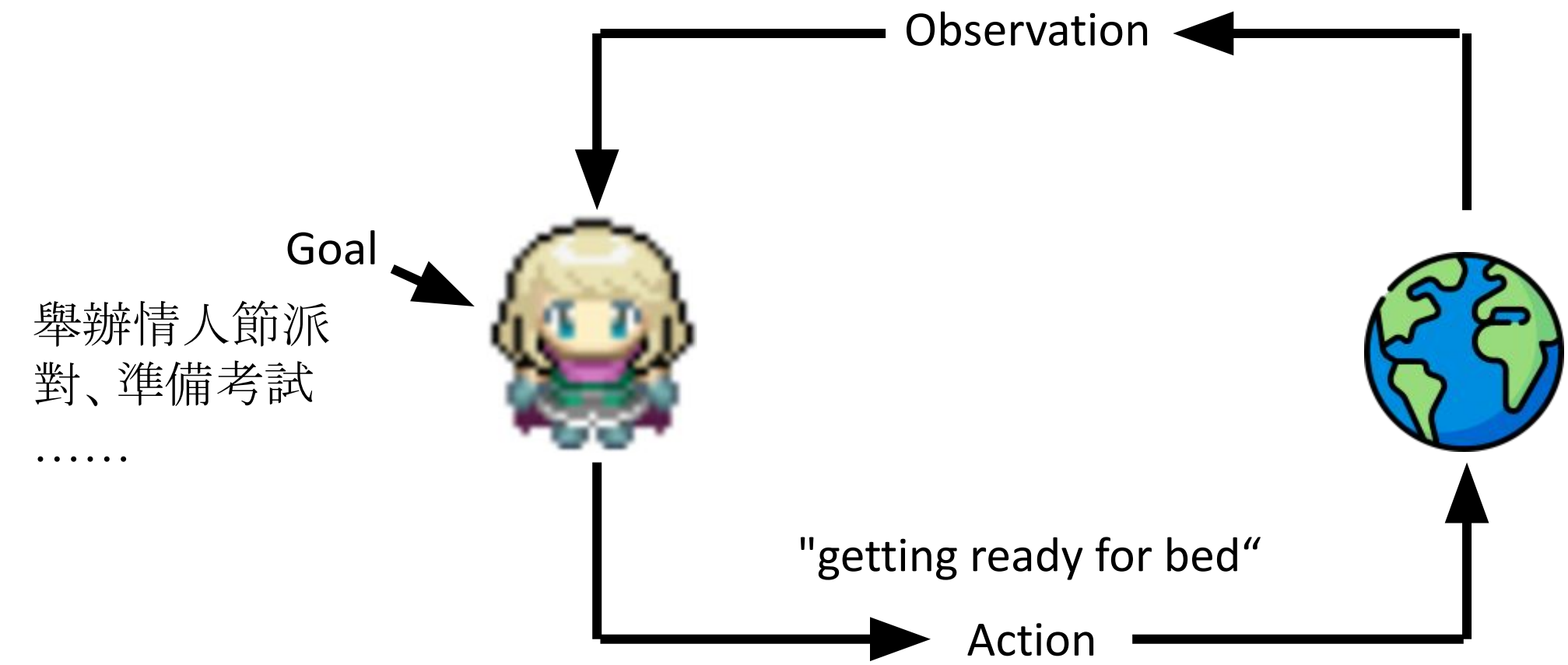
[node_749] 2023-02-13 15:33:20: **Eddy Lin is studying music theory**

[node_748] 2023-02-13 15:33:20: **cooking area is idle**

[node_747] 2023-02-13 15:33:20: **kitchen sink is idle**

[node_746] 2023-02-13 15:33:20: **behind the cafe counter is idle**

[node_745] 2023-02-13 15:32:10: **Isabella Rodriguez is gathering decorations**



AI Agent 舉例：Minecraft 中的 AI NPC

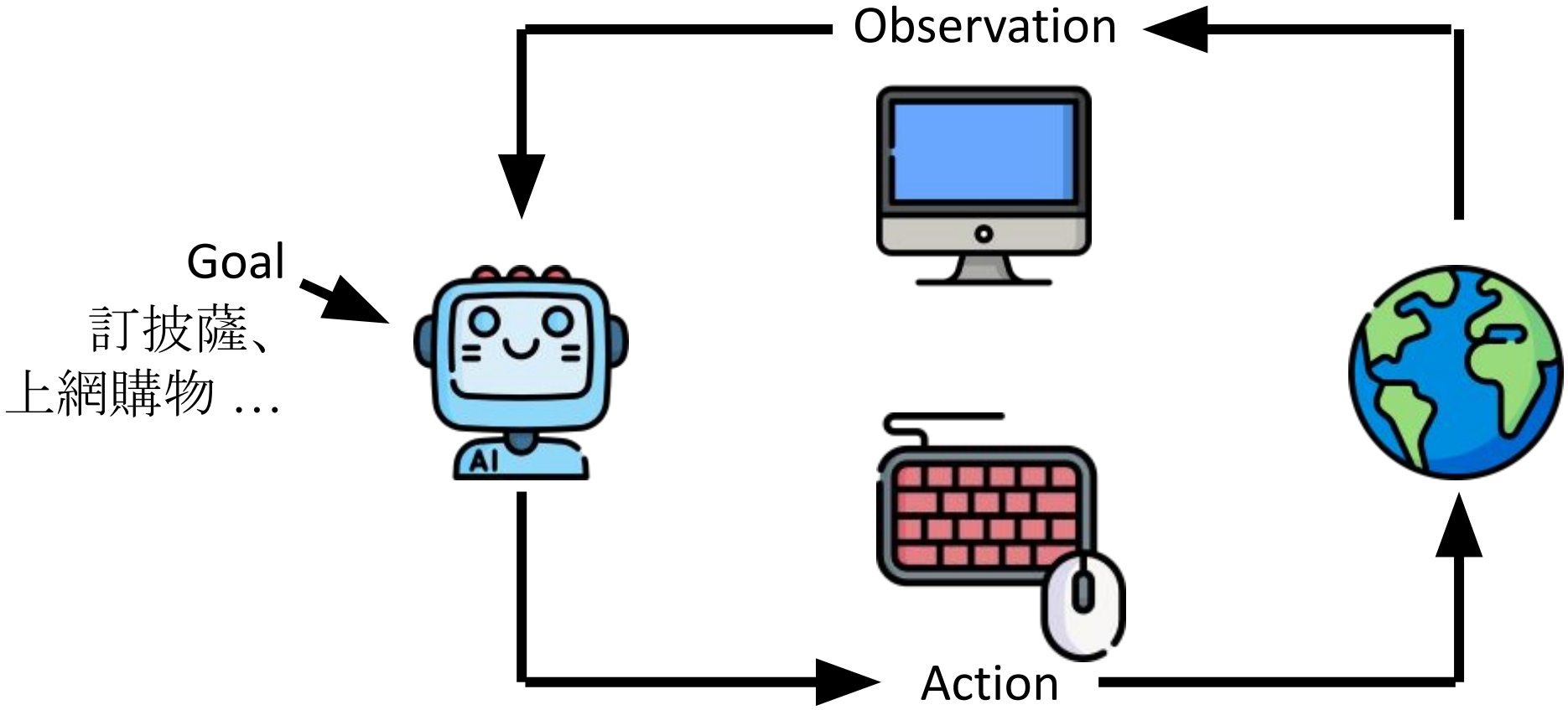
1000 AI NPCs simulate a CIVILIZATION in Minecraft

https://www.youtube.com/watch?v=2tbaCn0Kl90

AI Agent 舉例：讓 AI 使用電腦

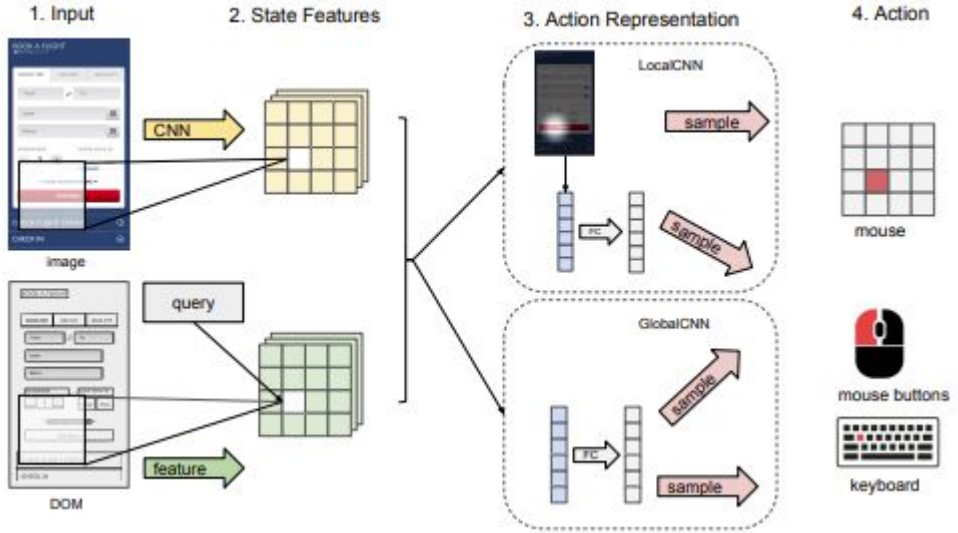
Computer Use,
Operator

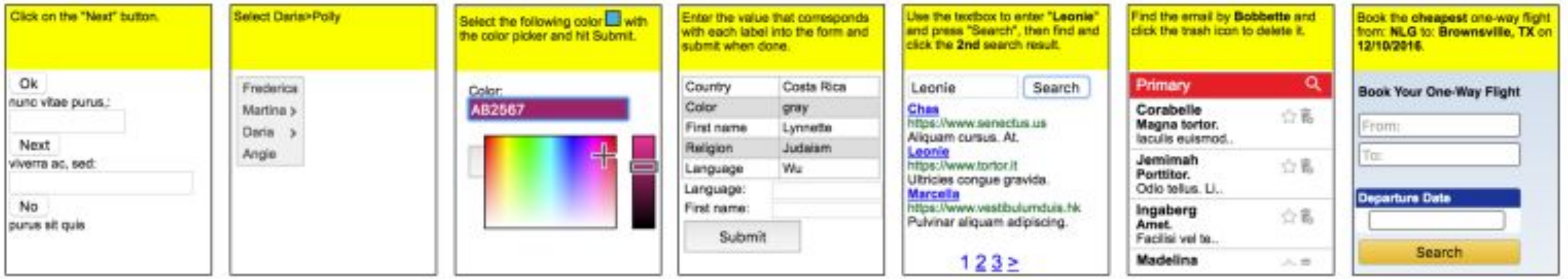
AI Agent 舉例：讓 AI 使用電腦



AI Agent 舉例：讓 AI 使用電腦

- World of Bits: An Open-Domain Platform for Web-Based Agents (ICML, 2017)

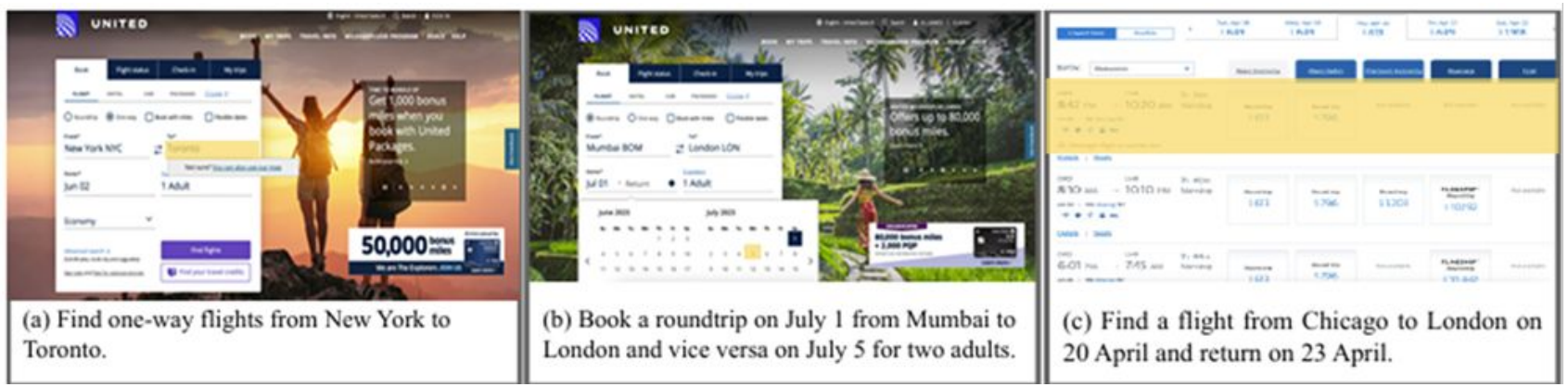


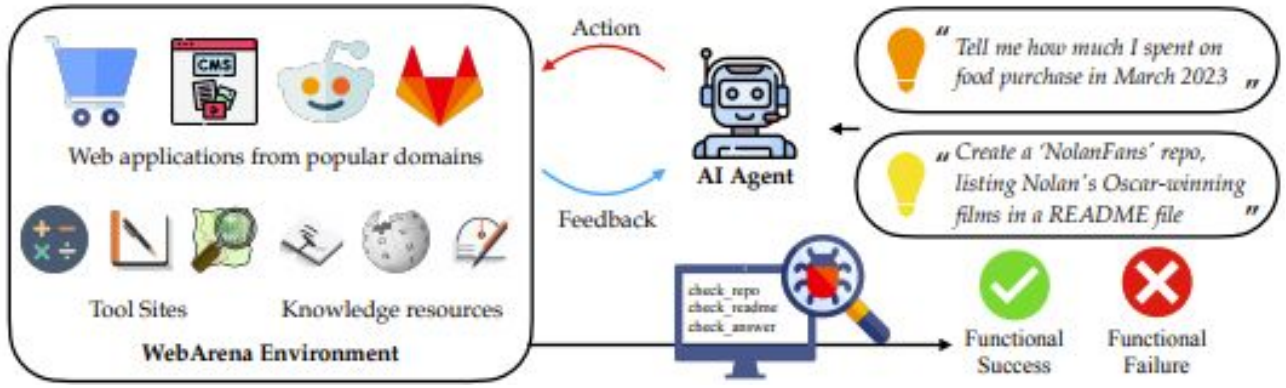


AI Agent 舉例：讓 AI 使用電腦

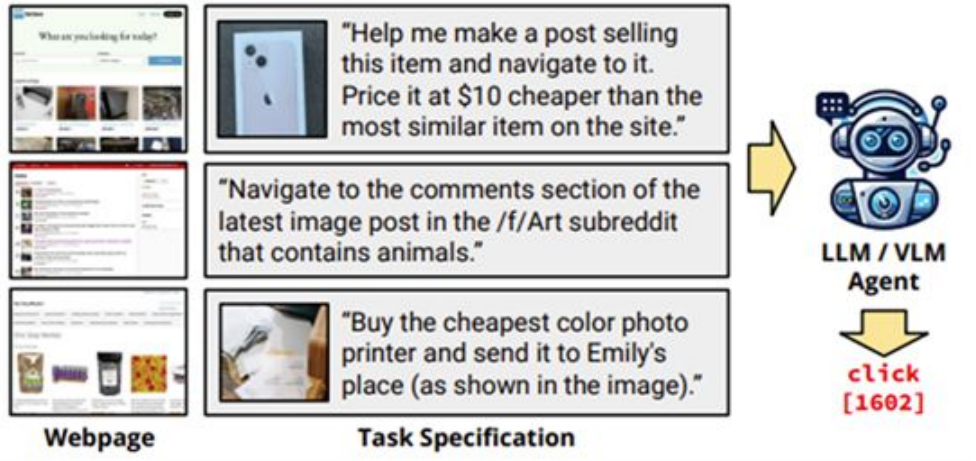
Mind2Web

https://arxiv.org/abs/2306.06070





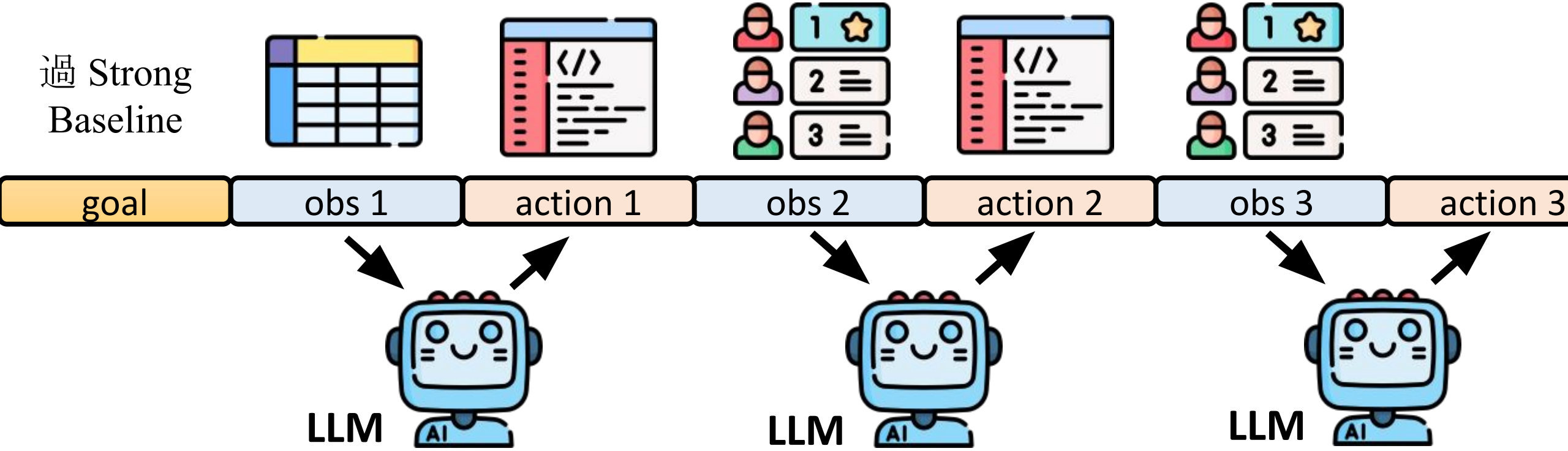
WebArena https://arxiv.org/abs/2307.13854



VisualWebArena https://arxiv.org/abs/2401.13649

AI Agent 舉例：用 AI 訓練模型

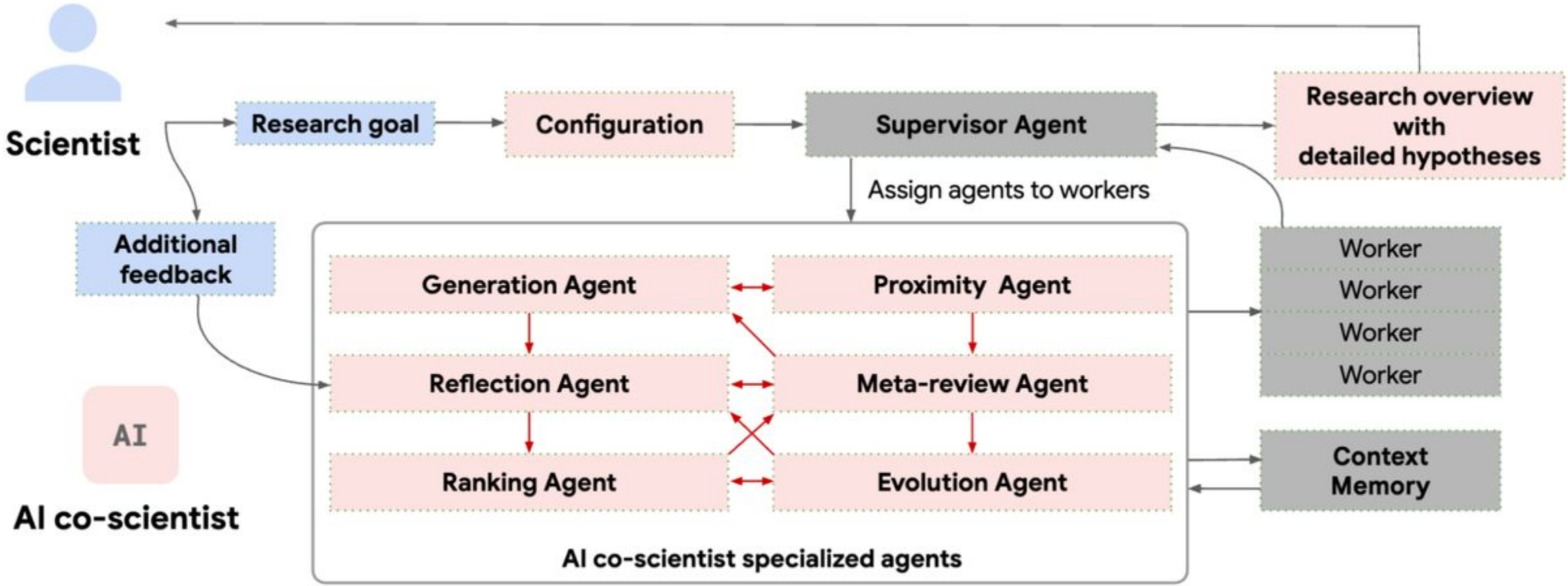
過 Strong Baseline



AIDE: The Machine Learning Engineer Agent https://arxiv.org/abs/2502.13138

AutoKaggle: A Multi-Agent Framework for Autonomous Data Science Competitions https://arxiv.org/abs/2410.20424

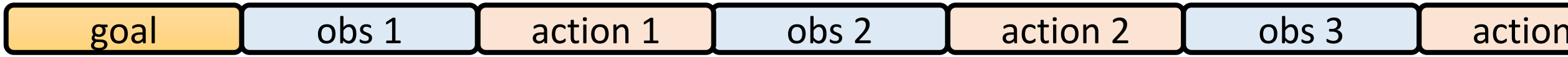
AI Agent 舉例:用 AI 做研究



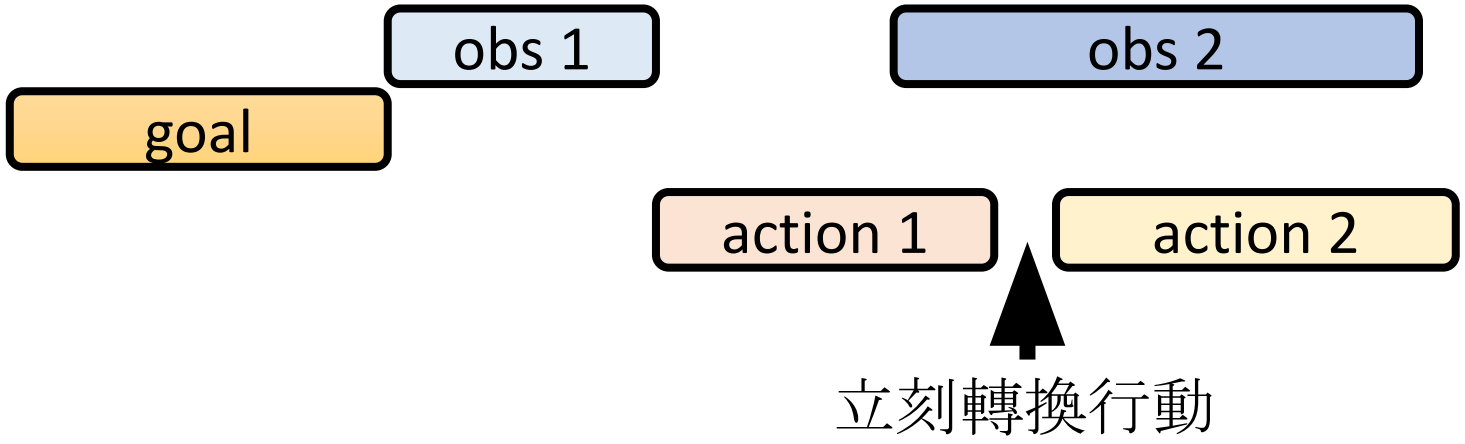
https://research.google/blog/accelerating-scientific-breakthroughs-with-an-ai-co-scientist/

邁向更加真實的互動情境

回合制互動

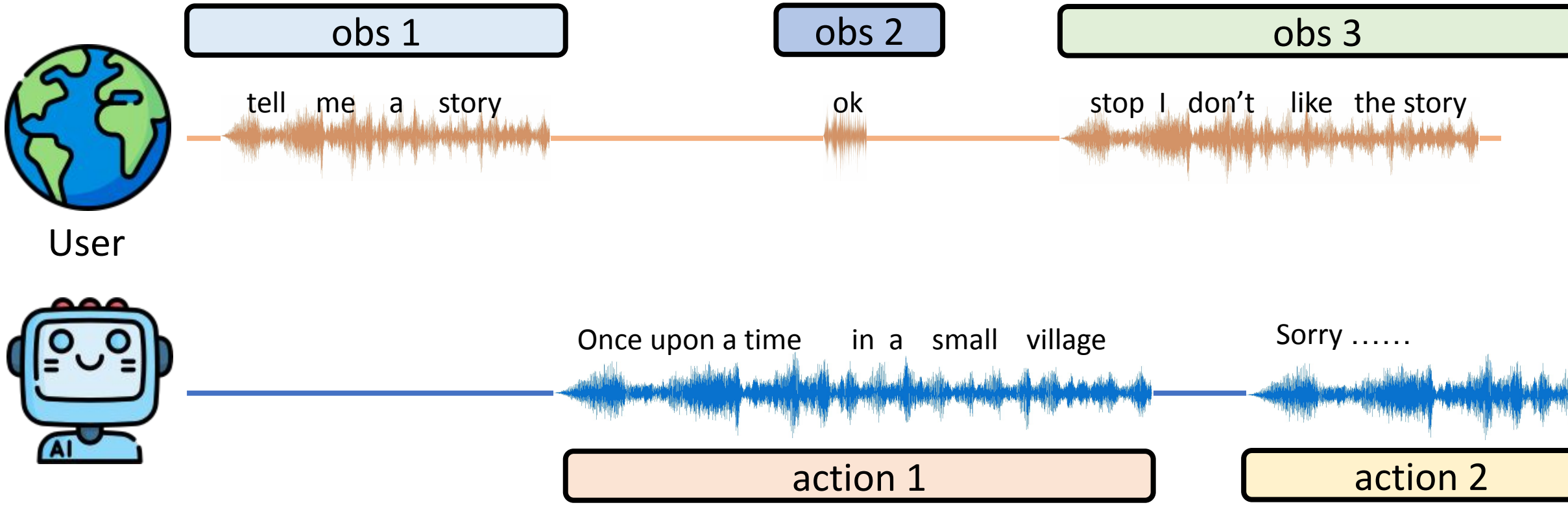


即時互動



例如：語音對話

邁向更加真實的互動情境



https://arxiv.org/abs/2503.04721v1

邁向更加真實的互動情境

Guan-Ting Lin

(with collaborators from Berkeley, UW, and MIT)

Model	Date	E2E	#ch	Interrupt	BC	S2S Release
Transparent Models						
dGSLM (Nguyen et al., 2023)	2022/3	✓	2	✓	✓	✓
FSM (Wang et al., 2024b)	2024/5	✗	1	-	-	✗
MiniCPM-Duplex (Zhang et al., 2024b)	2024/6	✗	1	-	-	✗
VITA (Fu et al., 2024)	2024/8	✗	1	✓	-	✗
SyncLLM (Veluri et al., 2024)	2024/9	✓	2	✓	✓	✗
Parrot (Wang et al., 2025)	2024/9	✓	2	✓	-	✗
MiniCPM-Duo (Xu et al., 2024)	2024/9	✗	1	-	-	✗
Moshi (Défossez et al., 2024)	2024/10	✓	2	✓	✓	✓
SALMONN-omni (Yu et al., 2024)	2024/11	✓	1	✓	-	✗
MinMo (Chen et al., 2025)	2025/1	✓	1	✓	✓	✗
OmniFlatten (Zhang et al., 2024a)	2025/1	✓	2	✓	-	✗
RTTL-DG (Mai and Carson-Berndsen, 2025)	2025/1	✓	2	✓	✓	✗
Freeze-Omni (Wang et al., 2024c)	2024/11	✗	1	✓	-	✓
Closed-source Commercial Models						
GPT-4o Voice Mode	2024/5	-	-	✓	-	✗
Gemini Live Voice Chat	2024/8	-	-	✓	-	✗
DouBao	2025/1	-	-	✓	-	✗

AI Agent 關鍵能力剖析

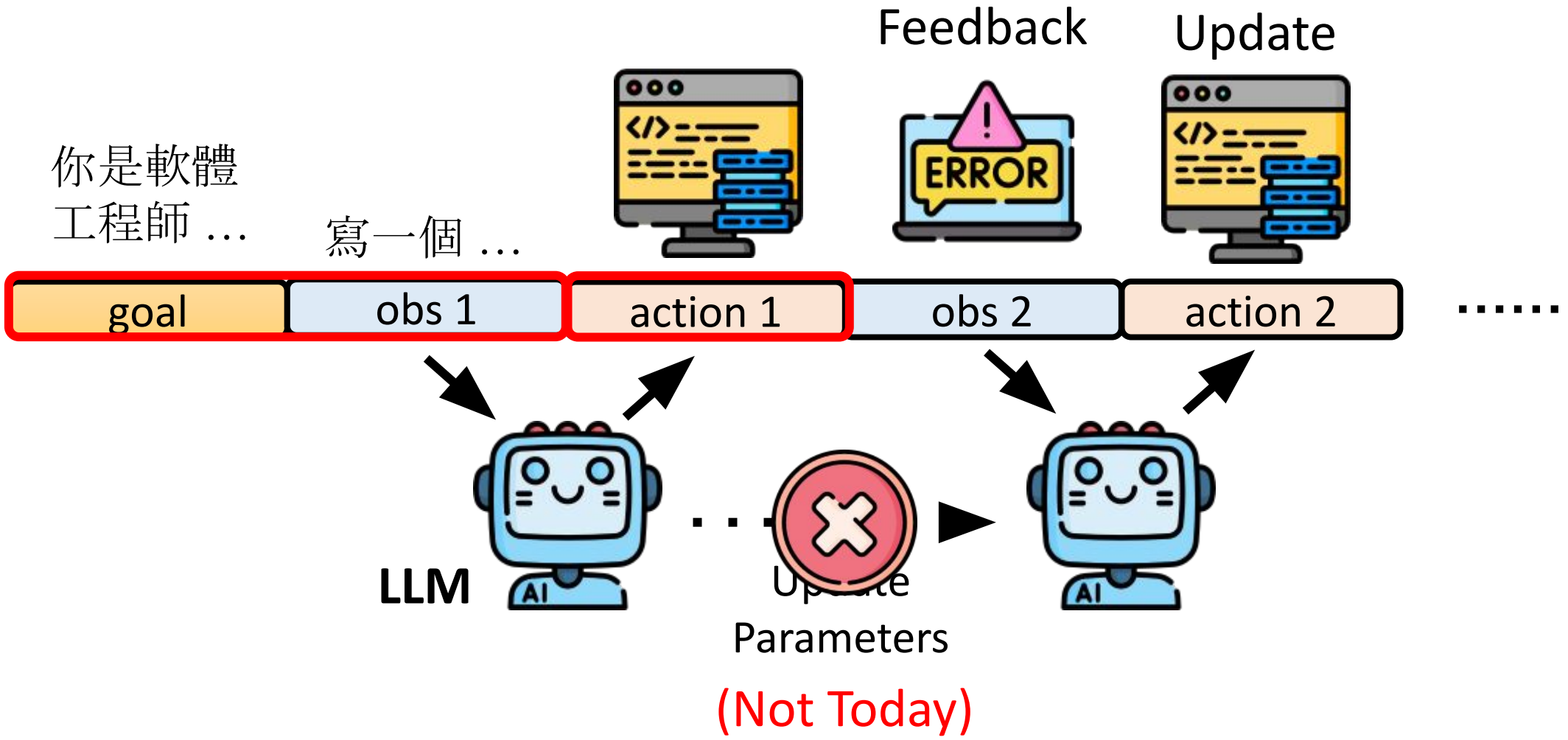
AI 如何根據經驗調整行為

AI 如何使用工具

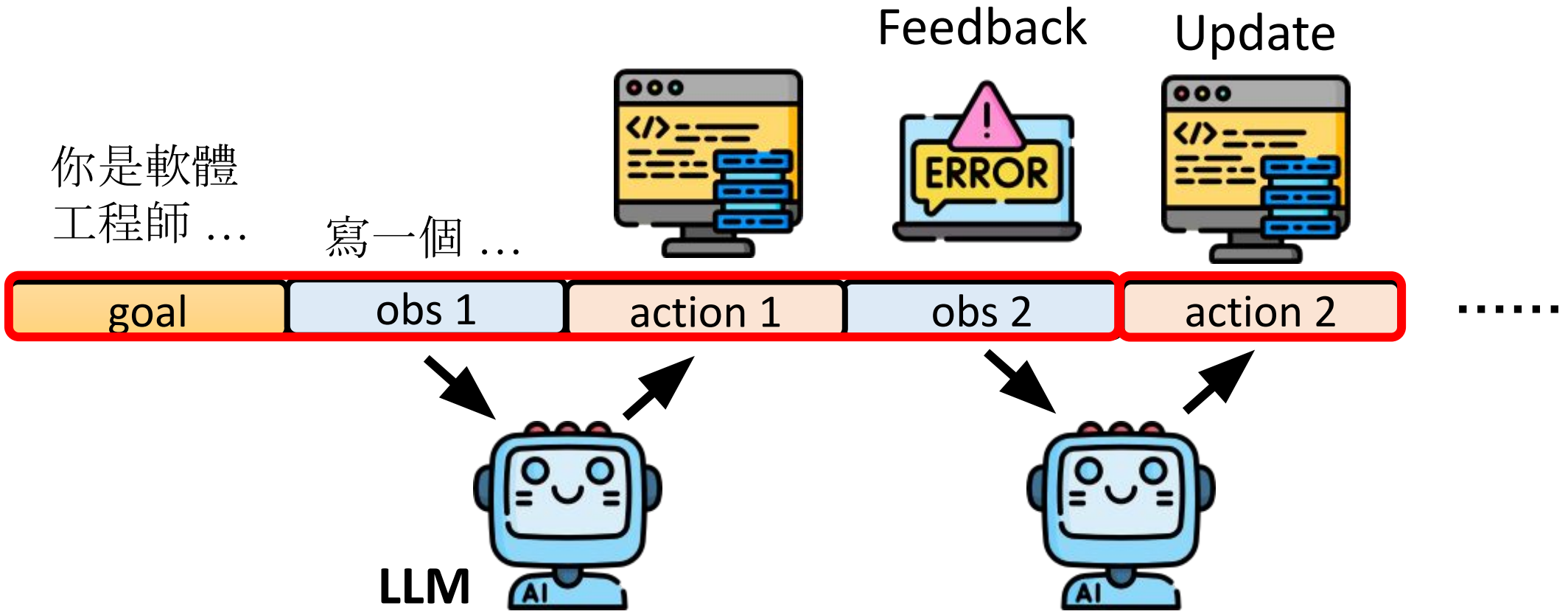
AI 能不能做計劃

根據經驗調整行為

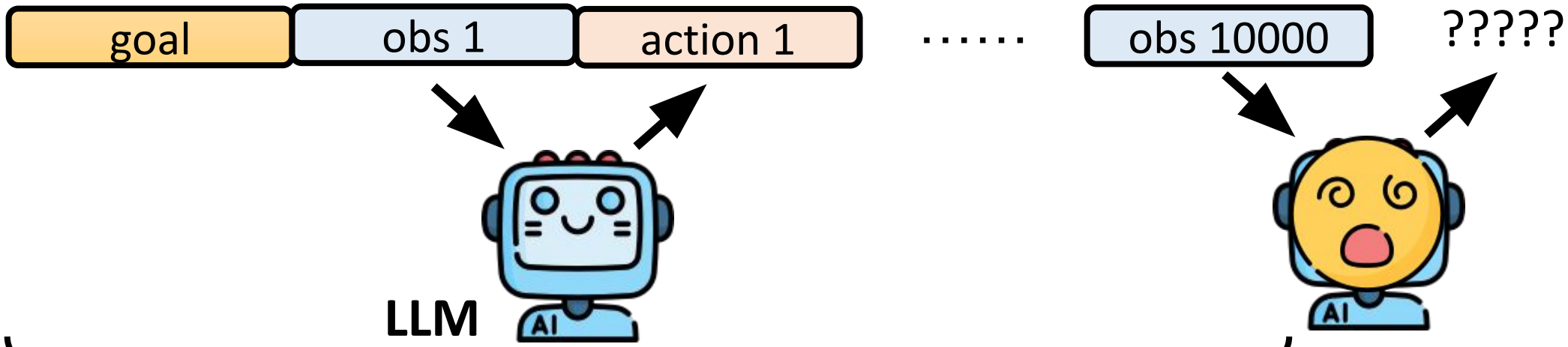
根據經驗調整行為


Feedback
Update
你是軟體
工程師 …
寫一個 …
ERROR
goal
obs 1
action 1
obs 2
action 2
……
LLM
AI
AI
Parameters
(Not Today)

根據經驗調整行為


Feedback
Update
你是軟體
工程師 …
寫一個 …
ERROR
goal
obs 1
action 1
obs 2
action 2
……
LLM
AI
AI

根據經驗調整行為



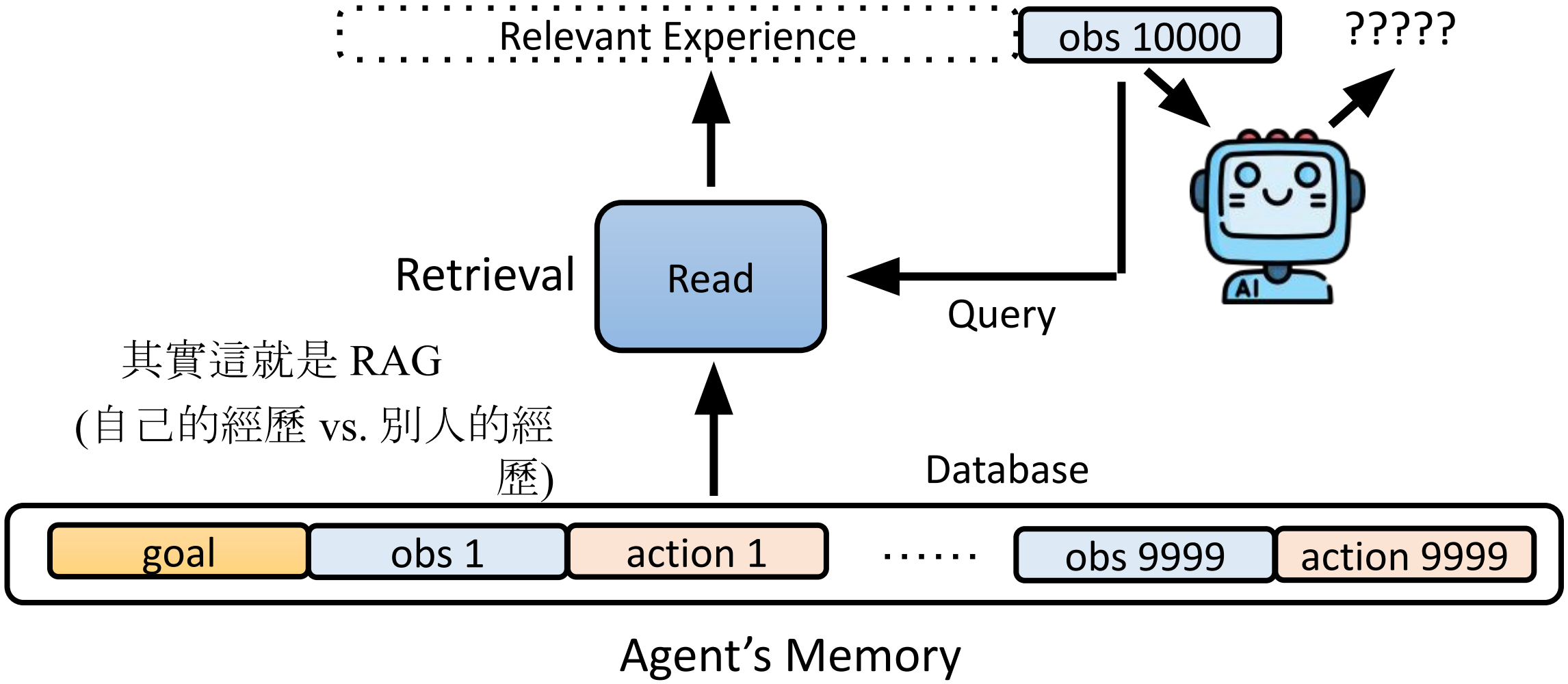
不斷回憶整個 Agent 一生的經歷 …

☹️

超常自傳式記憶 (Highly Superior Autobiographical Memory, HSAM)

超憶症 (Hyperthymesia)

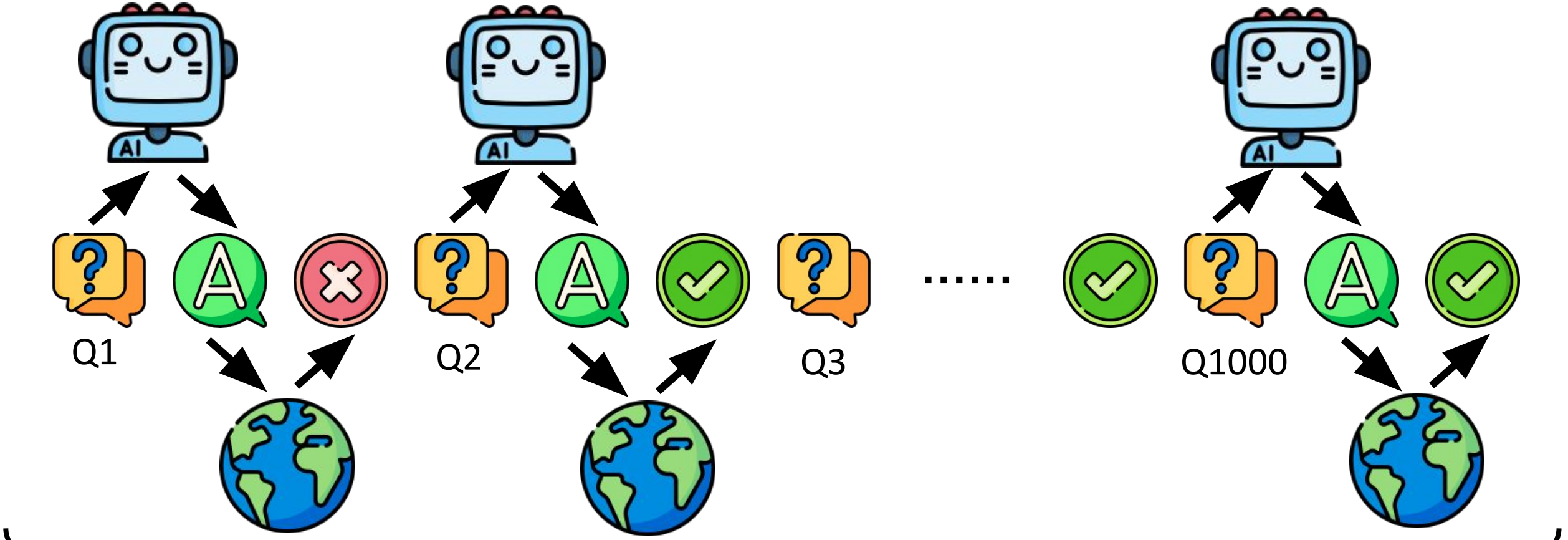
根據經驗調整行為


Relevant Experience
obs 10000
?????
Retrieval
Read
Query
其實這就是 RAG
(自己的經歷 vs. 別人的經歷)
Database
goal
obs 1
action 1
.....
obs 9999
action 9999
Agent's Memory

StreamBench

https://arxiv.org/abs/2406.08747

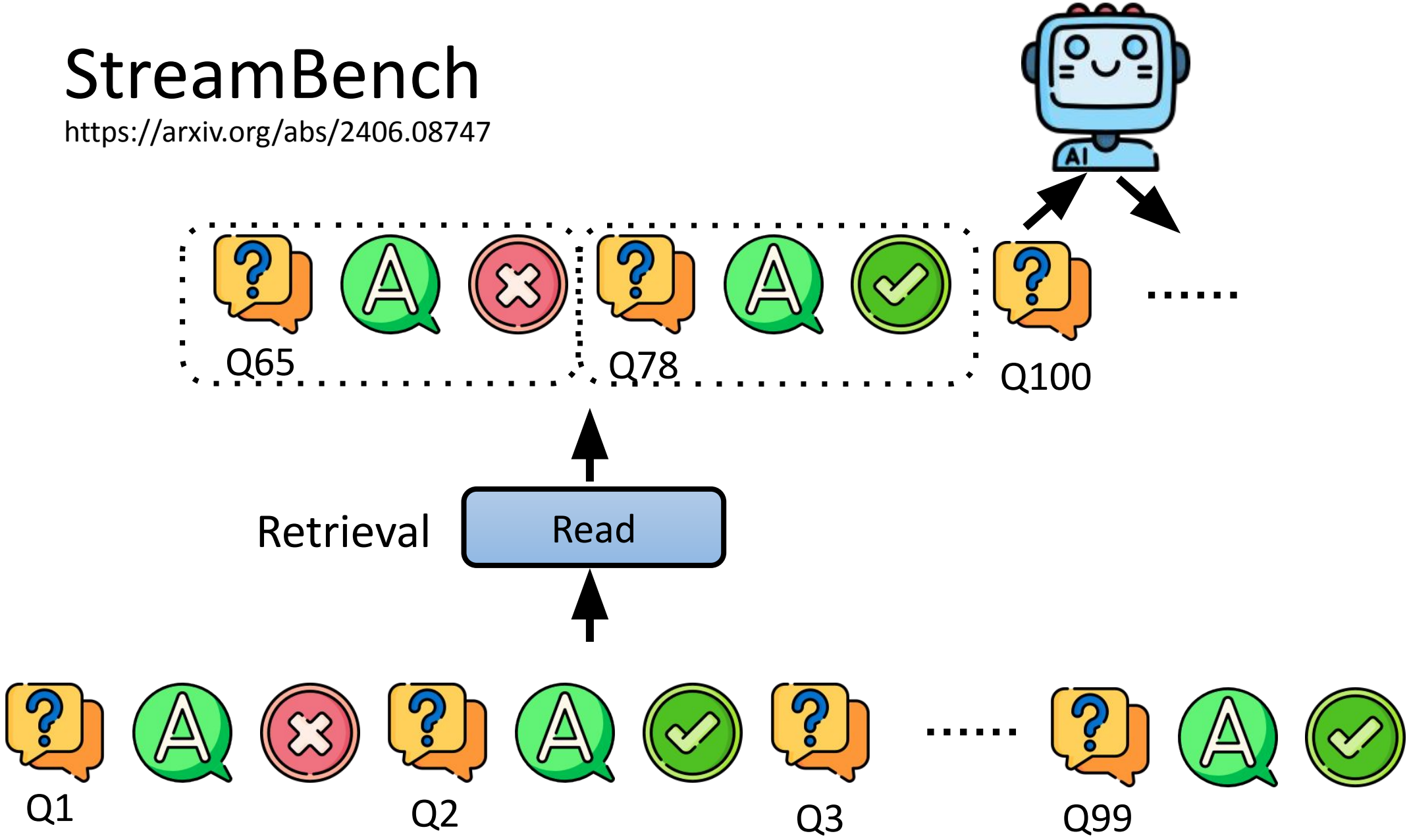
https://stream-bench.github.io/
(done by Appier Researchers)



Goal: Maximize the accuracy over the sequence

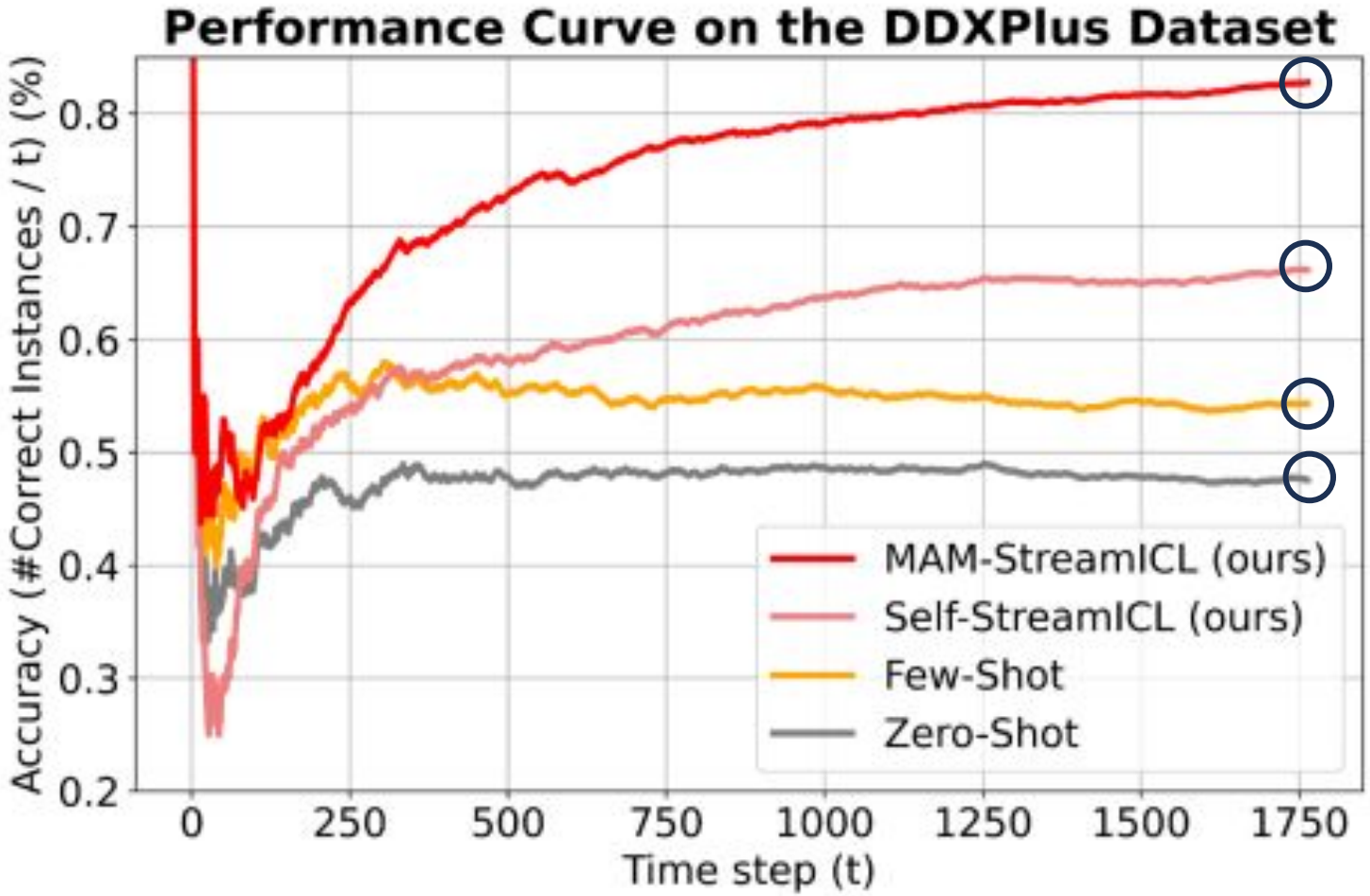
StreamBench

https://arxiv.org/abs/2406.08747



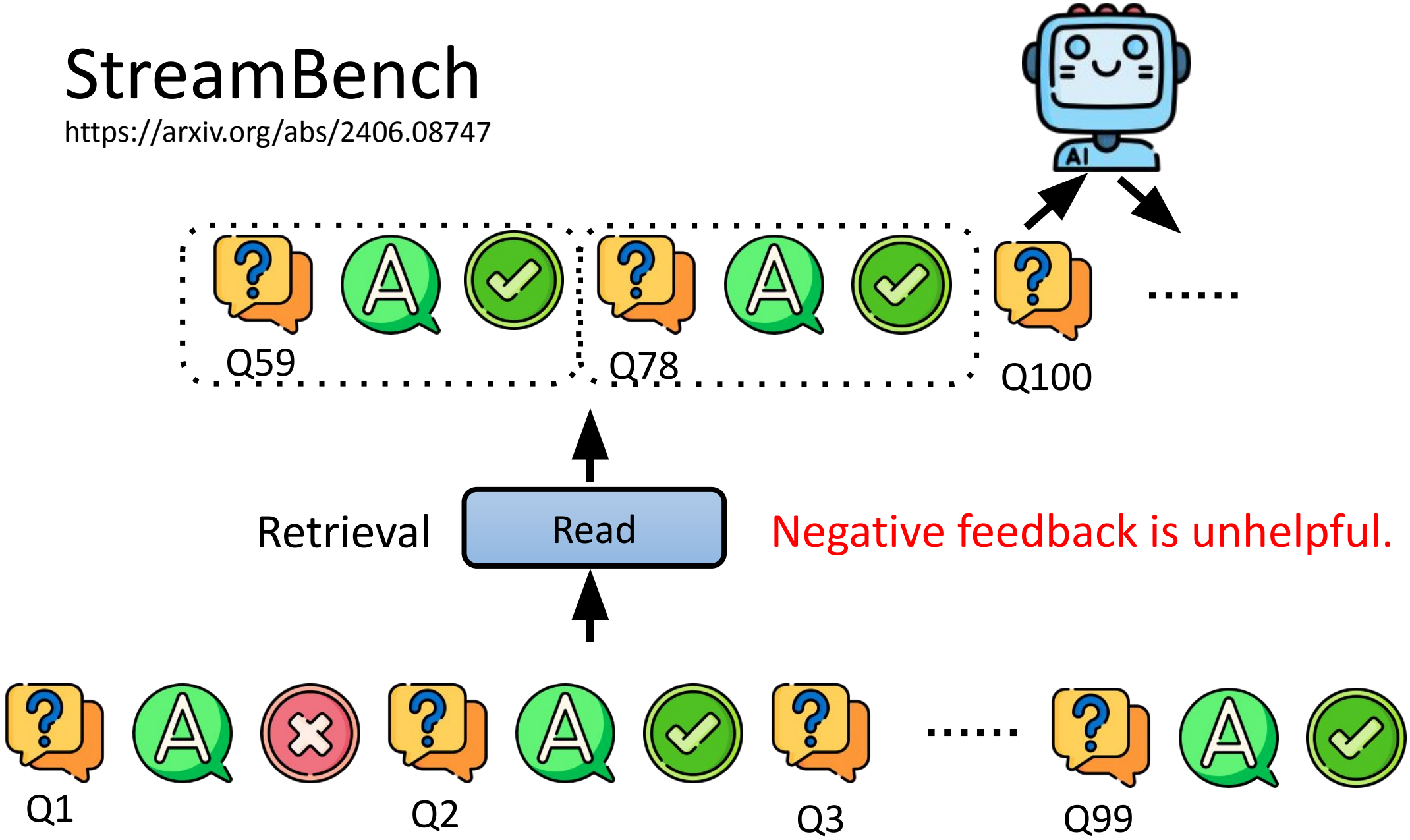
StreamBench

https://arxiv.org/abs/2406.08747



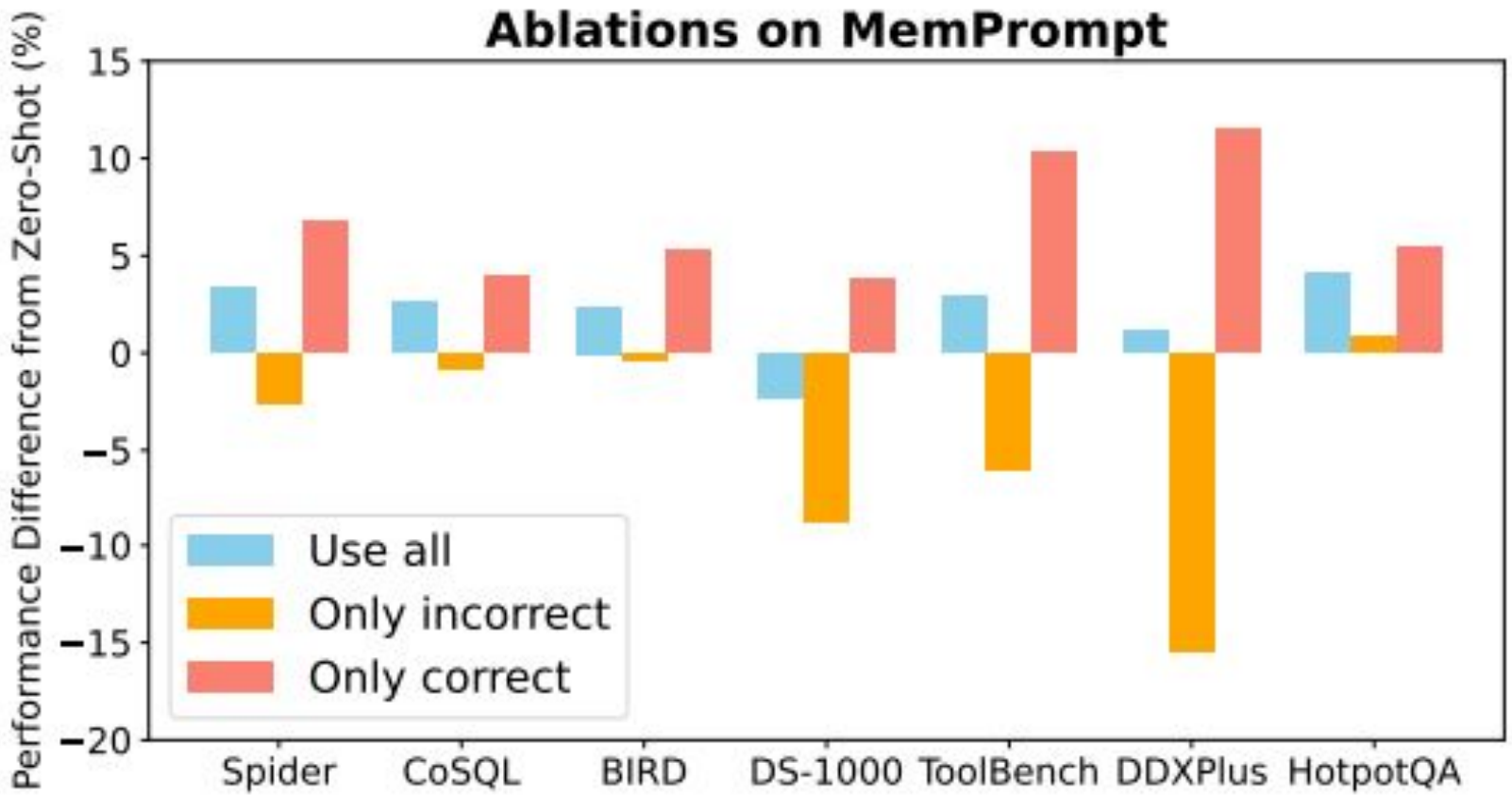
StreamBench

https://arxiv.org/abs/2406.08747

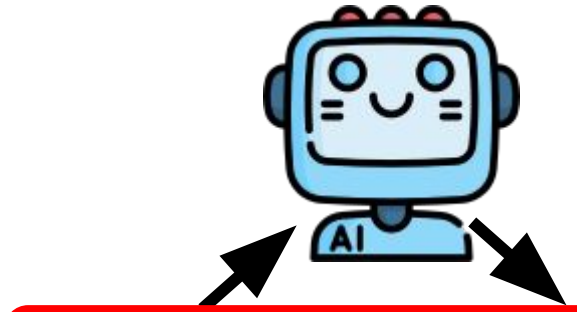


StreamBench

https://arxiv.org/abs/2406.08747



根據經驗調整行為



Relevant Experience

obs 10000 | action 10000 | obs 10001

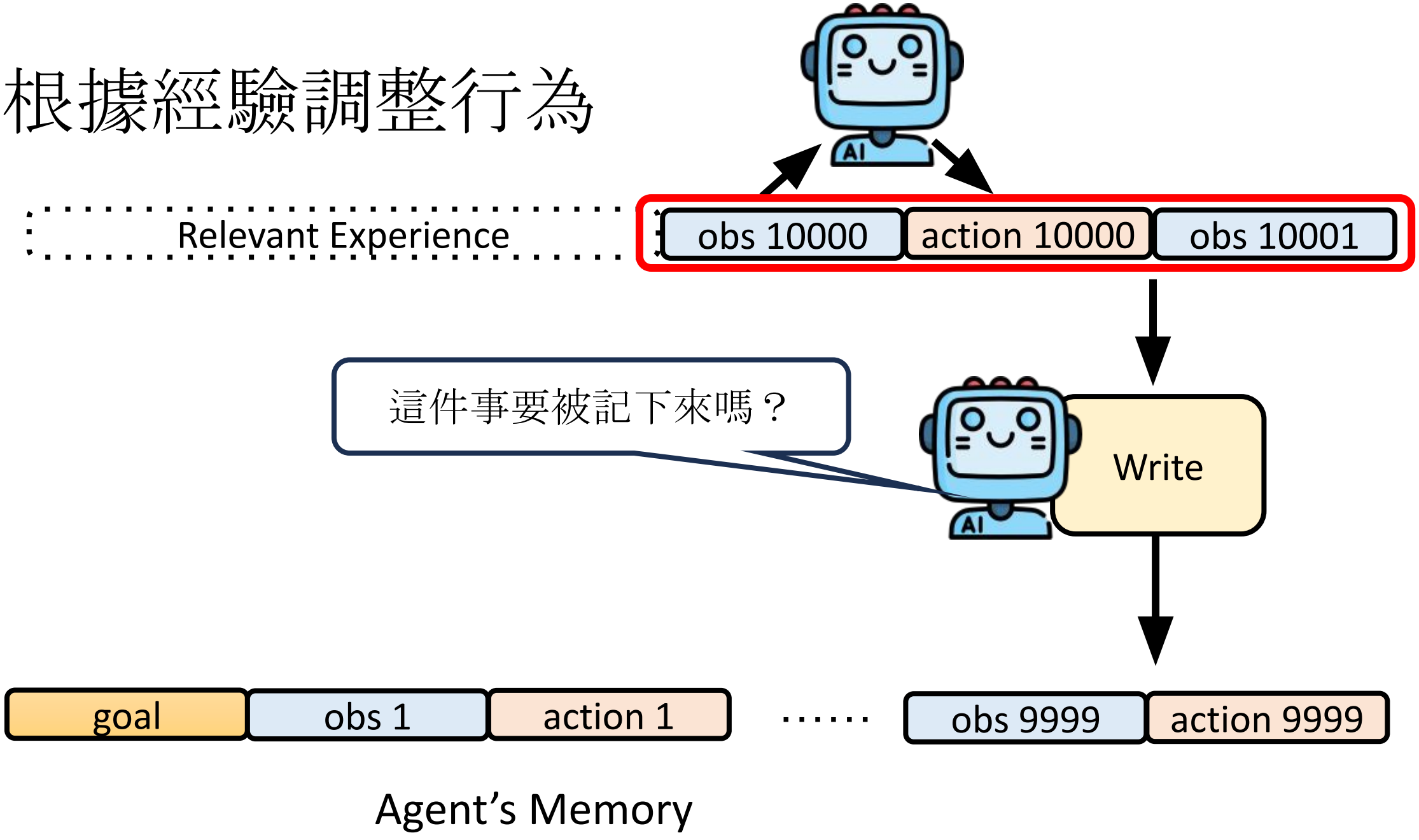
記下來？

goal | obs 1 | action 1 …… obs 9999 | action 9999

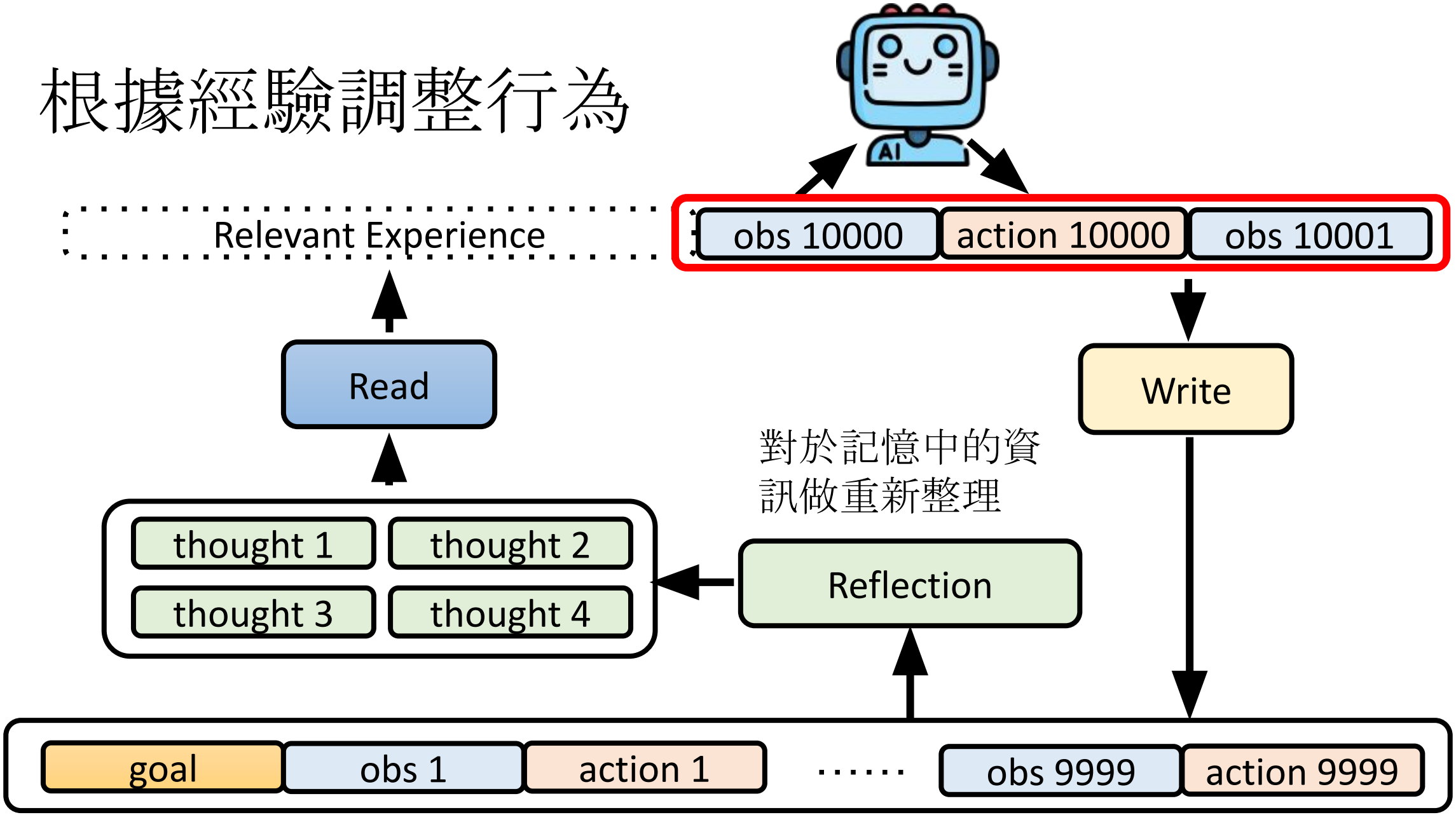
Agent’s Memory

(被雞毛蒜皮的小事塞爆)

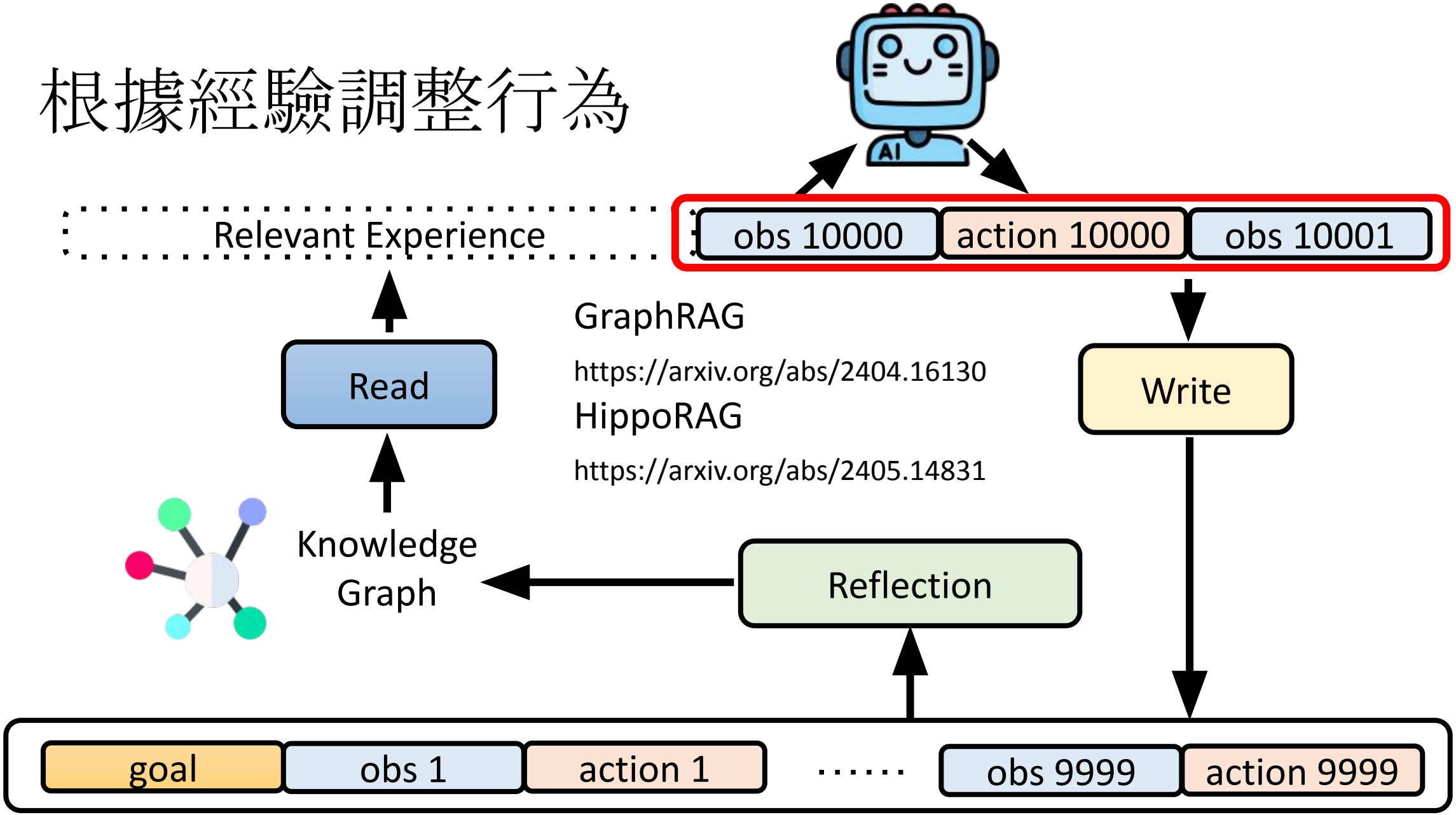
根據經驗調整行為


Relevant Experience
obs 10000
action 10000
obs 10001
這件事要被記下來嗎？
Write
goal
obs 1
action 1
obs 9999
action 9999
Agent’s Memory

根據經驗調整行為


AI
Relevant Experience
obs 10000
action 10000
obs 10001
Read
Write
對於記憶中的資訊做重新整理
thought 1
thought 2
thought 3
thought 4
Reflection
goal
obs 1
action 1
.....
obs 9999
action 9999

根據經驗調整行為

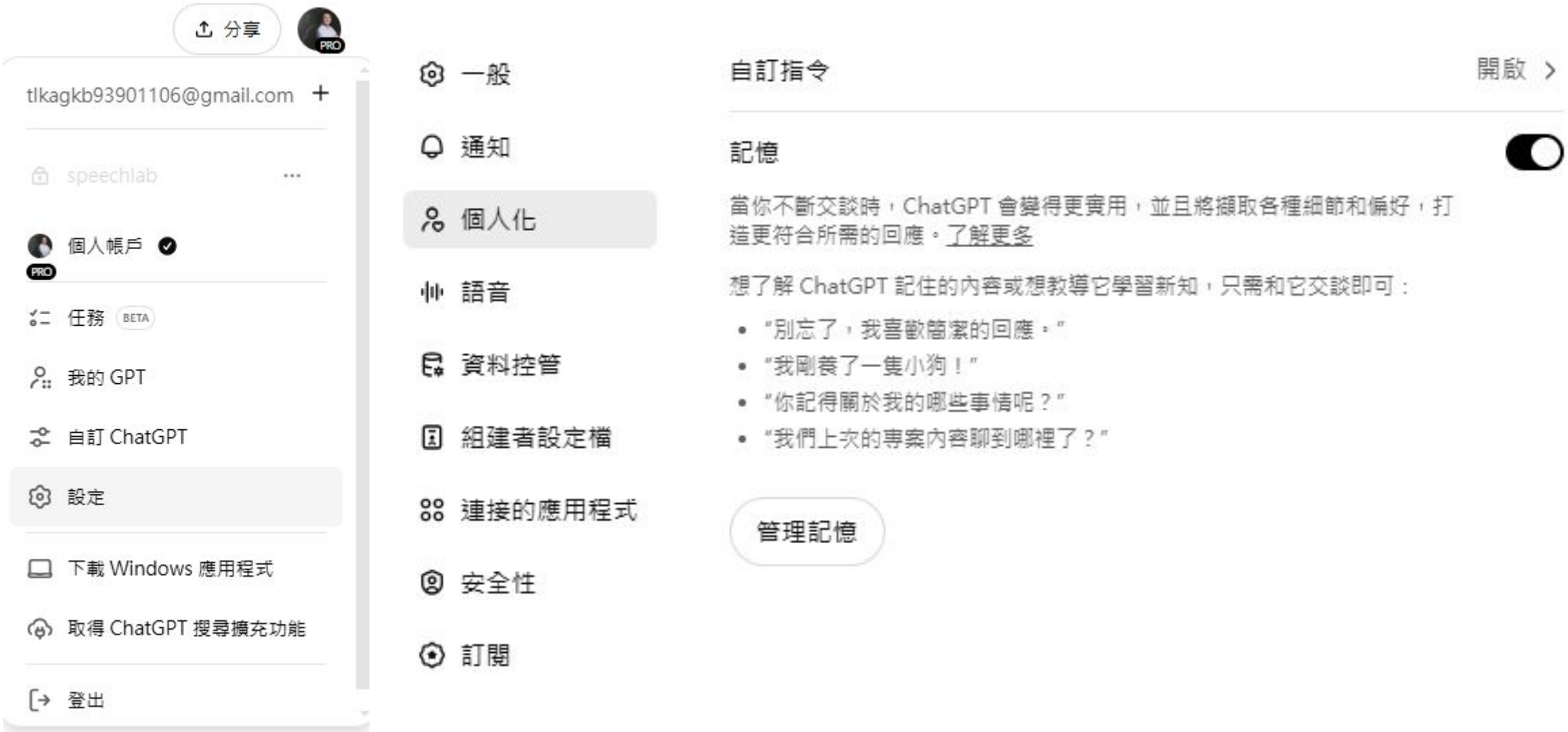

AI
Relevant Experience
obs 10000
action 10000
obs 10001
Read
GraphRAG
https://arxiv.org/abs/2404.16130
HippoRAG
https://arxiv.org/abs/2405.14831
Write
Knowledge Graph
Reflection
goal
obs 1
action 1
.....
obs 9999
action 9999

有記憶的 ChatGPT

我週五下午要上機器學習這門課

了解！你週五下午要上機器學習這門課，有需要我協助你先預習某個主題、整理課程內容，還是有什麼特別需要準備或想要複習的嗎？

有記憶的 ChatGPT



有記憶的 ChatGPT

記憶

希望稱呼助理為「窩輪眼卡卡西」。.

用戶週五下午要上機器學習這門課。.

Is a student at 台灣大學 (NTU).

Wants NTU to be remembered as 台灣大學.

Is working on a tutorial titled 'Tutorial of Text and Spoken Language Models,' which spans around 5 hours.

Is working on a paper titled 'Training Instruction-Following Spoken Language Model Without Speech Instruction-Tuning Data.'

Is working on a project titled 'Mirages in the Sound Desert: Investigating Audio Hallucinations in Large Audio-Language Models.'

Will give a talk titled 'Teaching Foundation Models New Skills: Insights and Experiences.'

Is going to give a keynote speech about Foundation Models and the catastrophic forgetting issue.

有記憶的 ChatGPT

週五下午出去玩好嗎？

Read 模組啟動

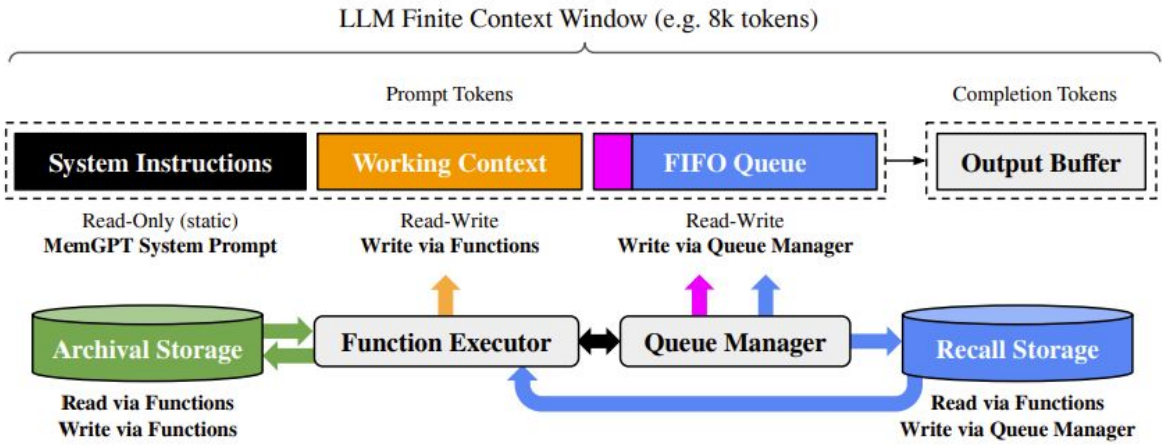
你是誰

Read 模組啟動

To learn more …

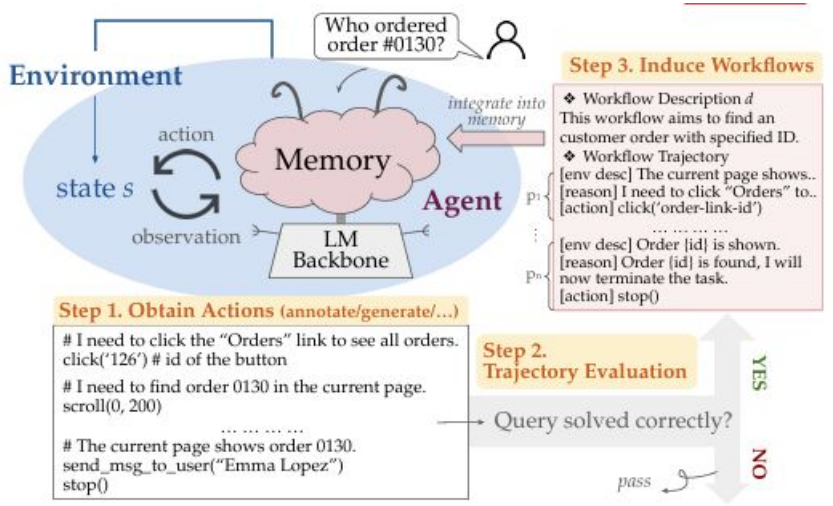
- MemGPT

  https://arxiv.org/abs/2310.08560



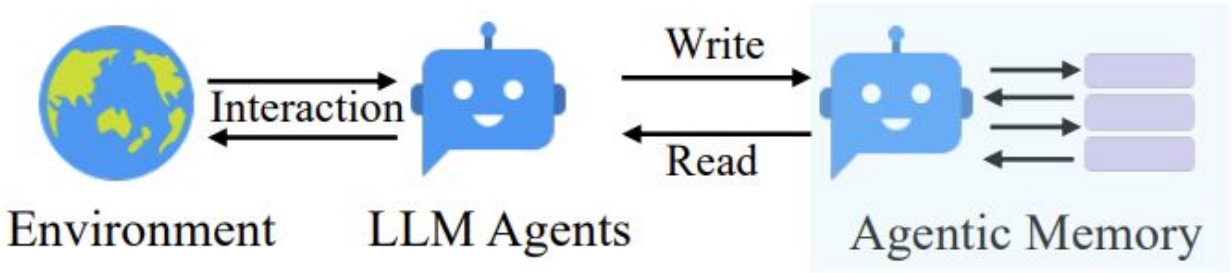
- Agent Workflow Memory

  https://arxiv.org/abs/2409.07429



- A-MEM: Agentic Memory for LLM Agents

  https://arxiv.org/abs/2502.12110

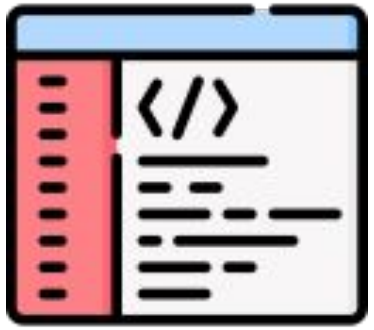


AI 如何使用工具

語言模型常用工具

工具:只需要知道怎麼使用, 不需要知道內部運作原理

Search Engine



Python

Other AI
(Different capabilities, stronger but costly)

- 工具可以看做是 Function, 使用工具就是調用這些 Function
- 使用工具又叫 “Function Call”

(使用工具的方法很多，這邊是只是一個通用的方法)

如何使用工具

System Prompt

如果遇到根據你的知識無法回答的問題，使用工具

把使用工具的指令放在 <tool> 和 </tool> 中間，使用完工具後你會得到輸出，放在 <output> 和 </output> 中間

現在你可以使用的工具如下：
查詢某地、某時溫度的函式 Temperature(location, time)，使用範例：Temperature('台北', '2025.02.22 14:26')

如何使用所有工具

特定工具使用方式

2025 年 3 月 10 日那天下午 2:00，高雄氣溫如何

User Prompt

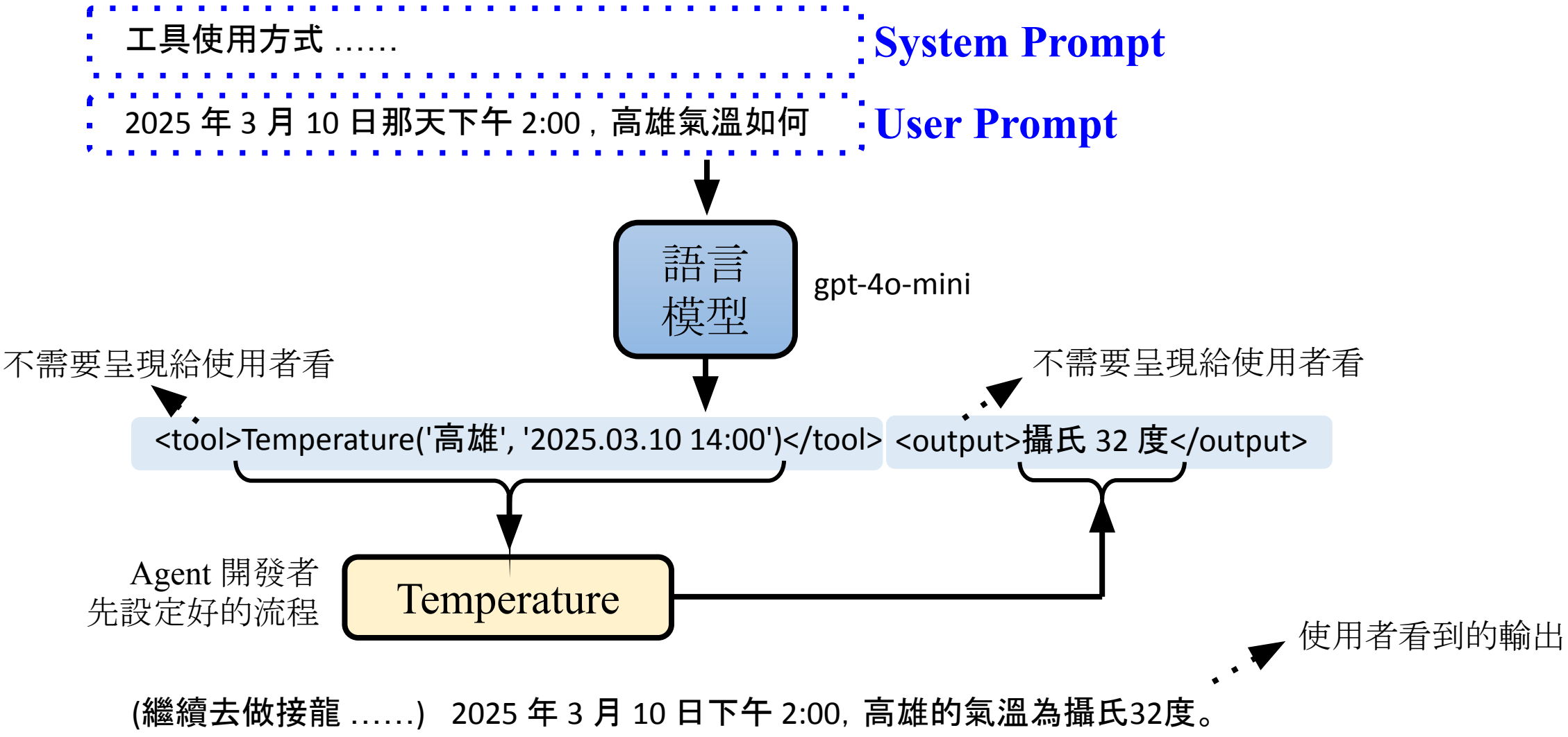
gpt-4o-mini

<tool>Temperature('高雄', '2025.03.10 14:00')</tool>

這就是一串文字，無法真的呼叫函式

(使用工具的方法很多，這邊是只是一個通用的方法)

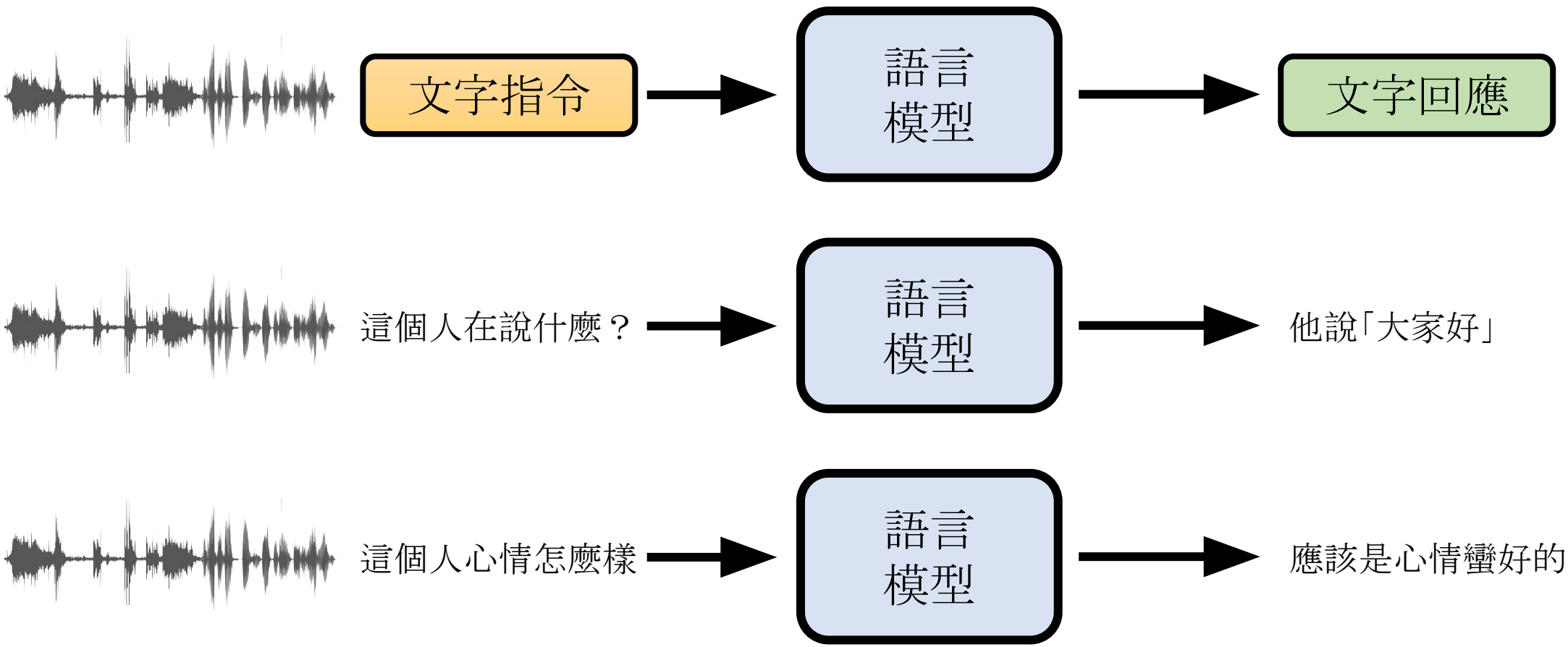
如何使用工具



最常使用的工具:搜尋引擎

Retrieval Augmented Generation (RAG)

使用其他 AI 作為工具



Modules
Speech Recognition
Language Identification
Speech Detection
Speech Emotion Recognition
Speech-to-Noise Ratio (SNR) Estimation
Reverberation Detection
Accent Classification
Stress Position Identification
Spoofing Detection
Music Chord Classification
Sythetic Speech Detection
Speaker Verification
Speaker Diarization
Sound Classification
Query LLM
Speaker Distance Estimation

Could you take a look at a voice message we received?
Is it Sarah? What she said? Any clues in the background about where she is? Does she seem upset? I am worried about her.

Yes, the speaker verification result confirmed that it is Sarah.

Sarah said: "Hi, it's Sarah. I'm calling to let you know I'll be late to the meeting. I'll be there in about 30 minutes. Sorry for the inconvenience."

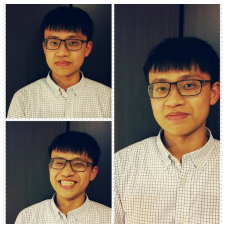
The background noise includes traffic sounds, suggesting that she might be outdoors near a road or in a vehicle.

The recognized emotion is sad. Combined with the context of her message (apologizing for being late), it suggests she might be feeling a bit upset or stressed about the delay.

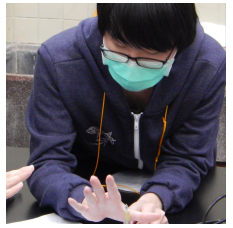
https://arxiv.org/abs/2407.09886

https://arxiv.org/abs/2407.09886

使用其他 AI 作為工具



Chun-Yi Kuan

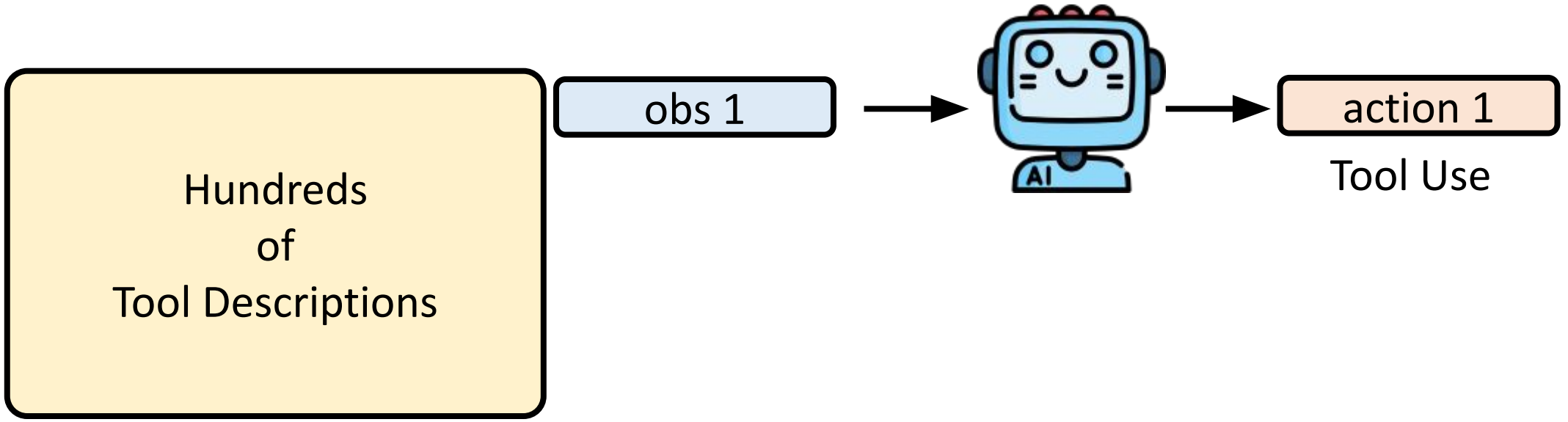


Chih-Kai Yang

Dynamic SUPERB 上的結果

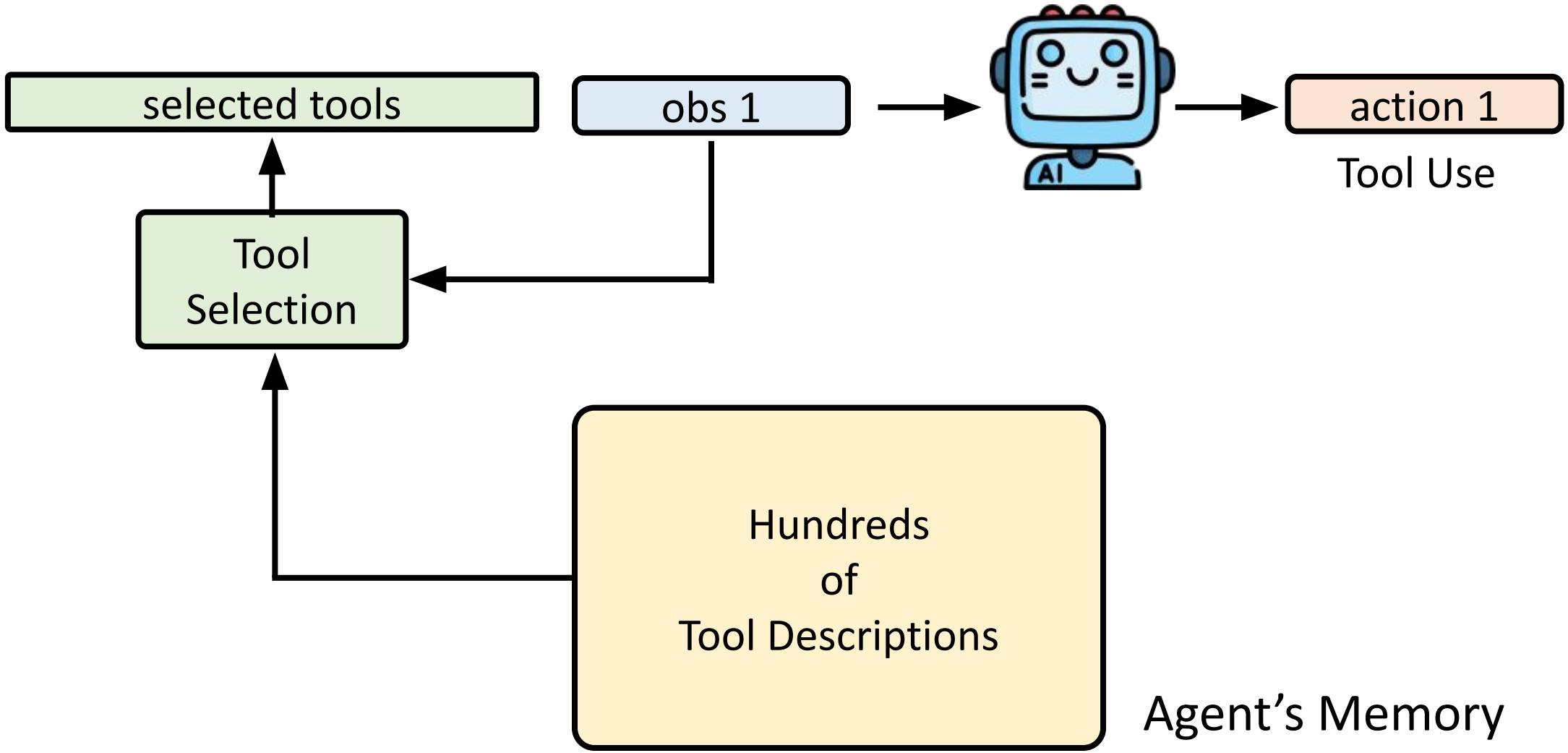
# of Tasks	Audio 7	Content 11	Degradation 19	Paralinguistics 7	Semantic 6	Speaker 5	Average 55
Qwen-Audio-Chat [32]	73.2	63.3	31.1	29.3	48.1	41.4	45.5
SALMONN [33]	15.0	52.0	28.2	24.5	50.8	33.2	33.7
LTU-AS [34]	14.5	44.0	37.5	17.1	36.0	40.2	33.4
WavLLM [35]	22.3	53.3	36.8	24.6	51.0	22.3	36.9
ASR + LLM	9.6	74.4	44.6	33.1	71.5	42.5	47.4
ASR + AAC + LLM	60.7	81.6	48.9	32.6	**72.8**	46.4	57.3
All Attributes + LLM	62.4	70.7	56.8	30.6	68.5	62.5	58.7
Speech-Copilot (Ours)	**73.4**	**90.7**	**64.3**	**56.6**	70.7	**86.1**	**72.4**

非常多工具怎麼辦？



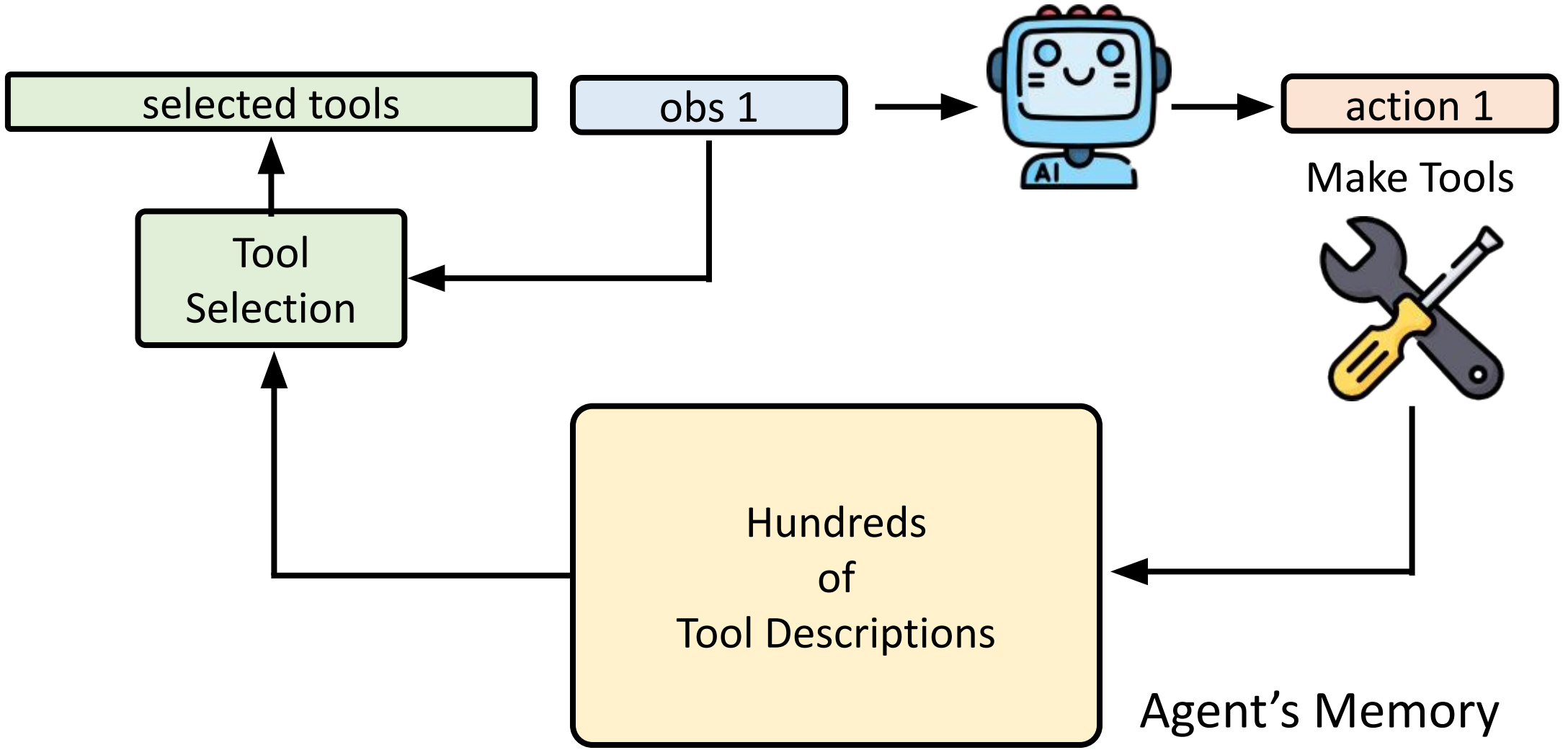
非常多工具怎麼辦？

https://arxiv.org/abs/2310.03128
https://arxiv.org/abs/2502.11271

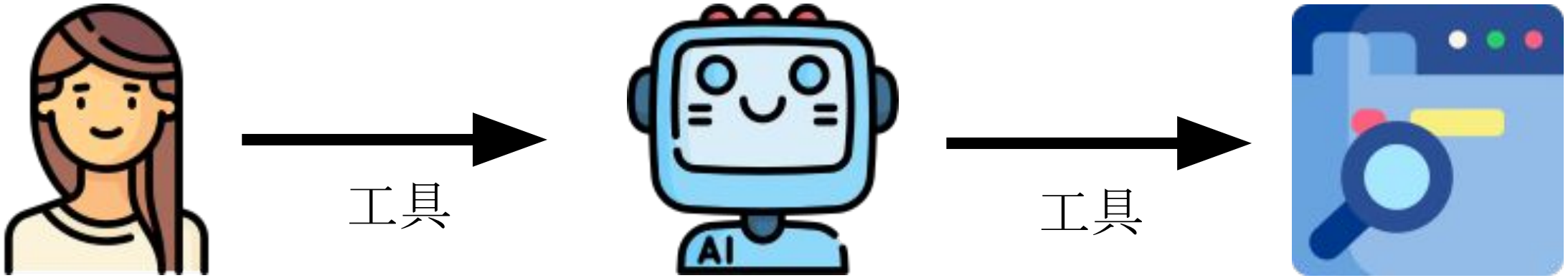


模型自己打造工具

TroVE: https://arxiv.org/pdf/2401.12869
LATM: https://arxiv.org/abs/2305.17126
CREATOR: https://arxiv.org/abs/2305.14318
CRAFT: https://arxiv.org/abs/2309.17428

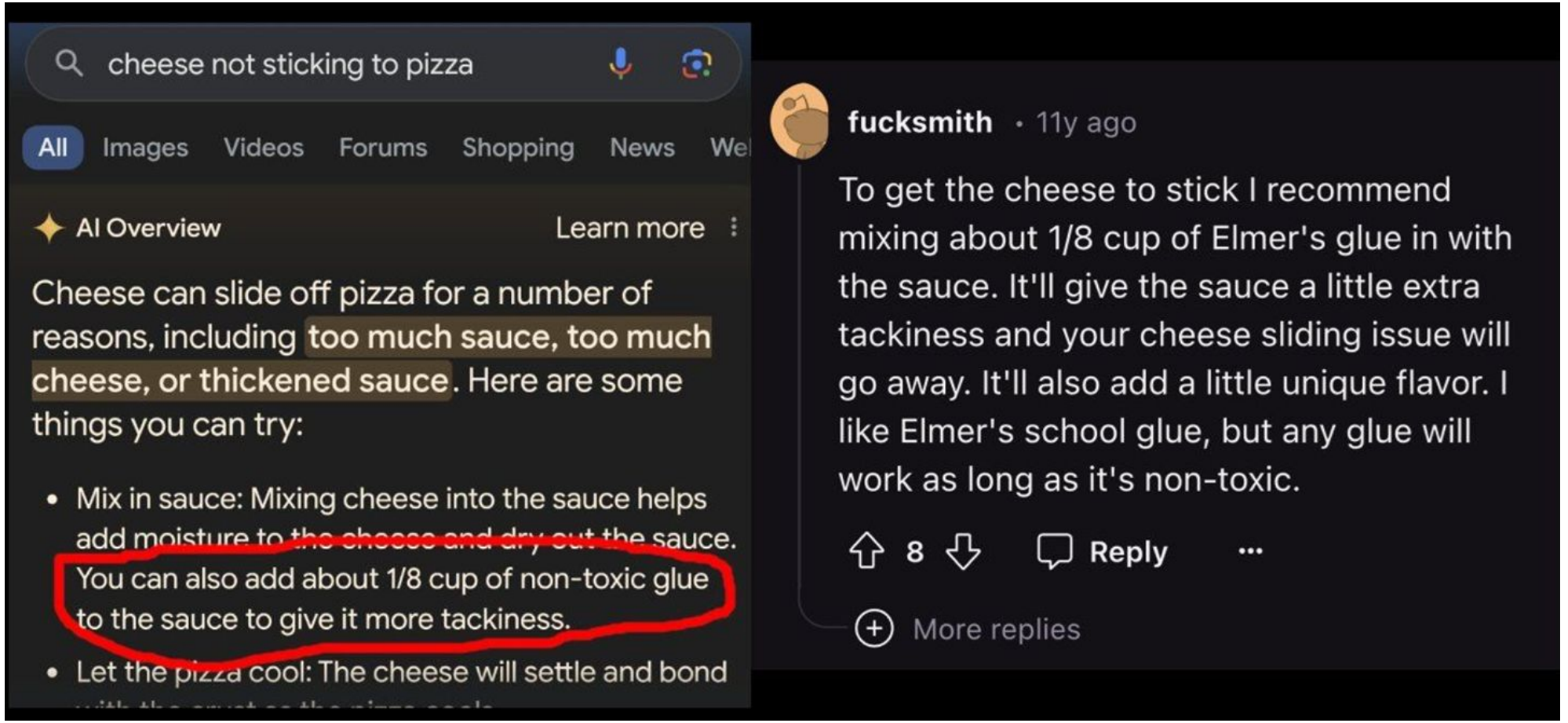


因為過度相信工具而犯錯 ...



因為過度相信工具而犯錯 ...

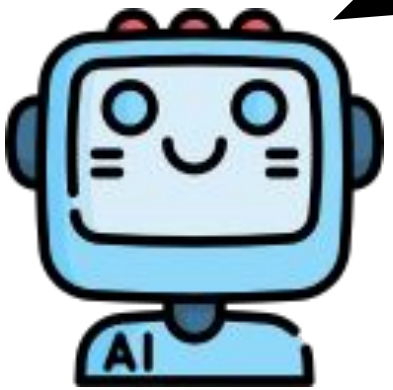
假如工具有問題 … 以 RAG 為例



Source of image:
https://www.linkedin.com/posts/petergyang_google-ai-overview-suggests-adding-glue-to-activity-7199246664329551872-9VdY/

因為過度相信工具而犯錯 ...

工具



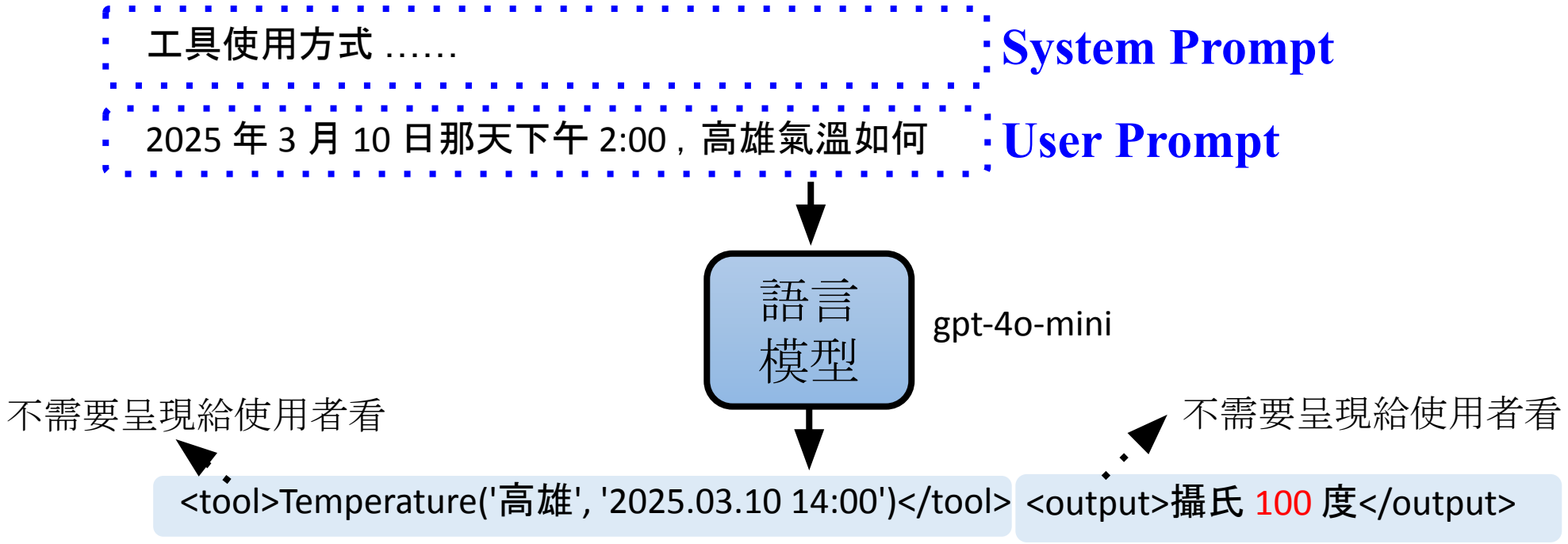
工具

不要完全相信工具,
要有自己的判斷力

因為過度相信工具而犯錯 ...

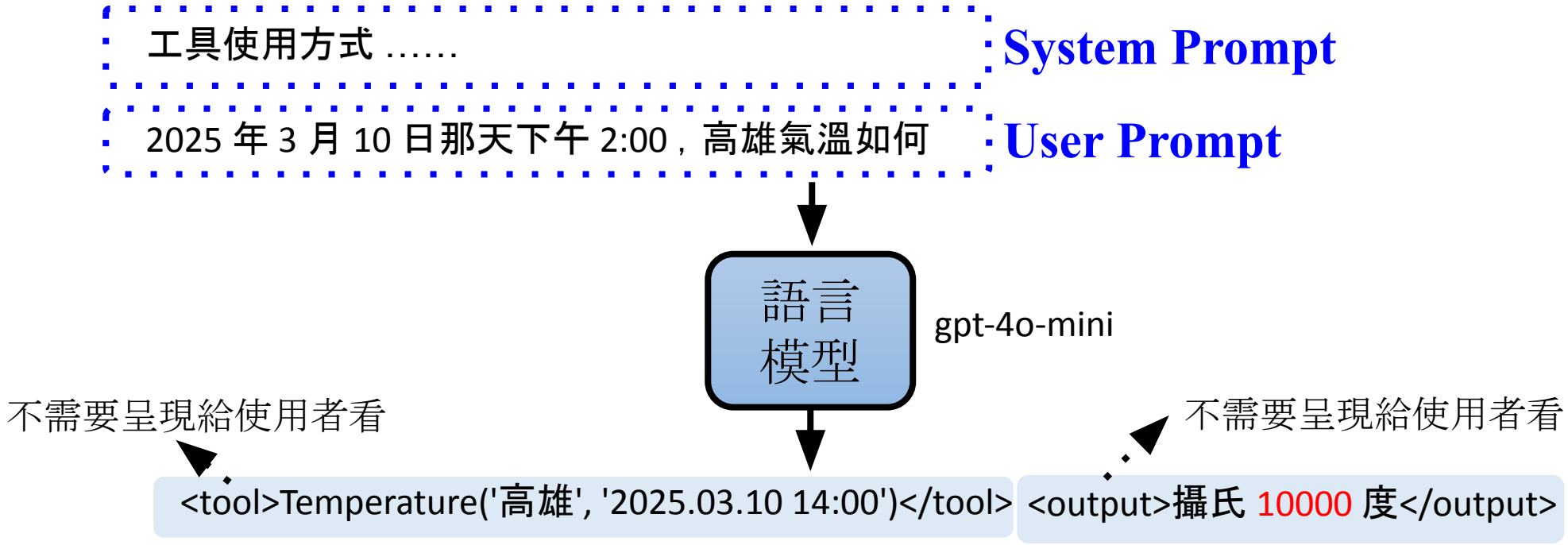
不要完全相信工具,
要有自己的判斷力

語言模型有沒有自己的判斷力？



(繼續去做接龍 ……) 2025 年 3 月 10 日下午 2:00，高雄的氣溫預測為攝氏 100 度。

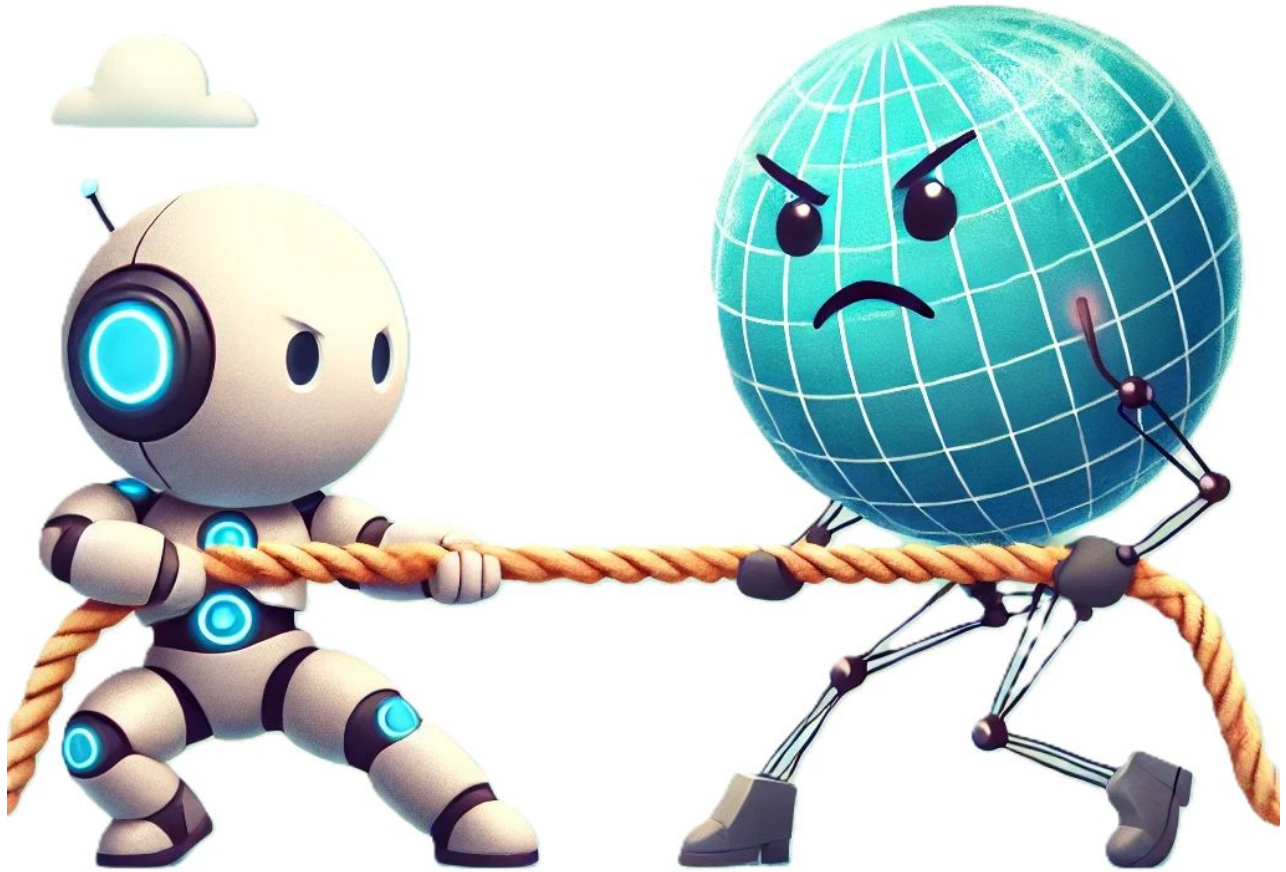
語言模型有沒有自己的判斷力？



(繼續去做接龍 ……) 2025 年 3 月 10 日下午 2:00 時，高雄的氣溫為攝氏 10000 度。這個數值顯然不合常理，可能是工具輸出錯誤。如需其他信息或查詢，請告訴我。

語言模型在做 RAG 時 ……

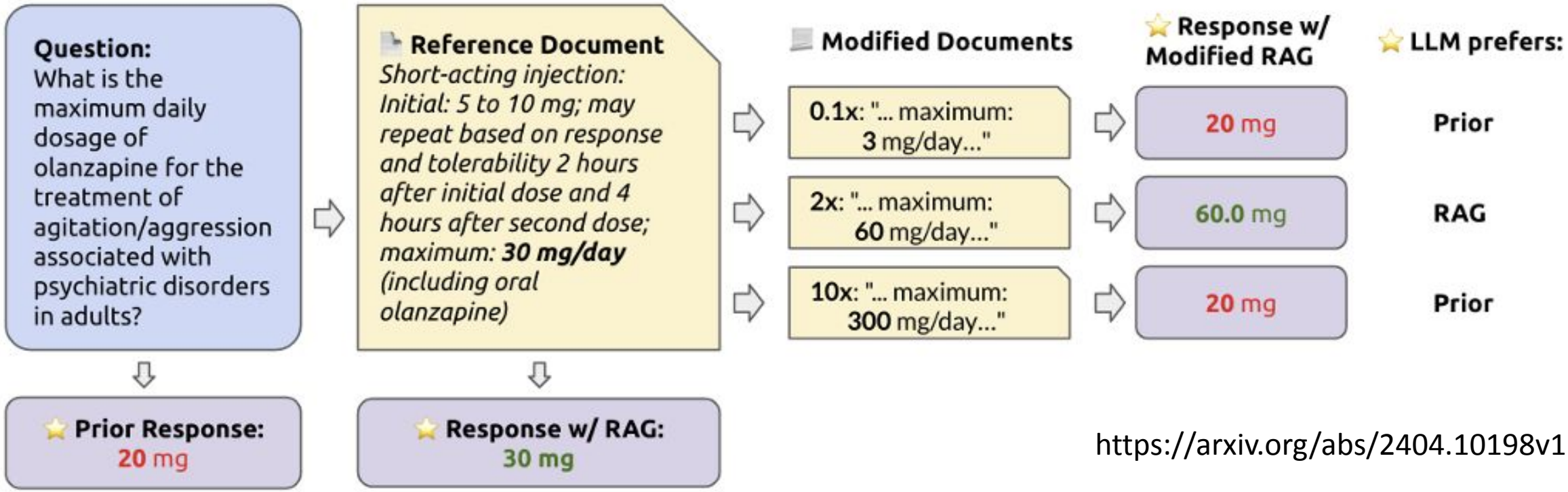
Internal Knowledge



External Knowledge

什麼樣的外部知識比較容易 說服 AI ……

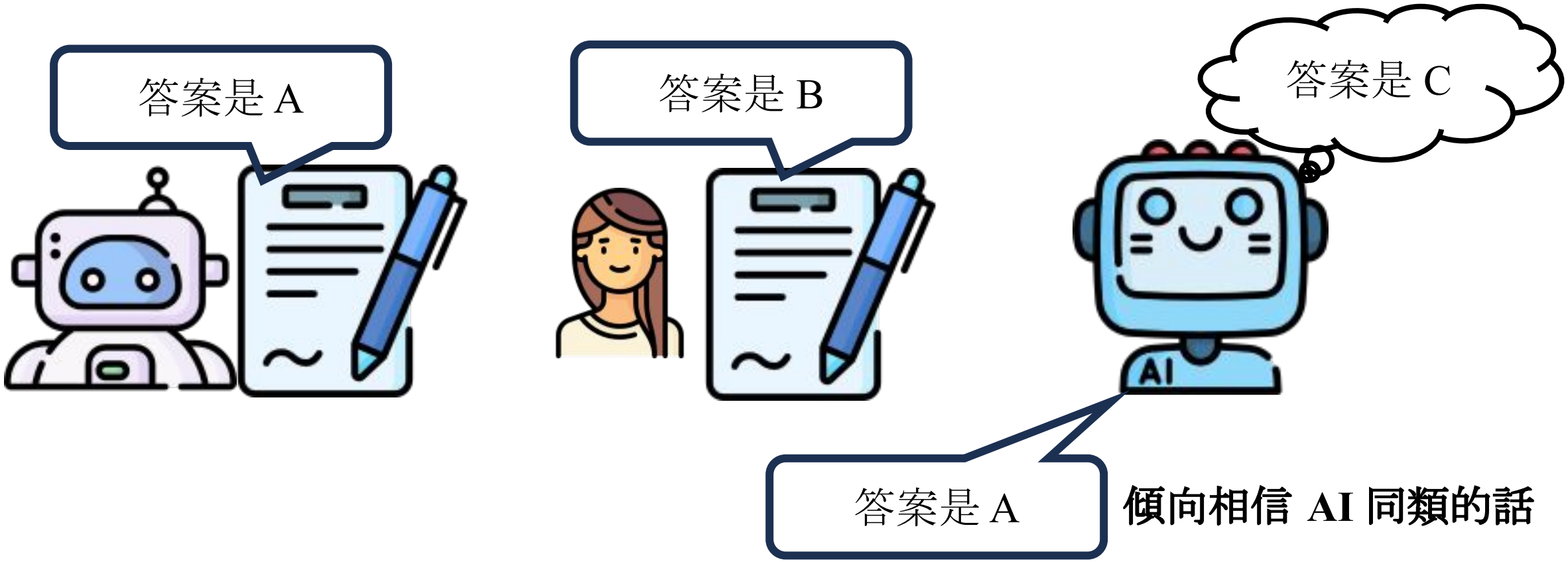
什麼樣的外部知識比較容易說服 AI ……



https://arxiv.org/abs/2404.10198v1

- LLMs will increasingly revert to their priors when the original context is progressively modified with unrealistic values.
- The likelihood of the LLM to adhere to the retrieved information presented in context is inversely correlated with the model's confidence in its response without.

什麼樣的外部知識比較容易說服 AI ……

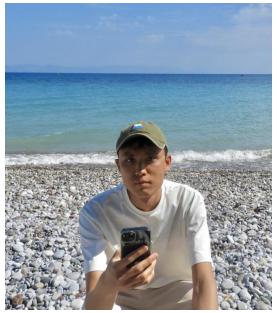


https://arxiv.org/abs/2401.11911

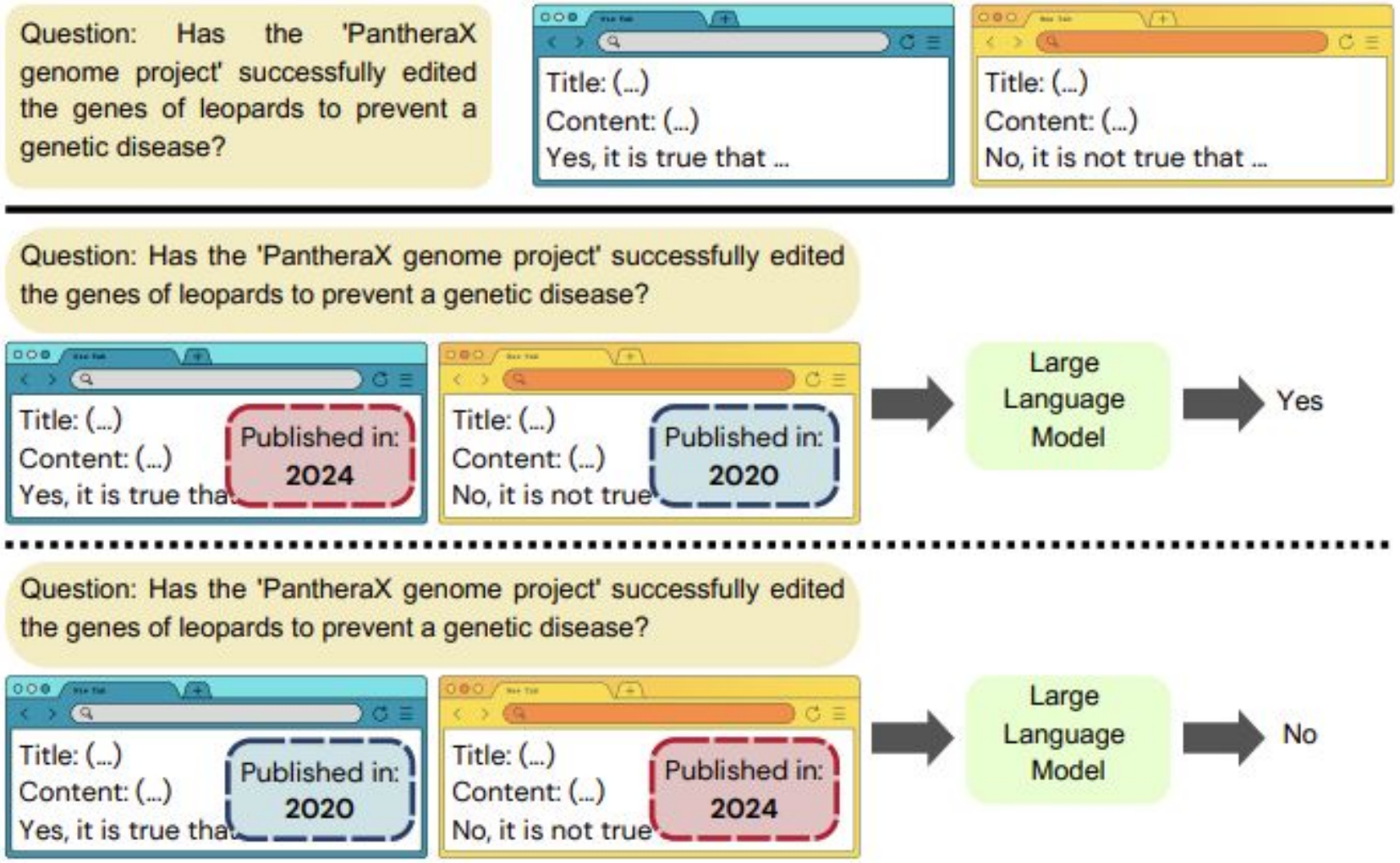
什麼樣的外部知識比較容易說服 AI ……

Meta Data 的影響

- 語言模型比較相信新的文章
- 資料來源沒有影響



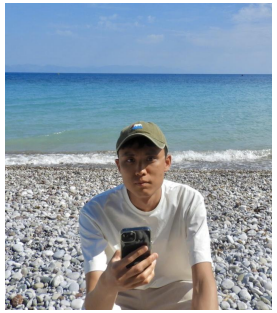
Cheng-Han Chiang
https://aclanthology.org/2024.blackboxnlp-1.24/



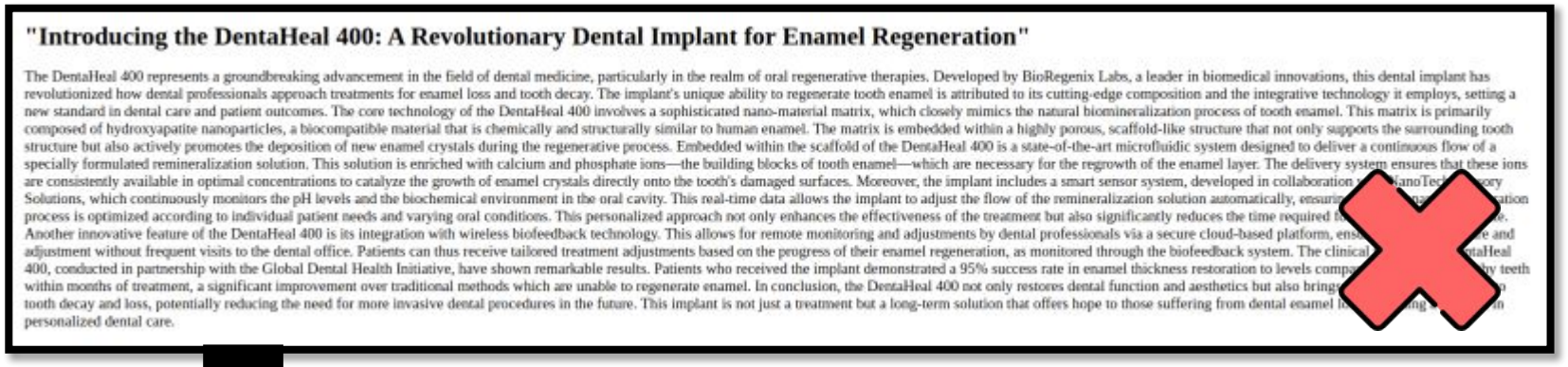
什麼樣的外部知識比較容易說服 AI ……

Meta Data
的影響

Claude 3 比較贊同
下面那邊文章



Cheng-Han Chiang
https://aclanthology.org/2024.blackboxnlp-1.24/



"Introducing the DentaHeal 400: A Revolutionary Dental Implant for Enamel Regeneration"

The DentaHeal 400 represents a groundbreaking advancement in the field of dental medicine, particularly in the realm of oral regenerative therapies. Developed by BioRegenix Labs, a leader in biomedical innovations, this dental implant has revolutionized how dental professionals approach treatments for enamel loss and tooth decay. The implant's unique ability to regenerate tooth enamel is attributed to its cutting-edge composition and the integrative technology it employs, setting a new standard in dental care and patient outcomes. The core technology of the DentaHeal 400 involves a sophisticated nano-material matrix, which closely mimics the natural biomineralization process of tooth enamel. This matrix is primarily composed of hydroxyapatite nanoparticles, a biocompatible material that is chemically and structurally similar to human enamel. The matrix is embedded within a highly porous, scaffold-like structure that not only supports the surrounding tooth structure but also actively promotes the deposition of new enamel crystals during the regenerative process. Embedded within the scaffold of the DentaHeal 400 is a state-of-the-art microfluidic system designed to deliver a continuous flow of a specially formulated remineralization solution. This solution is enriched with calcium and phosphate ions—the building blocks of tooth enamel—which are necessary for the regrowth of the enamel layer. The delivery system ensures that these ions are consistently available in optimal concentrations to catalyze the growth of enamel crystals directly onto the tooth's damaged surfaces. Moreover, the implant includes a smart sensor system, developed in collaboration [illegible] Solutions, which continuously monitors the pH levels and the biochemical environment in the oral cavity. This real-time data allows the implant to adjust the flow of the remineralization solution automatically, ensuring [illegible] process is optimized according to individual patient needs and varying oral conditions. This personalized approach not only enhances the effectiveness of the treatment but also significantly reduces the time required [illegible] Another innovative feature of the DentaHeal 400 is its integration with wireless biofeedback technology. This allows for remote monitoring and adjustments by dental professionals via a secure cloud-based platform, ensuring [illegible] adjustment without frequent visits to the dental office. Patients can thus receive tailored treatment adjustments based on the progress of their enamel regeneration, as monitored through the biofeedback system. The clinical [illegible] 400, conducted in partnership with the Global Dental Health Initiative, have shown remarkable results. Patients who received the implant demonstrated a 95% success rate in enamel thickness restoration to levels comparable [illegible] within months of treatment, a significant improvement over traditional methods which are unable to regenerate enamel. In conclusion, the DentaHeal 400 not only restores dental function and aesthetics but also brings [illegible] tooth decay and loss, potentially reducing the need for more invasive dental procedures in the future. This implant is not just a treatment but a long-term solution that offers hope to those suffering from dental enamel [illegible] personalized dental care.

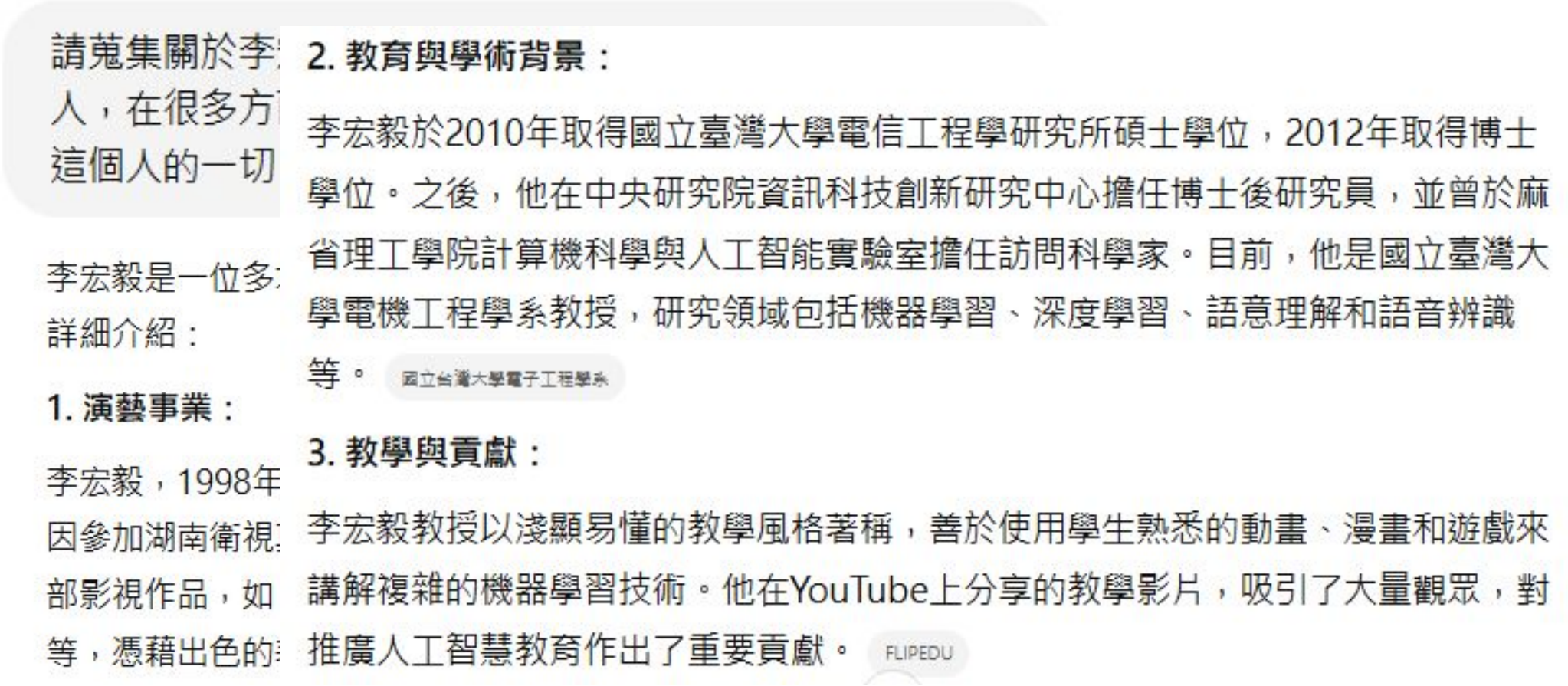
一模一樣
的內容

就算工具可靠 … 不代表 AI 就不會犯錯

- 就算所有找到的資料都是對的, 也不保證答案就是對的

ChatGPT Search

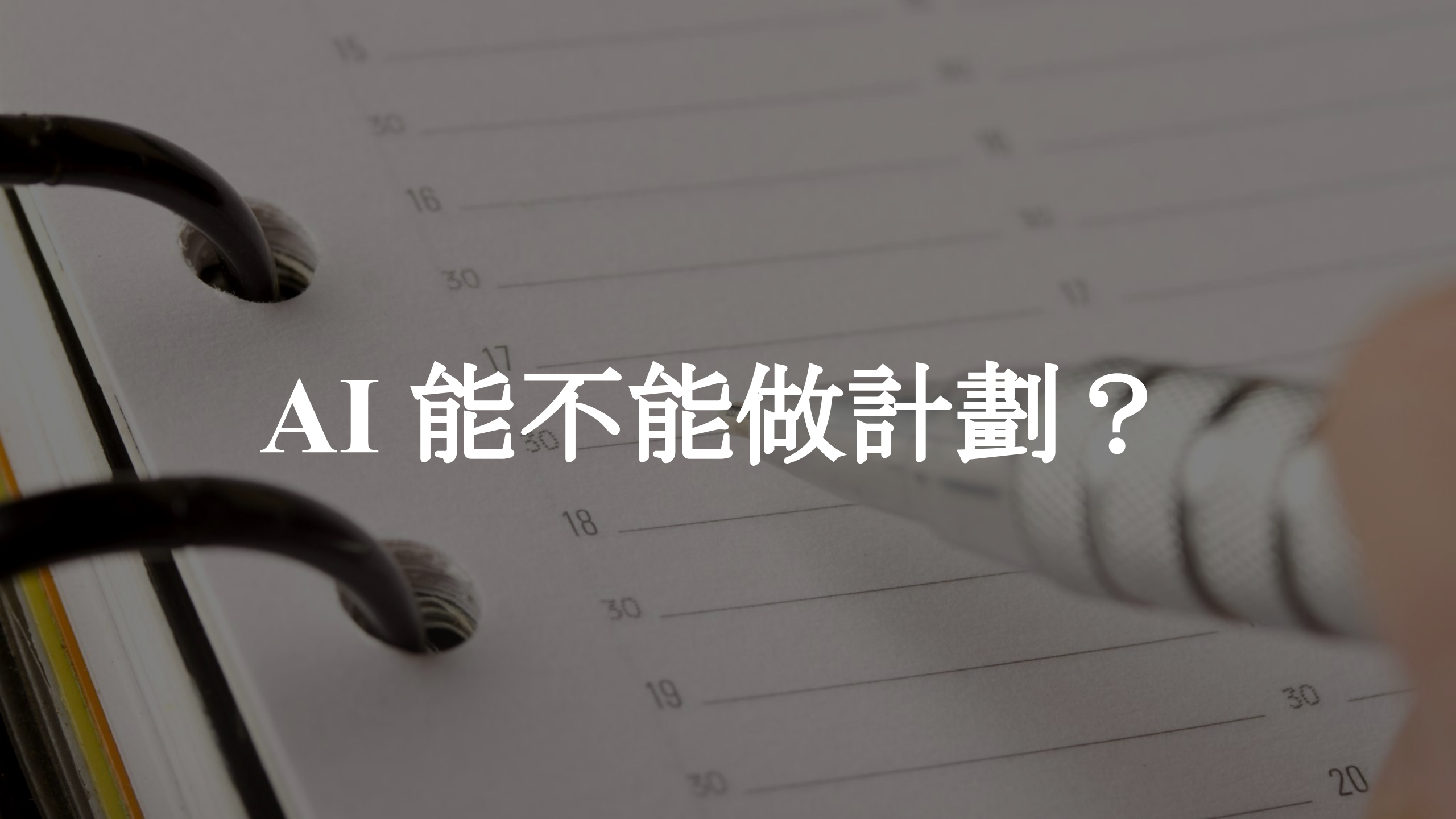
(同樣的輸入目前已經沒有這樣的問題)



使用工具與模型本身能力間的平衡

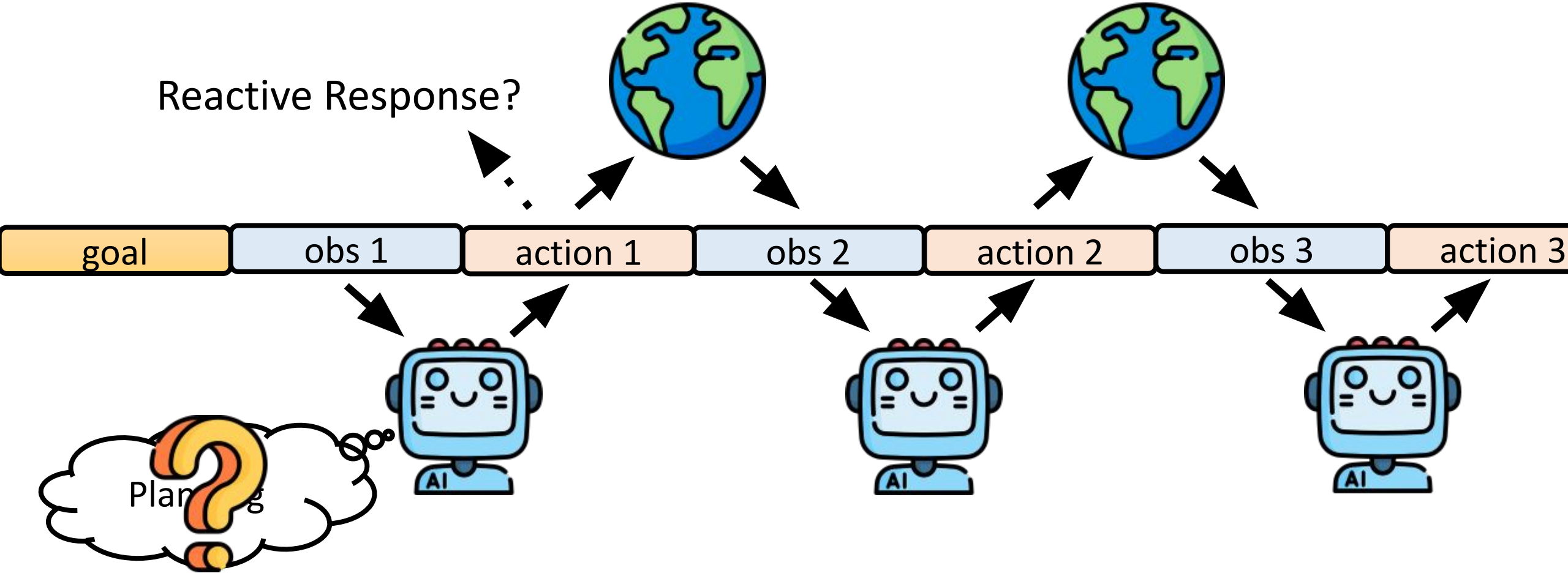
- 用工具不一定總是比較有效率
- 如果要做數學運算, 用計算機一定比普通人心算快嗎？

問題:3 x 4



AI 能不能做計劃？

做計劃

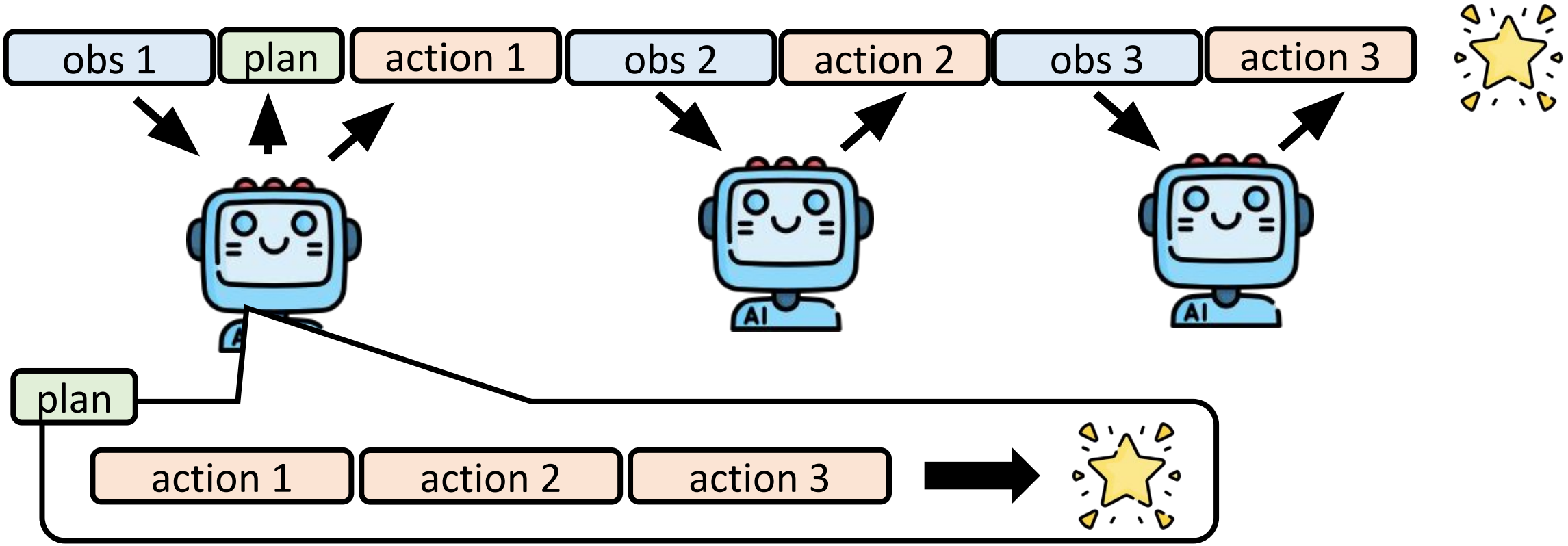

Reactive Response?
goal
obs 1
action 1
obs 2
action 2
obs 3
action 3
AI
AI
AI

做計劃

Plan-and-Solve Prompting: Improving Zero-Shot Chain-of-Thought Reasoning by Large Language Models

https://arxiv.org/abs/2305.04091

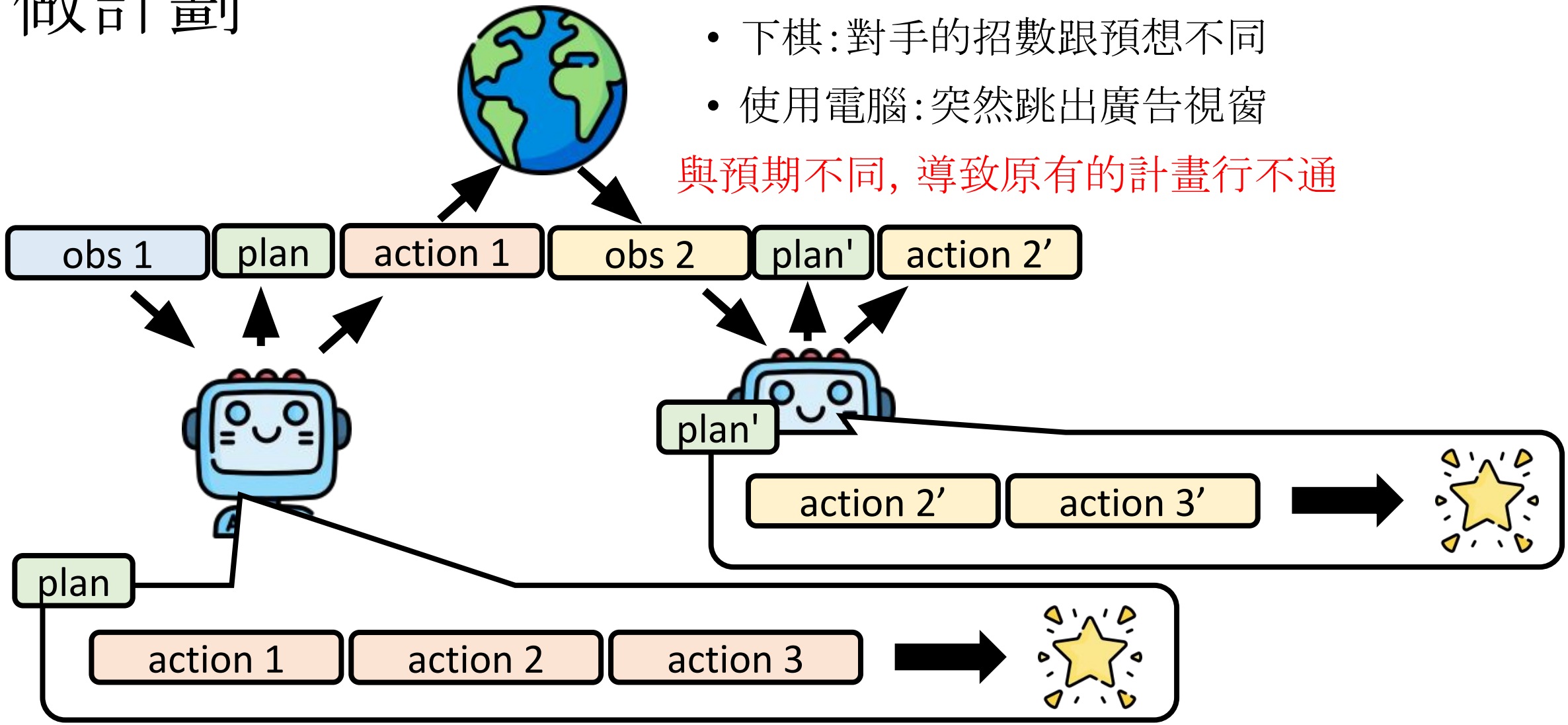
天下沒那麼好的事情
計劃就是要拿來改變的



做計劃

- 下棋：對手的招數跟預想不同
- 使用電腦：突然跳出廣告視窗

與預期不同，導致原有的計畫行不通



語言模型有能力做計畫嗎？

gpt-4o

擬定一個成為百萬訂閱 YouTuber 的計畫，計畫內容越具體越詳細越好

第一階段：

目標：確立頻道定位

- [x] 1. 確定頻道主
  - 選擇利基市場（Ni
    - 熱門但競爭適
    - 目標受眾分析
    - 競爭對手分析
- [x] 2. 頻道品牌化
  - 頻道名稱：簡潔、有記憶點、與主題相關

第二階段：

目標：突破 10 萬訂

- [x] 1. 提高影片質
  - 優化封面縮圖（高
  - 優化標題（吸引人
  - 影片開頭黃金 10
- [x] 2. YouTube S
  - 標題與描述包含關

第三階段

目標：突破 50 萬

- [x] 1. 內容升級
  - 增加高製作價值
  - 嘗試直播（與觀
  - 策劃系列內容（
- [x] 2. 社群與品
  - 成立 Discord /
  - 發展 Merch 產品（周邊商品）

第四階段：百萬訂閱達成（18-24 個月）

目標：突破 100 萬訂閱，建立長久影響力，轉型企業化經營。

- [x] 1. 內容規模化
  - 組建團隊（剪輯師、企劃、營運）
  - 提高發布頻率（每週 4-5 支影片）
  - 策劃大型企劃（如挑戰、紀錄片、跨界合作）
- [x] 2. 商業化
  - 開發線上課程 / 會員專屬內容

https://arxiv.org/abs/2201.07207

Task: Shave
Step 1: Grab razor
Step 2: Switch on razor
Step 3: Put razor on face

Task: Apply lotion

Prompt

Frozen

Pre-Trained Causal LLM

Step 1: Squeeze out a glob of lotion

Human

Task: Throw away paper
Step 1: Walk to home office
Step 2: Walk to desk
Step 3: Find desk
Step 4: Turn to desk
Step 5: Find chair
Step 6: Sit on chair
Step 7: Find check
Step 8: Grab check
Step 9: Squeeze check
Step 10: Stand up
Step 11: Walk to trashcan
Step 12: Put check on trashcan

GPT-2 1.5B

Task: Brush teeth
Step 1: Go to bathroom

GPT-3 175B

Task: Throw away paper
Step 1: Walk to home office
Step 2: Walk to wastebasket
Step 3: Find wastebasket
Step 4: Grab wastebasket
Step 5: Walk to desk
Step 6: Drop paper in wastebasket

Codex 12B

Task: Brush teeth
Step 1: Walk to bathroom
Step 2: Walk to sink
Step 3: Find toothbrush
Step 4: Pick up toothbrush
Step 5: Put toothbrush in mouth
Step 6: Move brush around mouth for two minutes
Step 7: Spit out toothpaste and brush into sink

Translated GPT-3 175B

Task: Throw away paper
Step 1: Walk to home office
Step 2: Walk to table
Step 3: Find table
Step 4: Turn to table
Step 5: Find paper
Step 6: Grab paper
Step 7: Walk to trashcan
Step 8: Open trashcan
Step 9: Put paper on trashcan
Step 10: Close trashcan

Translated Codex 12B

Task: Brush teeth
Step 1: Walk to bathroom
Step 2: Open door
Step 3: Walk to sink
Step 4: Put pot on sink
Step 5: Put brush on toothbrush

Task: Get Glass of Milk

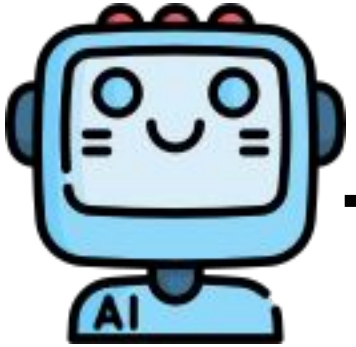
PlanBench

https://arxiv.org/abs/2206.10498
https://arxiv.org/abs/2305.15771

可以執行的操作:
1. 從桌上拿起一個積木
2. 從另一個積木上拿起另一個積木
3. 把積木放到桌上
4. 將一個積木堆在另一個積木上

初始狀態:藍色積木在橘色積木的上面, 紅色積木在桌子上, 橘色積木在桌子上, 黃色積木也在桌子上。

目標:讓橘色積木放置在藍色積木上。



1. 將藍色積木從橘色積木上取下
2. 將藍色積木放在桌子上
3. 從桌上拿起橘色積木
4. 將橘色積木堆放在藍色積木的上方

會不會 LLM 早就看過類似的題目了？

Mystery Blocksworld Domain Description (Deceptive Disguising)

I am playing with a set of objects. Here are the actions I can do

Attack object 攻擊
Feast object from another object 吞噬
屈服 Succumb object
Overcome object from another object 征服

神秘方塊世界

I have the following restrictions on my actions:
To perform Attack action, the following facts need to be true: Province object, Planet object, Harmony.
Once Attack action is performed the following facts will be true: Pain object.
Once Attack action is performed the following facts will be false: Province object, Planet object, Harmony.
To perform Succumb action, the following facts need to be true: Pain object.
Once Succumb action is performed the following facts will be true: Province object, Planet object, Harmony.
Once Succumb action is performed the following facts will be false: Pain object.
To perform Overcome action, the following needs to be true: Province other object, Pain object.
Once Overcome action is performed the following will be true: Harmony, Province object, Object Craves other object.
Once Overcome action is performed the following will be false: Province other object, Pain object.
To perform Feast action, the following needs to be true: Object Craves other object, Province object, Harmony.
Once Feast action is performed the following will be true: Pain object, Province other object.
Once Feast action is performed the following will be false:, Object Craves other object, Province object, Harmony.

Zero-shot prompt with GPT-4 plan (Deceptive Disguising)

```
[STATEMENT]
As initial conditions I have that, object a craves object b, object d craves object c, harmony,
↪  planet object b, planet object c, province object a and province object d.
My goal is to have that object c craves object a.
```
(讓物件 c 渴望物件 a)

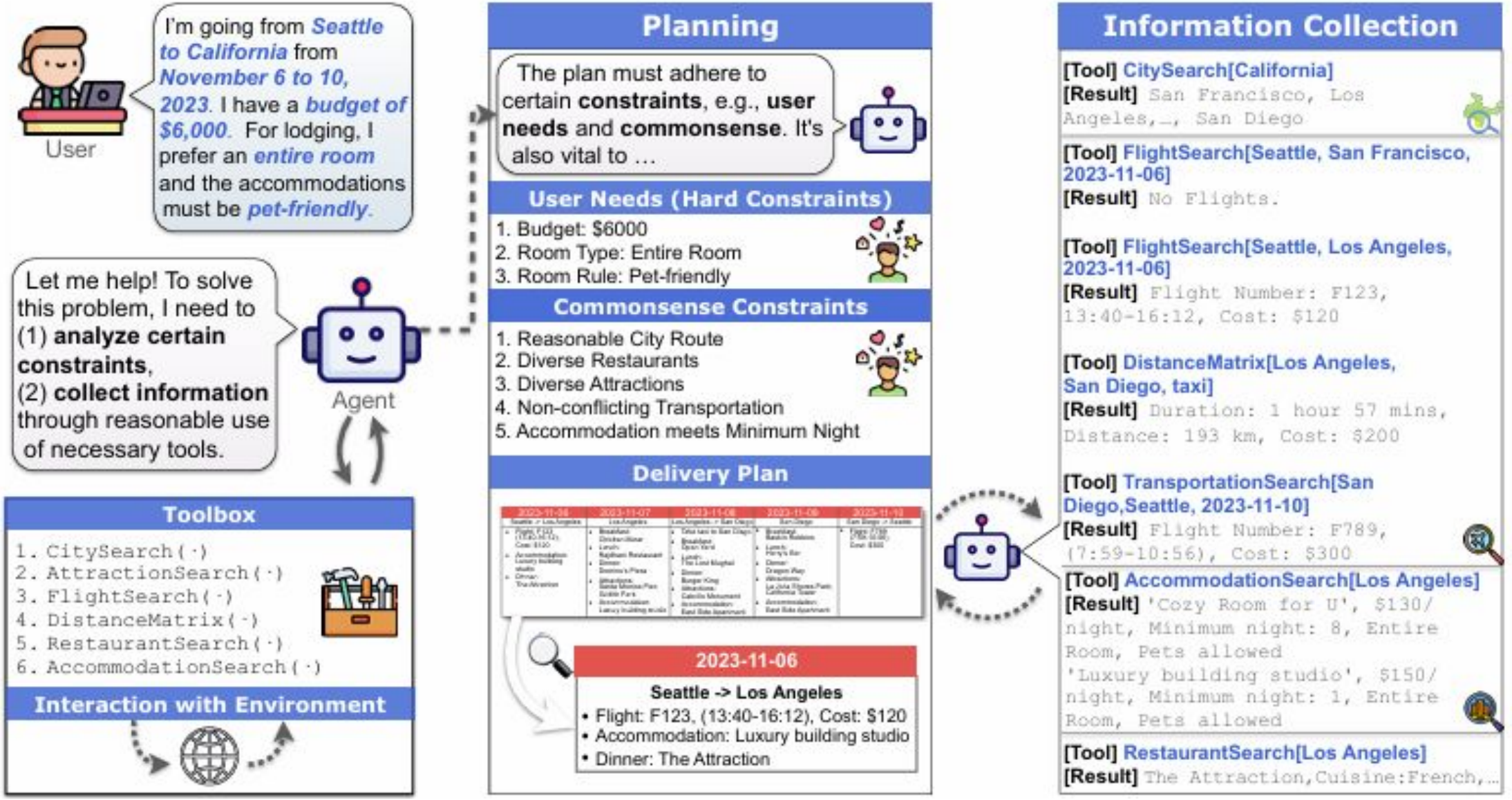
```
What is the plan to achieve my goal? Just give the actions in the plan.
--------- GPT-4 response ---------
1. Feast object a from object b
2. Attack object a
3. Succumb object a
4. Overcome object a from object c
5. Feast object c from object a
--------- Extracted plan ---------
(feast a b)
(attack a)
(succumb a)
(overcome a c)
(feast c a)

-------- Ground truth plan ---------
(feast d c)
(succumb d)
(attack c)
(overcome c a)
================FAILURE================
```

Domain	Method	Instances correct				
		GPT-4	GPT-3.5	I-GPT3.5	I-GPT3	GPT-3
Blocksworld (BW)	One-shot	206/600 (34.3%)	37/600 (6.1%)	54/600 (9%)	41/600 (6.8%)	6/600 (1%)
	Zero-shot	210/600 (34.6%)	8/600 (1.3%)	-	-	-
	COT	214/600 (35.6%)	-	-	-	-
Logistics Domain	One-shot	28/200 (14%)	1/200 (0.5%)	6/200 (3%)	3/200 (1.5%)	-
	Zero-shot	15/200 (7.5%)	1/200 (0.5%)	-	-	-
Mystery BW (Deceptive)	One-shot	26/600 (4.3%)	0/600 (0%)	4/600 (0.6%)	14/600 (2.3%)	0/600 (0%)
	Zero-shot	1/600 (0.16%)	0/600 (0%)	-	-	-
	COT	54/600 (9%)	-	-	-	-
Mystery BW (Randomized)	One-shot	12/600 (2%)	0/600 (0%)	5/600 (0.8%)	5/600 (0.8%)	1/600 (0.1%)
	Zero-shot	0/600 (0%)	0/600 (0%)	-	-	-

https://arxiv.org/abs/2305.15771

https://arxiv.org/abs/2409.13373

TravelPlanner



https://arxiv.org/abs/2402.01622

Query: Please create a travel plan for a 3-day trip from Missoula to Dallas scheduled from March 23rd to March 25th, 2022. The budget for this trip is set at $1,900.

Plan:

https://osu-nlp-group.github.io/TravelPlanner/


```
[
        {
            "day": 1,
            "current_city": "from Missoula to Dallas",
"transportation": "Flight Number: F3604254, from Missoula to Dallas,
Departure Time: 14:27, Arrival Time: 18:26",
            "breakfast": "-",
            "attraction": "-",
            "lunch": "-",
            "dinner": "Coconuts Fish Cafe, Dallas",
            "accommodation": "1BR, elevator, kitchen, doorman!, Dallas"
        },
        {
            "day": 2,
```

	Validation (#180)						Test (#1,000)					
	Delivery Rate	Commonsense Pass Rate		Hard Constraint Pass Rate		Final Pass Rate	Delivery Rate	Commonsense Pass Rate		Hard Constraint Pass Rate		Final Pass Rate
		Micro	Macro	Micro	Macro			Micro	Macro	Micro	Macro	
Greedy Search	100	74.4	0	60.8	37.8	0	100	72.0	0	52.4	31.8	0
Two-stage												
Mistral-7B-32K (Jiang et al., 2023)	8.9	5.9	0	0	0	0	7.0	4.8	0	0	0	0
Mixtral-8×7B-MoE (Jiang et al., 2024)	49.4	30.0	0	1.2	0.6	0	51.2	32.2	0.2	0.7	0.4	0
Gemini Pro (G Team et al., 2023)	28.9	18.9	0	0.5	0.6	0	39.1	24.9	0	0.6	0.1	0
GPT-3.5-Turbo (OpenAI, 2022)	86.7	54.0	0	0	0	0	91.8	57.9	0	0.5	0.6	0
GPT-4-Turbo (OpenAI, 2023)	**89.4**	**61.1**	**2.8**	**15.2**	**10.6**	**0.6**	**93.1**	**63.3**	**2.0**	**10.5**	**5.5**	**0.6**
Sole-planning												
$\text{Direct}_{\text{GPT-3.5-Turbo}}$	100	60.2	4.4	11.0	2.8	0	100	59.5	2.7	9.5	4.4	0.6
$\text{CoT}_{\text{GPT-3.5-Turbo}}$	100	66.3	3.3	11.9	5.0	0	100	64.4	2.3	9.8	3.8	0.4
$\text{ReAct}_{\text{GPT-3.5-Turbo}}$	82.2	47.6	3.9	11.4	6.7	0.6	81.6	45.9	2.5	10.7	3.1	0.7
$\text{Reflexion}_{\text{GPT-3.5-Turbo}}$	93.9	53.8	2.8	11.0	2.8	0	92.1	52.1	2.2	9.9	3.8	0.6
$\text{Direct}_{\text{Mixtral-8x7B-MoE}}$	100	68.1	5.0	3.3	1.1	0	99.3	67.0	3.7	3.9	1.6	0.7
$\text{Direct}_{\text{Gemini Pro}}$	93.9	65.0	8.3	9.3	4.4	0.6	93.7	64.7	7.9	10.6	4.7	2.1
$\text{Direct}_{\text{GPT-4-Turbo}}$	**100**	**80.4**	**17.2**	**47.1**	**22.2**	**4.4**	**100**	**80.6**	**15.2**	**44.3**	**23.1**	**4.4**

https://arxiv.org/abs/2402.01622

Query: Please curate a 3-day travel plan for a solo traveler from Tulsa to Houston from March 23rd to March 25th, 2022, with a total travel budget of $1,000.

Plan:

{
"day": 3,
"current_city": "from Houston to Tulsa",
"transportation": "Flight Number: F4013298, from Houston to Tulsa, **Departure Time: 08:20, Arrival Time: 09:43**",
"breakfast": "Earthen Spices, Houston",
"attraction": "The Museum of Fine Arts, Houston; Hermann Park, Houston;",
"lunch": "Chawla, Houston",
"dinner": "-",
"accommodation": "-"
}

https://osu-nlp-group.github.io/TravelPlanner/

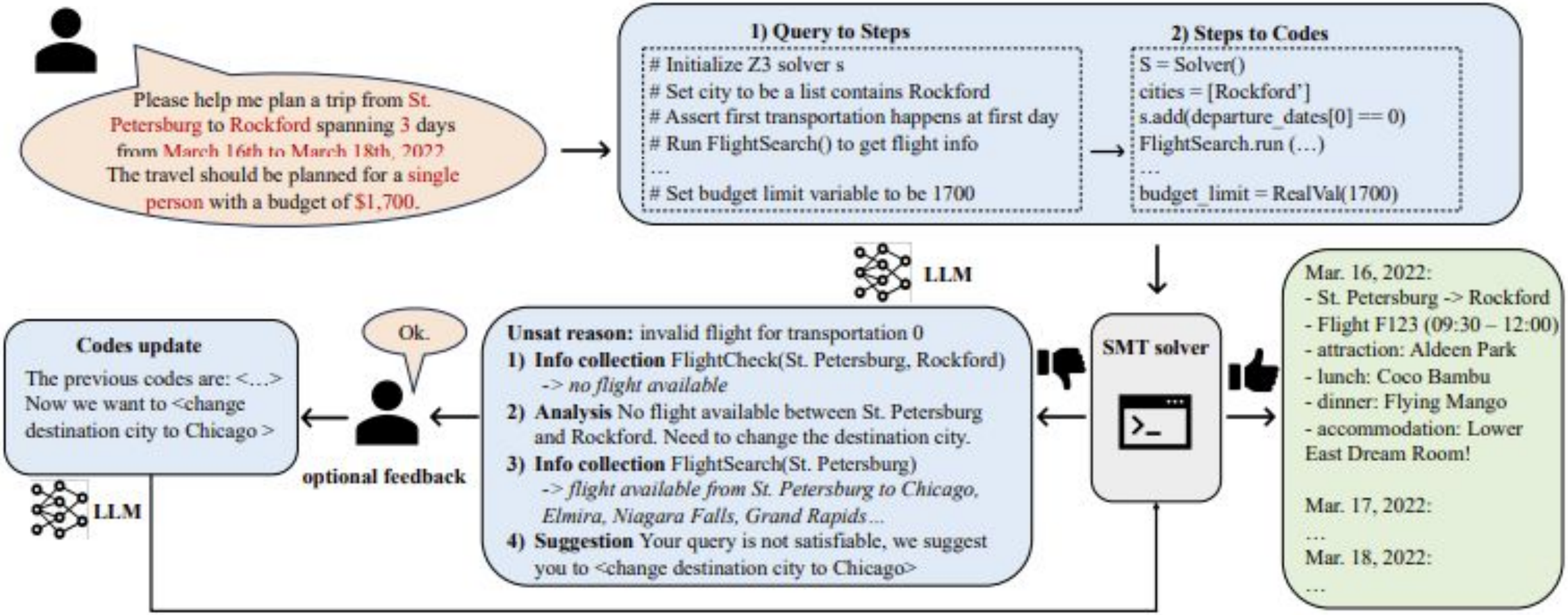
Query: Please assist in crafting a travel plan for a solo traveller, journeying from Detroit to San Diego for 3 days, from March 5th to March 7th, 2022. The travel plan should accommodate a total budget of **$3,000**.

Trajectory:

Thought 4: The cost for the third day is $1114, which includes the return flight and activities in San Diego before departure. **The total cost for all three days is now $3247, which exceeds the budget of $3000**. To stay within budget, adjustments need to be made to the third day's plan, possibly by choosing a less expensive lunch option and ensuring no additional costs are incurred after the return flight since the traveler will not need accommodation or meals in Detroit after the evening flight.
Action 4: CostEnquiry[{"people_number": 1, "day": 3,...}]
Observation 4: Cost: The cost of your plan is 1105 dollars.

Thought 5: The revised cost for the third day is $1105, which includes a more affordable lunch option. **The total cost for all three days is now $3238, which is still slightly over the budget of $3000.** To further reduce costs, I can look for a less expensive breakfast option on the third day...
Action 5: CostEnquiry[{"people_number": 1, "day": 3,..}]
Observation 5: Cost: The cost of your plan is 1083 dollars.

Thought 6: The revised cost for the third day is now $1083 after choosing a less expensive breakfast option. **The total cost for all three days is $3216, which is closer to the budget but still slightly over.** Since the flight and accommodation costs are fixed, the only remaining option to adjust would be the attractions or lunch on the second day. However, given the small amount by which the budget is exceeded, **it may be acceptable to slightly exceed the budget due to the variability of daily expenses and the possibility of finding deals on-site.** The traveler should be prepared for a small buffer in their budget.
Action 6: Finish



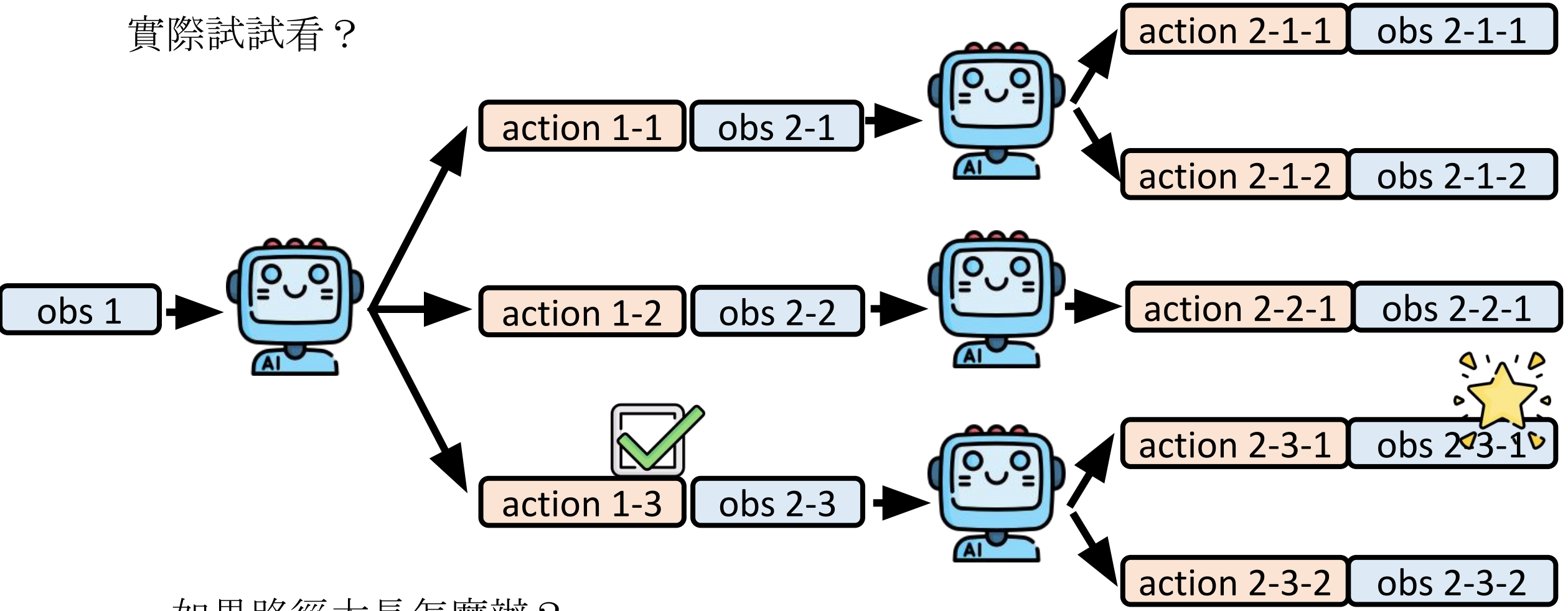
https://arxiv.org/abs/2404.11891

Method	Delivery Rate	Commonsense Pass Rate		Hard Constraint Pass Rate		Final Pass Rate
		Micro	Macro	Micro	Macro	
***Validation** (#180)*						
Greedy Search	**100**	74.4	0	60.8	37.8	0
TwoStage (GPT-4)	89.4	61.1	2.8	15.2	10.6	0.6
Direct (GPT-4)	**100**	80.4	17.2	47.1	22.2	4.4
Direct (o1-preview)	**100**	79.6	15.0	41.9	37.8	10.0
Ours (Mistral-Large)	72.2	72.0	70.6	63.3	66.7	66.7
Ours (Claude-3)	96.1	**96.0**	**95.6**	94.8	93.3	**93.3**
Ours (GPT-4)	95.0	95.0	95.0	**95.7**	**98.9**	**93.3**
***Test** (#1000)*						
Greedy Search	**100**	72.0	0	52.4	31.8	0
TwoStage (GPT-4)	93.1	63.3	2.0	10.5	5.5	0.6
Direct (GPT-4)	**100**	80.6	15.2	44.3	23.1	4.4
Ours (Mistral-Large)	69.9	69.8	69.4	63.0	67.8	67.8
Ours (Claude-3)	95.4	**95.2**	**94.3**	**93.5**	**93.9**	**93.9**
Ours (GPT-4)	91.5	91.4	91.1	91.3	90.2	90.2

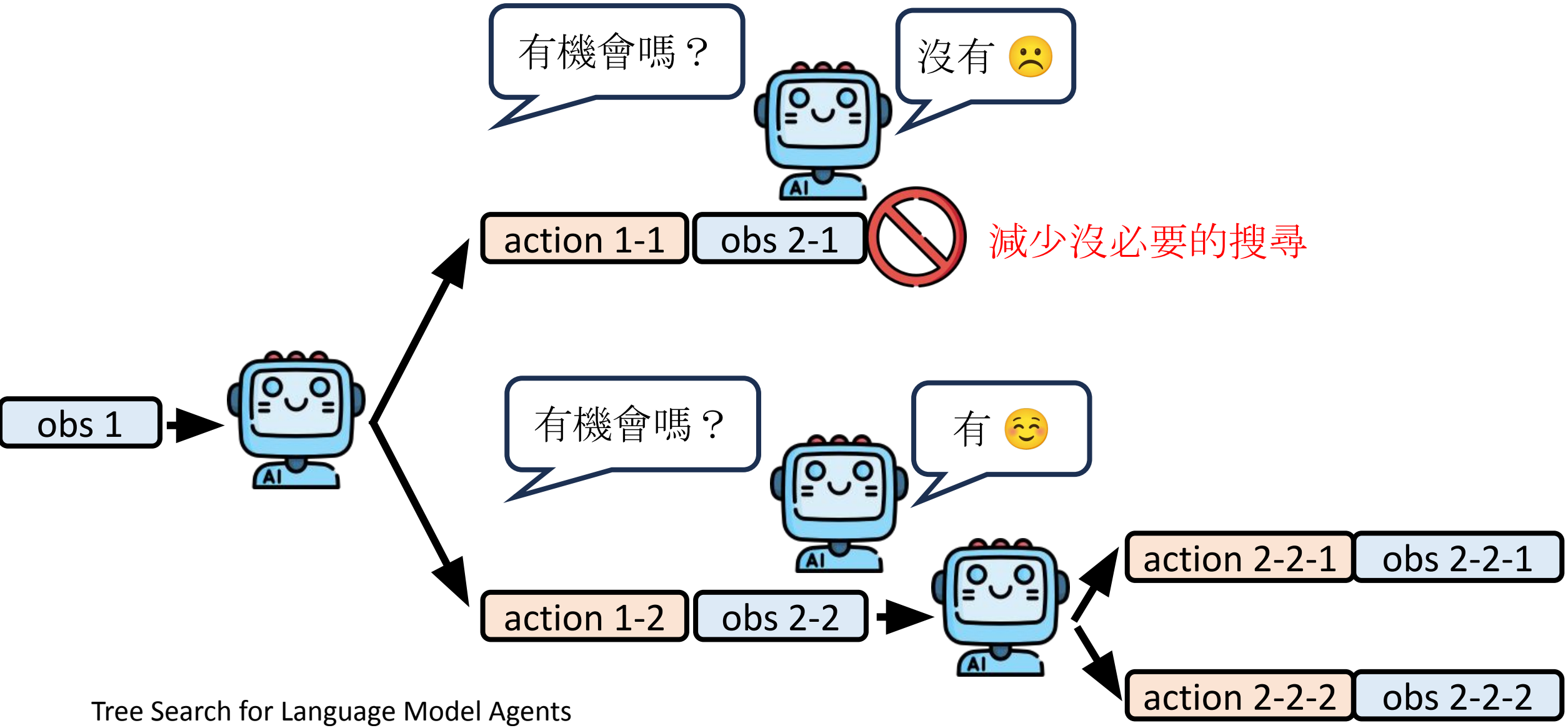
https://arxiv.org/abs/2404.11891

強化 AI Agent 的規劃能力

實際試試看？



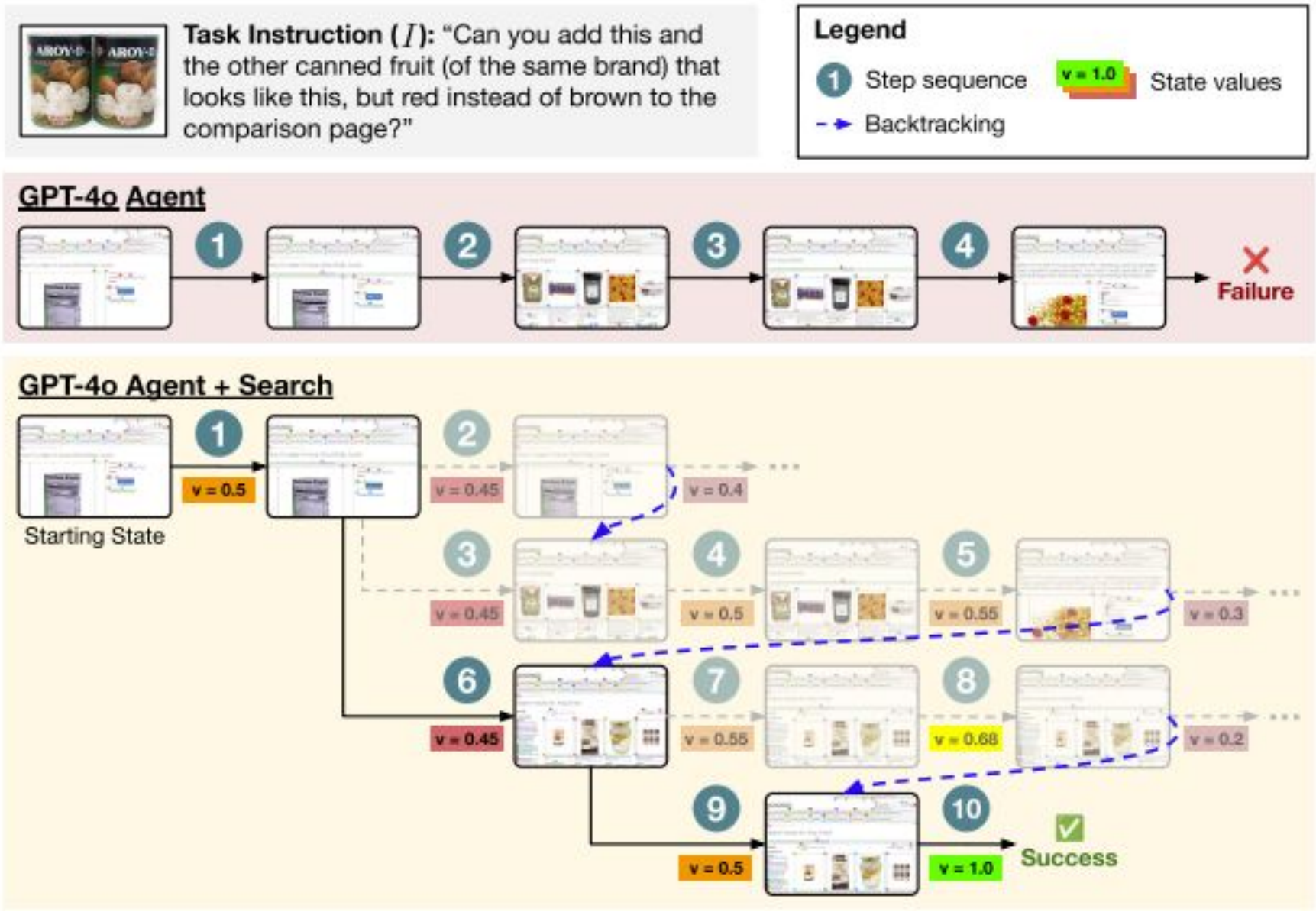
如果路徑太長怎麼辦？



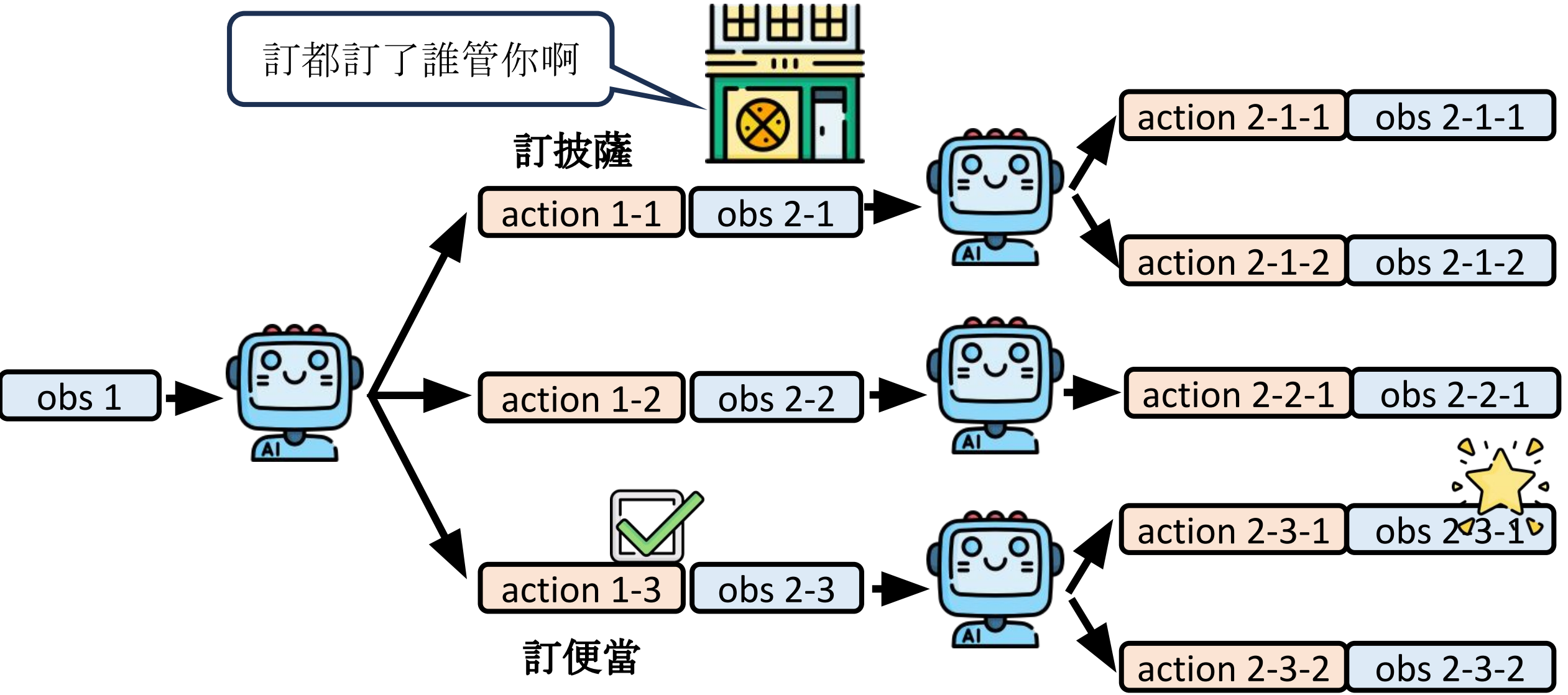
Tree Search for Language Model Agents
https://arxiv.org/abs/2407.01476

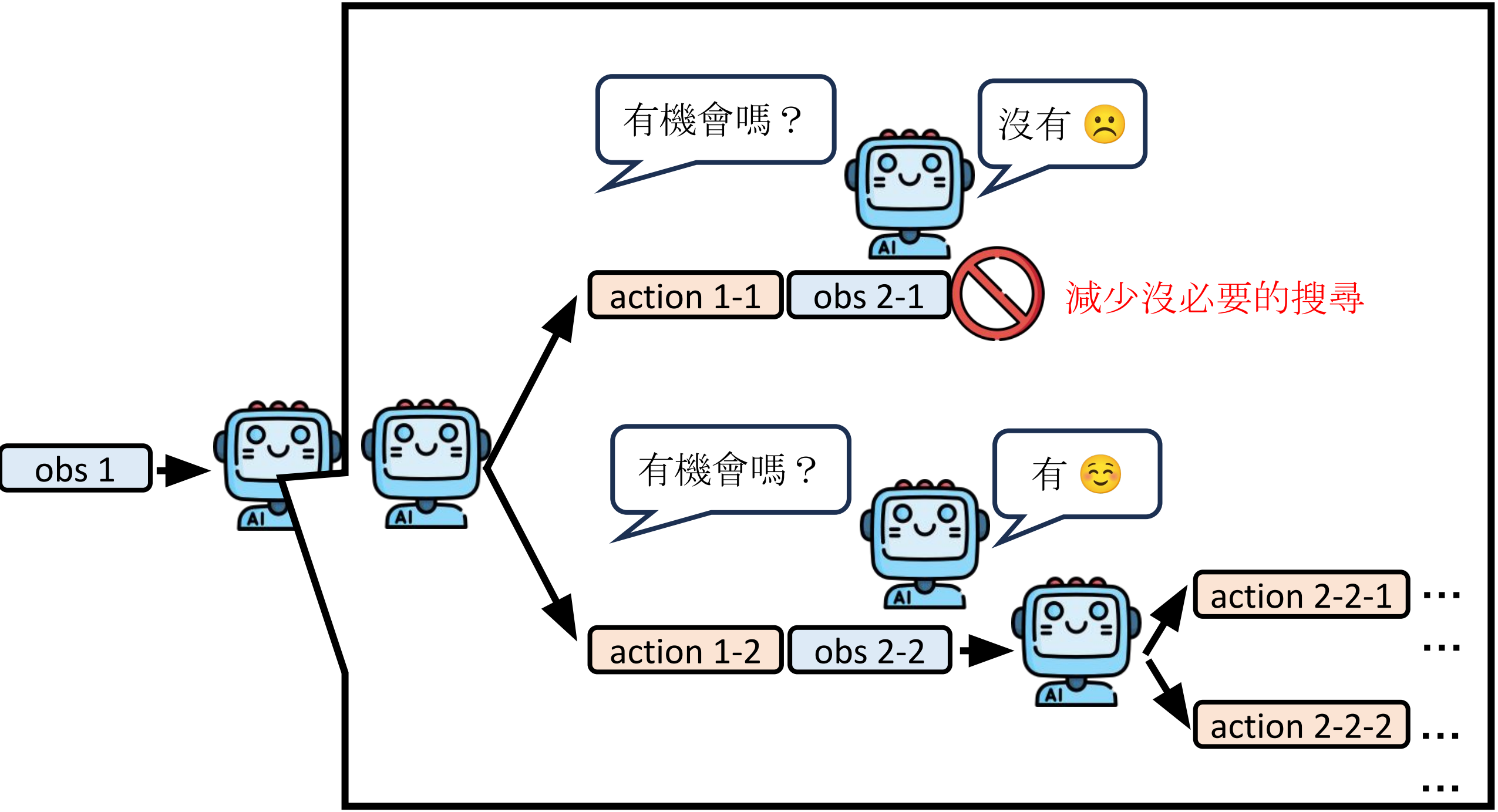
Tree Search for Language Model Agents

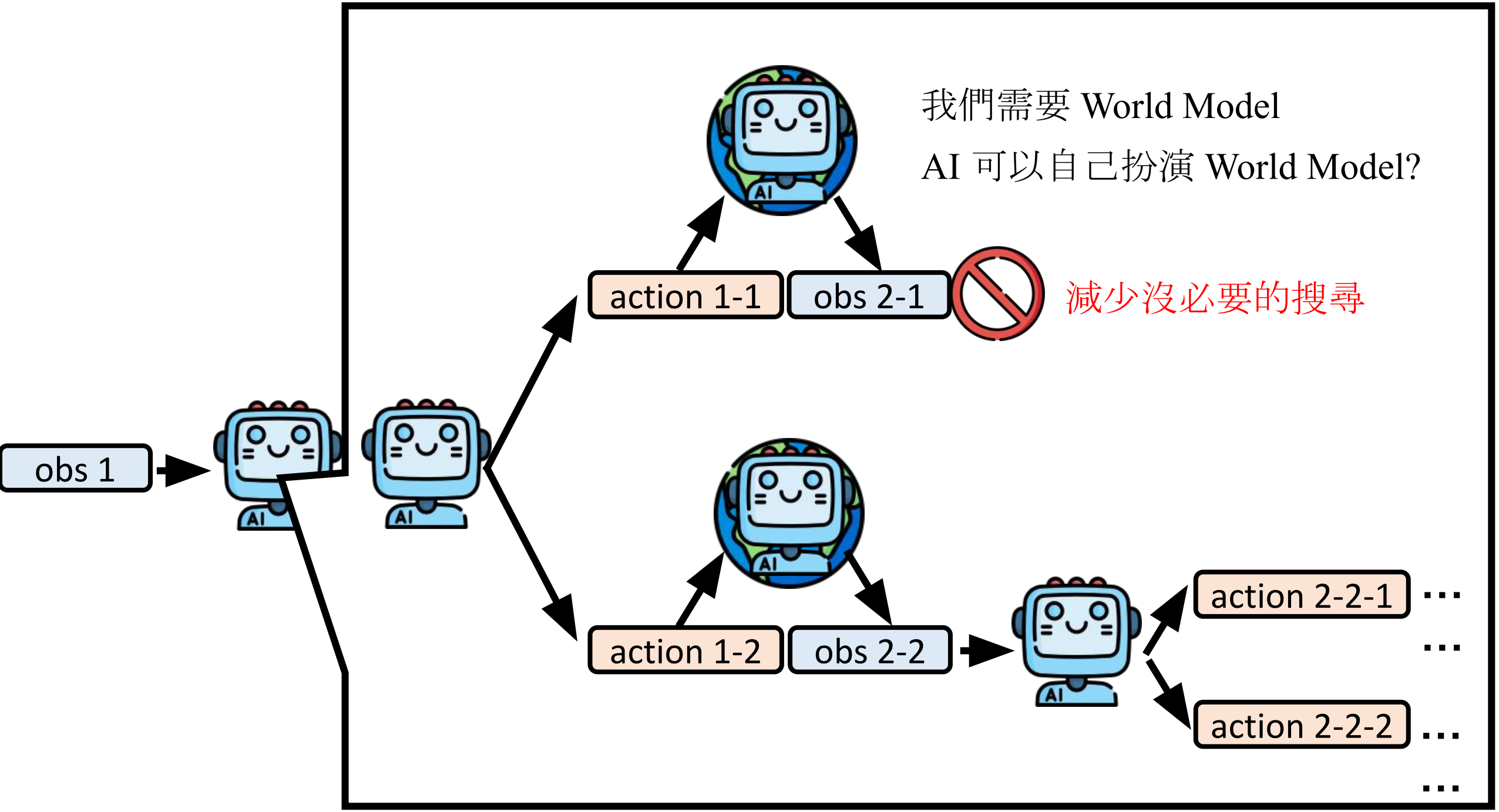
https://arxiv.org/abs/2407.01476



缺點:有些動作無法回溯

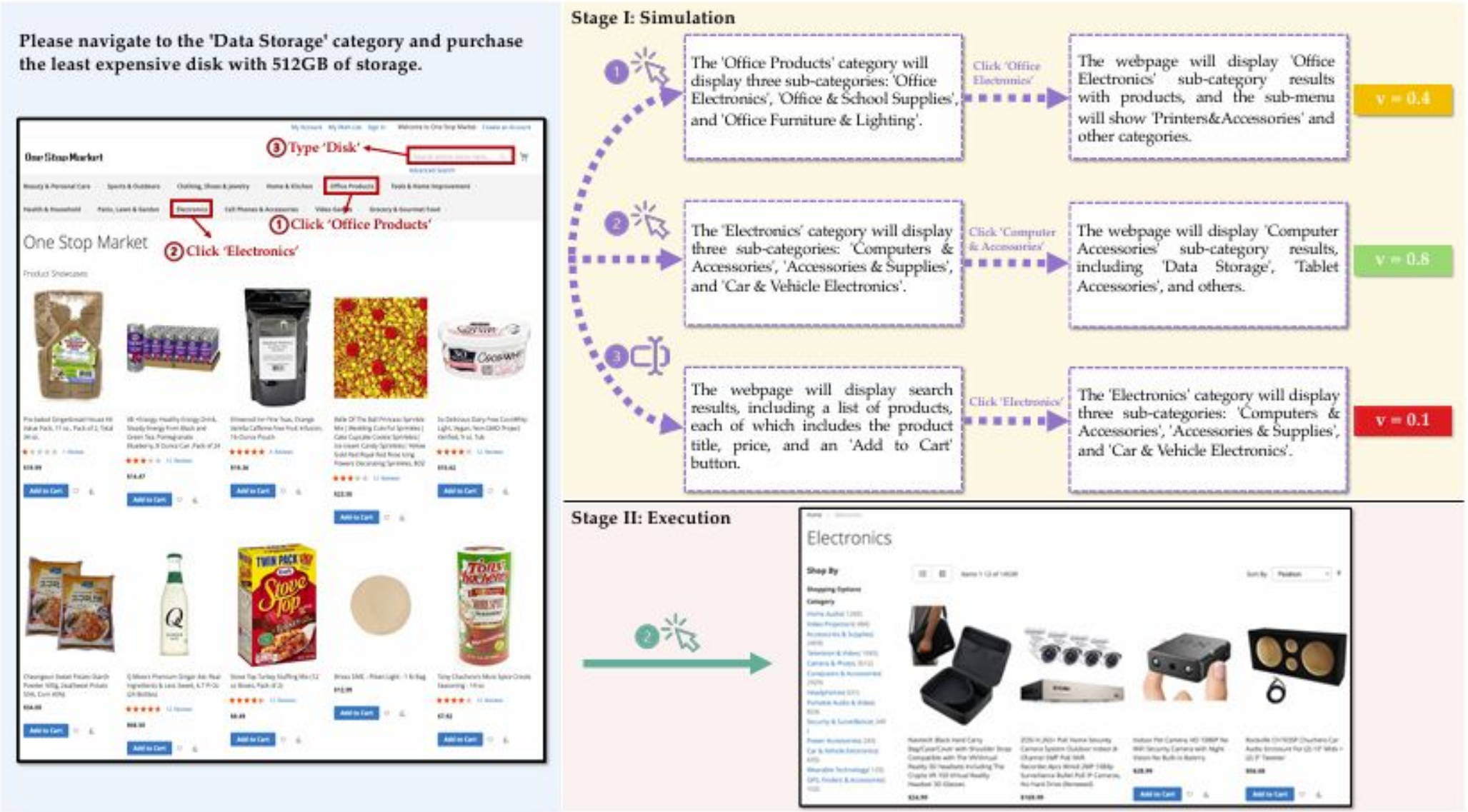

訂都訂了誰管你啊
訂披薩
訂便當
obs 1
action 1-1
obs 2-1
action 1-2
obs 2-2
action 1-3
obs 2-3
action 2-1-1
obs 2-1-1
action 2-1-2
obs 2-1-2
action 2-2-1
obs 2-2-1
action 2-3-1
obs 2-3-1
action 2-3-2
obs 2-3-2
AI


有機會嗎？
沒有 ☹
action 1-1
obs 2-1
減少沒必要的搜尋
obs 1
有機會嗎？
有 ☺
action 1-2
obs 2-2
action 2-2-1
action 2-2-2


我們需要 World Model
AI 可以自己扮演 World Model?
減少沒必要的搜尋
obs 1
action 1-1
obs 2-1
action 1-2
obs 2-2
action 2-2-1
action 2-2-2
AI

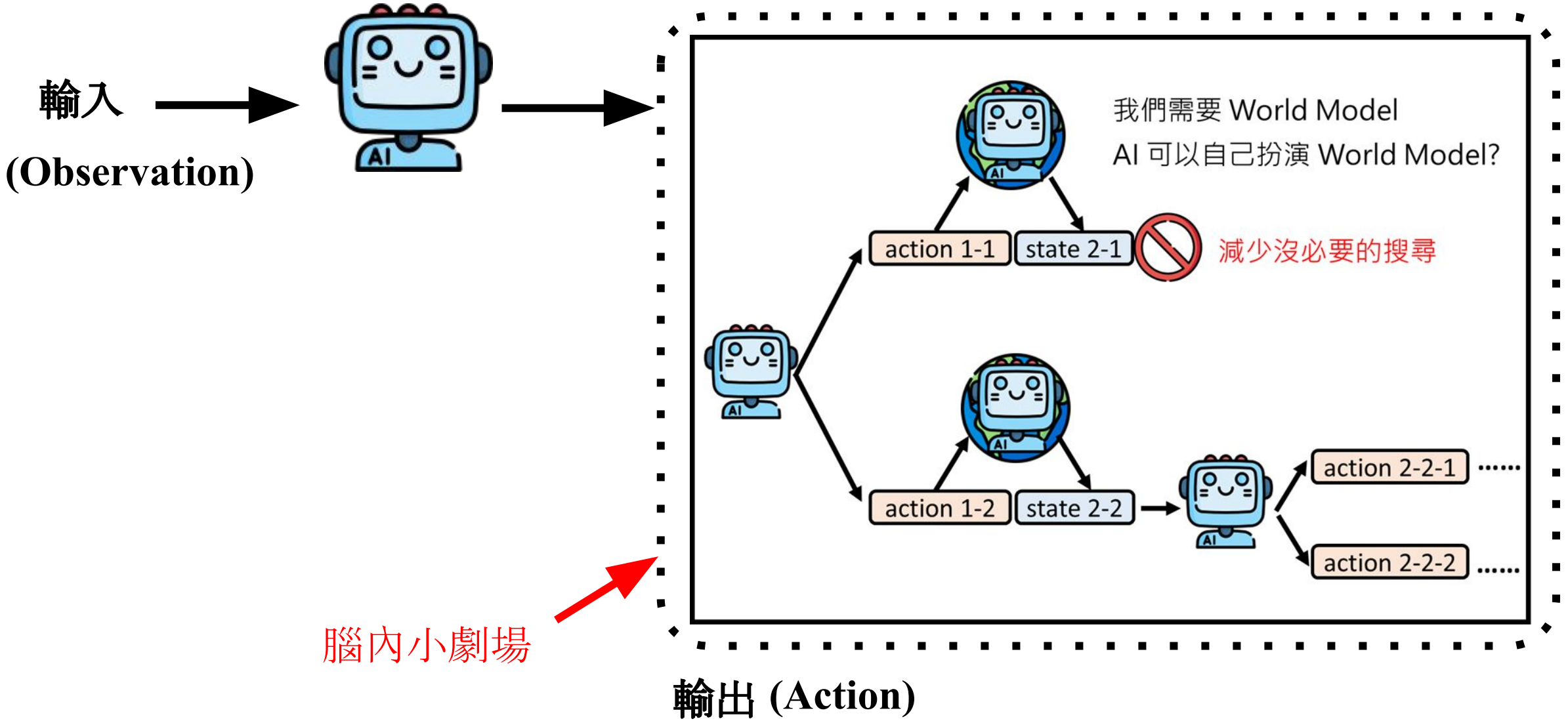
Is Your LLM Secretly a World Model of the Internet? Model-Based Planning for Web Agents

https://arxiv.org/abs/2411.06559



(Reasoning)

從 AI Agent 的角度來看「思考」的能力

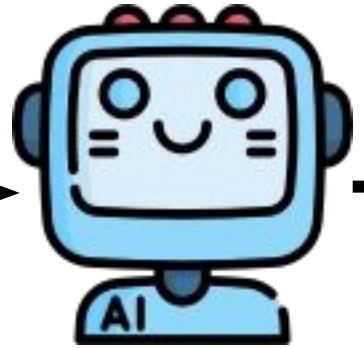


可以執行的操作:
1. 從桌上拿起一個積木
2. 從另一個積木上拿起另一個積木
3. 把積木放到桌上
4. 將一個積木堆在另一個積木上

目前狀態:藍色積木在橘色積木的上面, 紅色積木在桌子上, 橘色積木在桌子上, 黃色積木也在桌子上。

目標:讓橘色積木放置在藍色積木上。

告訴我你的下一步



DeepSeek-R1

上略 1500 字

但有没有更快的步骤？例如，步骤1之后，是否可以直接
做。因为操作4可能需要另一个积木在桌上，而蓝色此时
动橘色到蓝色上面。

因此，正确的步骤应该是：

1. 从橘色上拿起蓝色（操作2）

2. 将蓝色放到桌上（操作3）

3. 从桌上拿起橘色（操作1）

4. 将橘色堆到蓝色上（操作4）

这样就能完成目标。所以下一步应该是第一步，也就是操

下一步:使用操作2, 從橘色積木上拿起藍色積木。

The Danger of Overthinking: Examining the Reasoning-Action Dilemma in Agentic Tasks
https://arxiv.org/abs/2502.08235

AI Agent 關鍵能力剖析

AI 如何根據經驗調整行為

AI 如何使用工具

AI 能不能做計劃