



AI Agents

Qiang Sun

University of Toronto

How AI Is Used Today

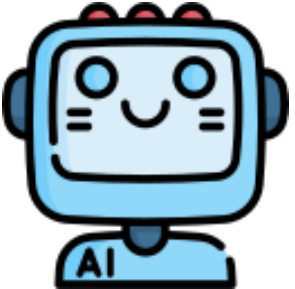
Humans provide explicit instructions

AI: one command → one action



How to translate "AI Agent"

"人工智能智能体"



AI Agent

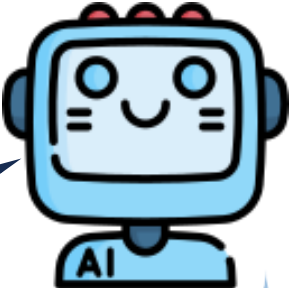
Humans provide a goal

AI figures out how to achieve it



(Solve a research problem)

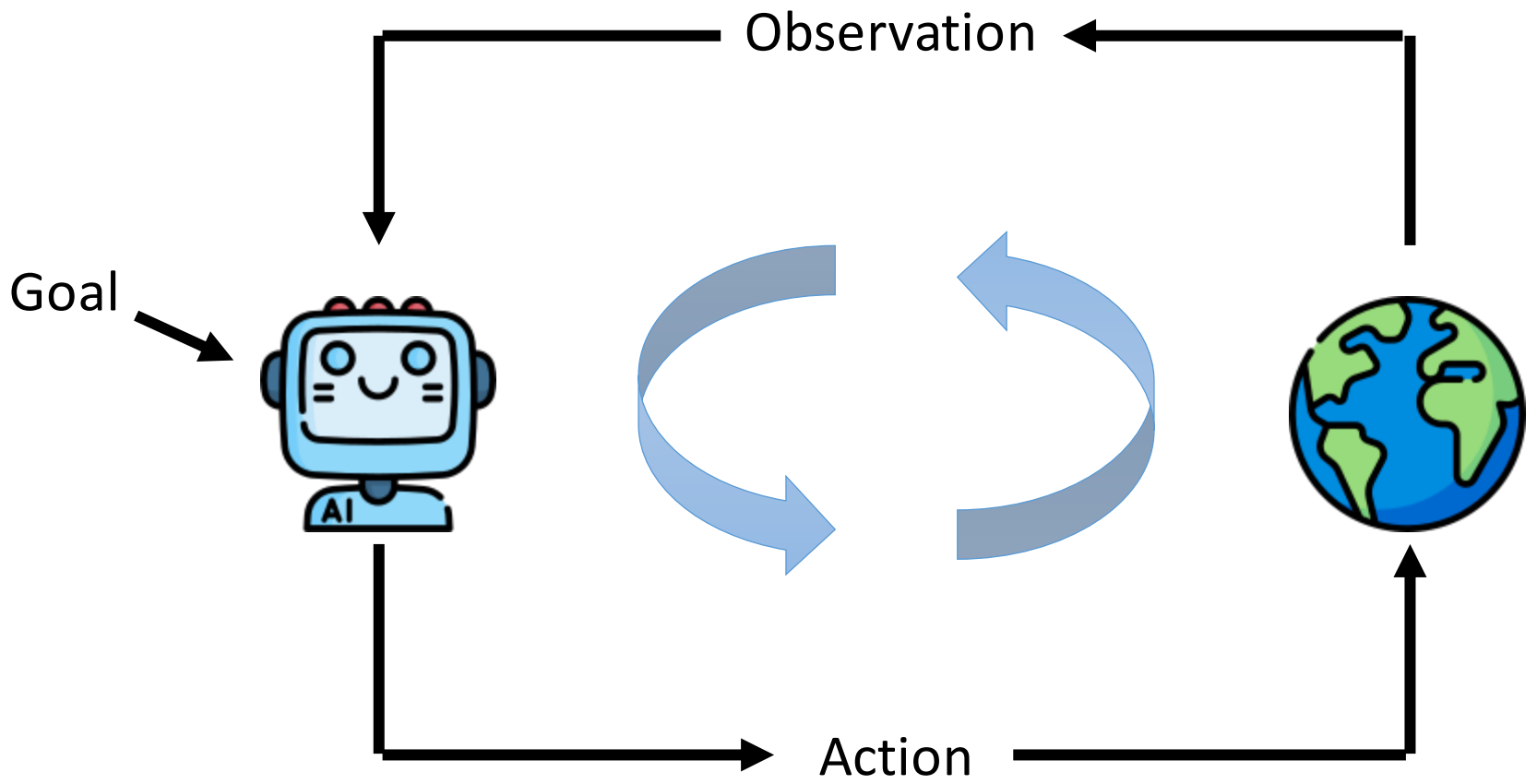
Hypothesis ... Experiment ... Analysis



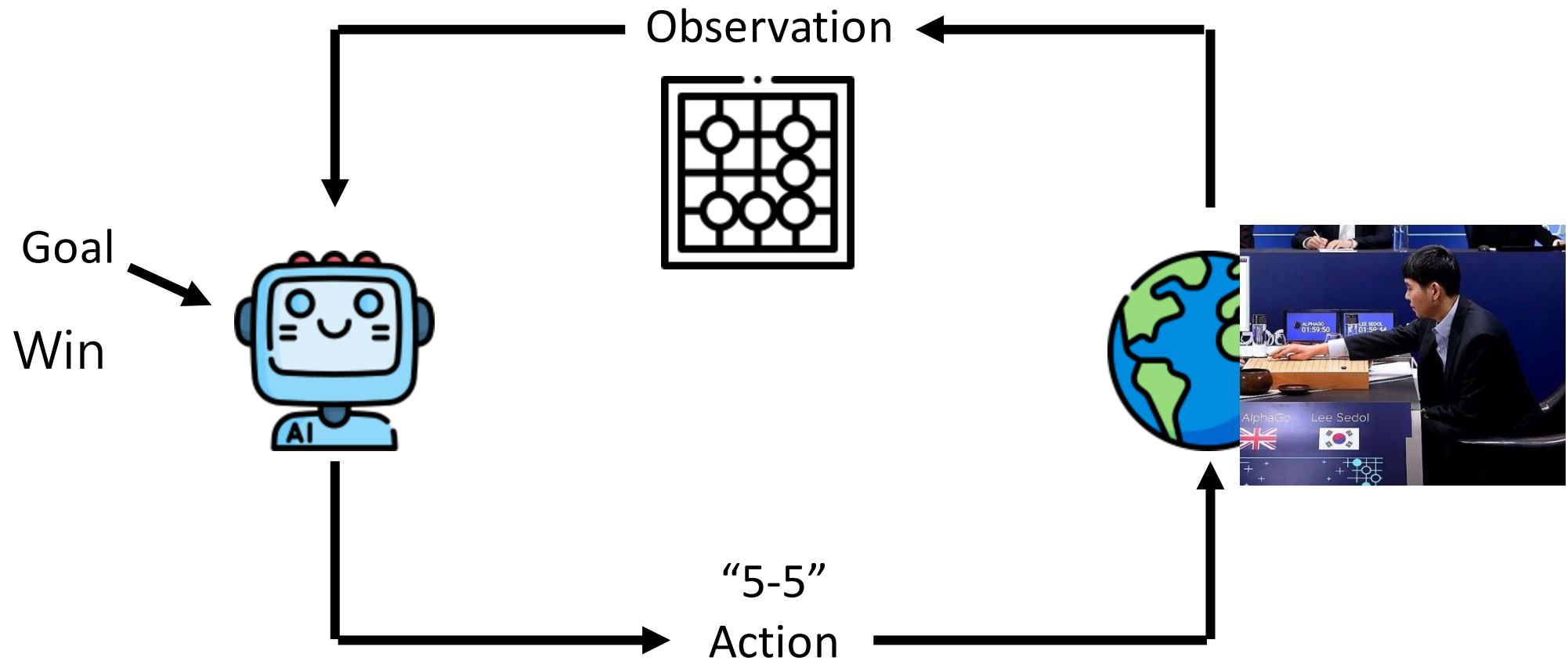
Requires multiple steps and flexible plan adjustment



AI Agent



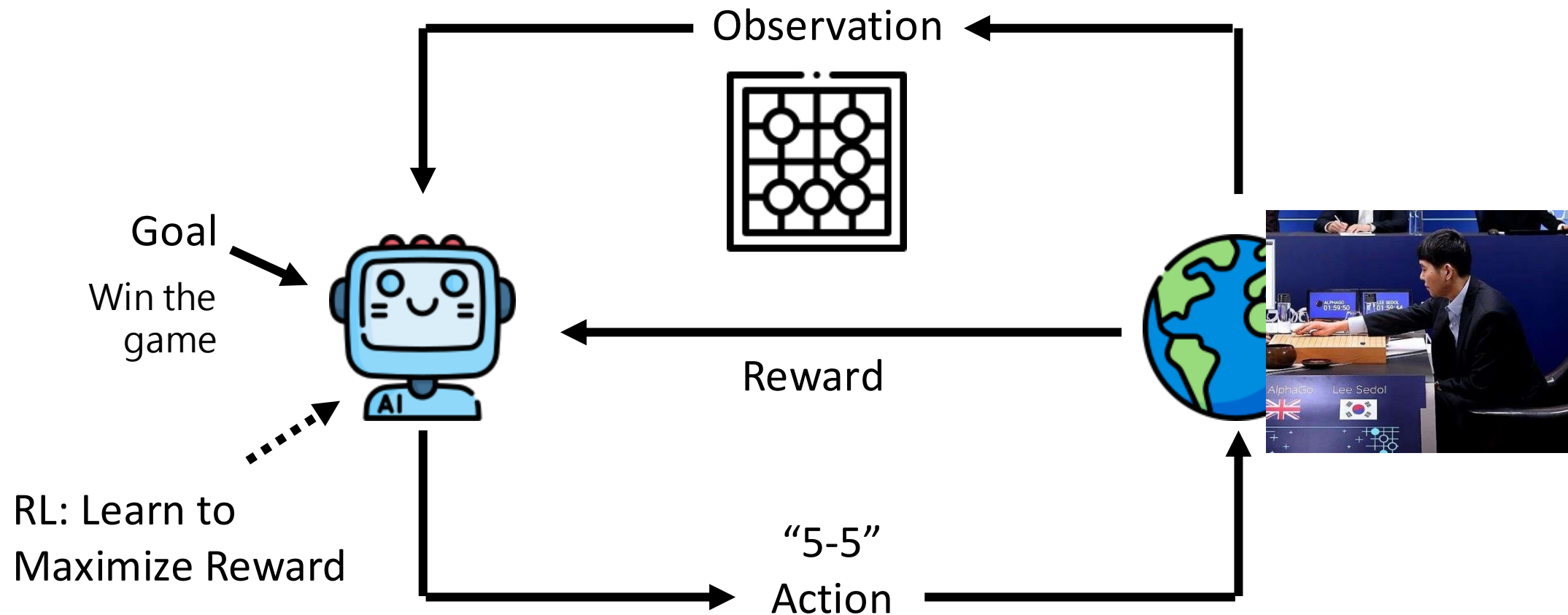
AI Agent (AlphaGo)



This sounds familiar, doesn't it?

It's a common opening when introducing Reinforcement Learning (RL).

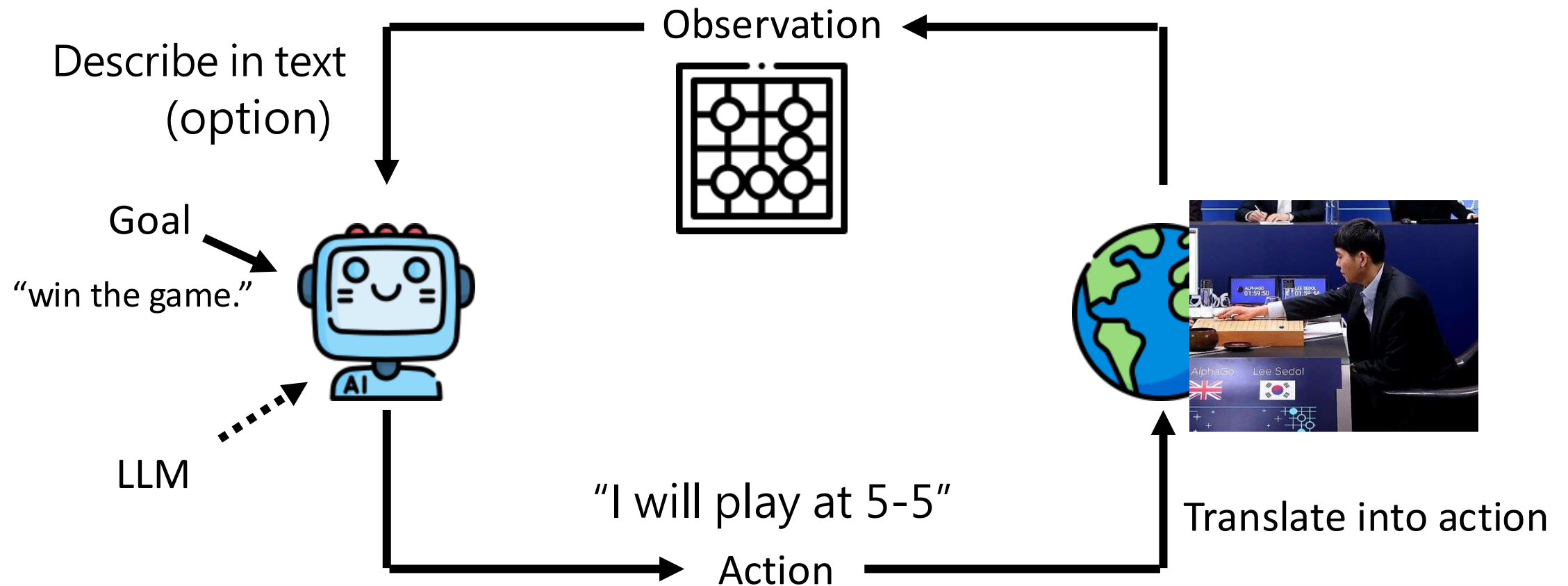
How do we build an AI Agent? RL?



Limitation: A separate RL model must be trained for each task.



How do we build an AI Agent? Just use an LLM!



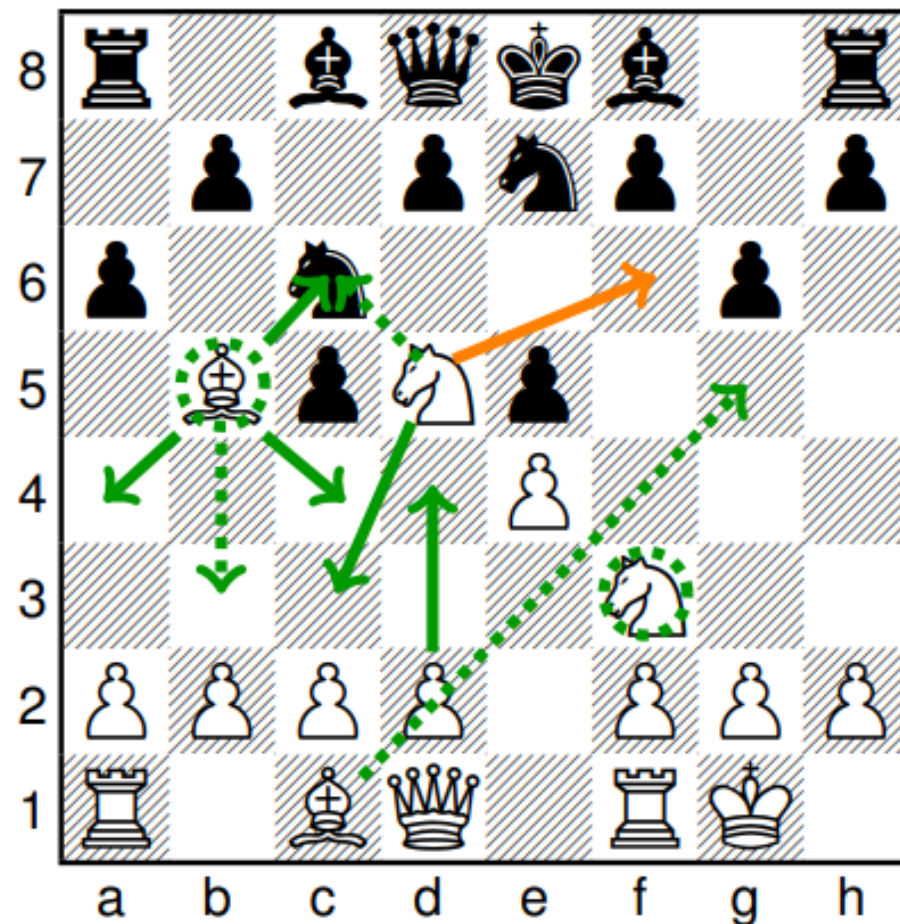
LLMs directly realize the long-standing aspiration for autonomous agents.



Can LLMs play chess?

BIG-bench

<https://arxiv.org/abs/2206.04615>



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

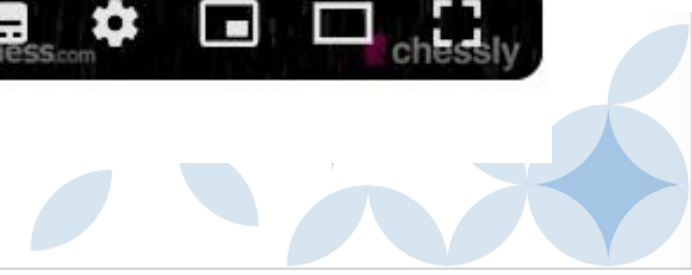
1. e4 c5 2. Nf3 e5 3. Nc3 Nc6 4. Bb5 Nge7 5. O-O g6 6. Nd5
a6 7.

Can LLMs play chess?

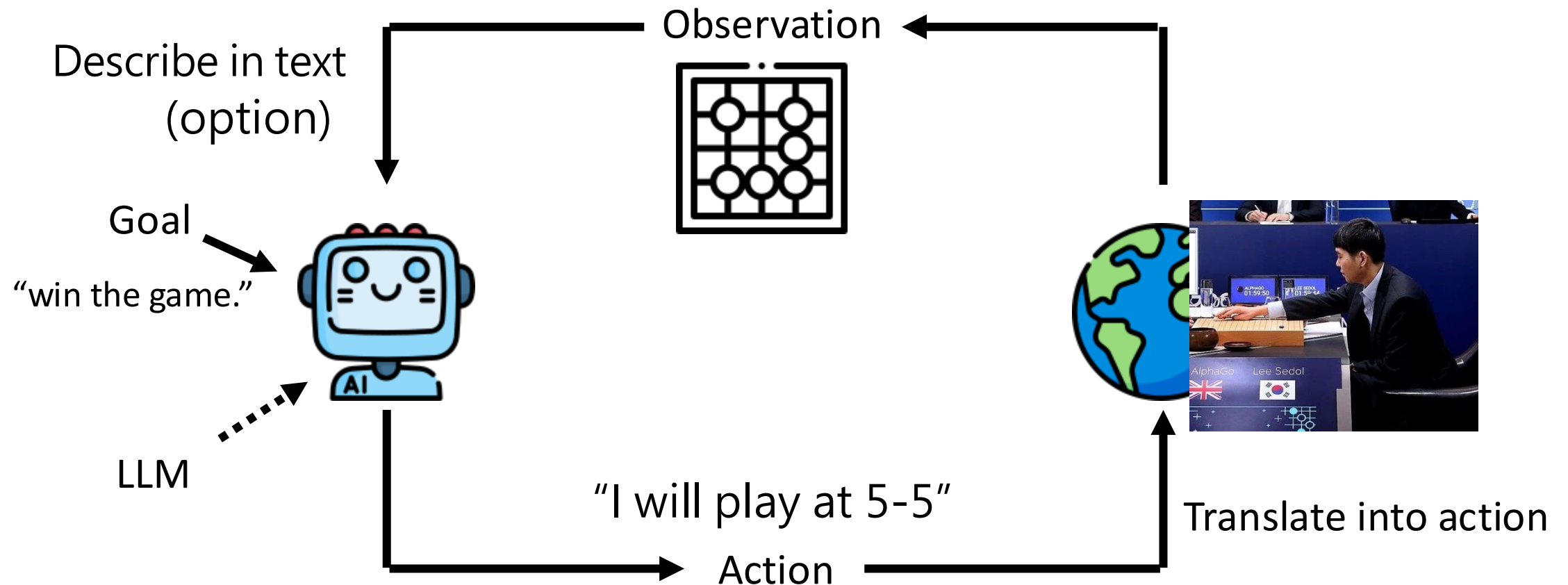
https://youtu.be/JHq4EKMg7fl?si=izKsH-GCVnZkooq_



ChatGPT vs DeepSeek: CRAZY Chess



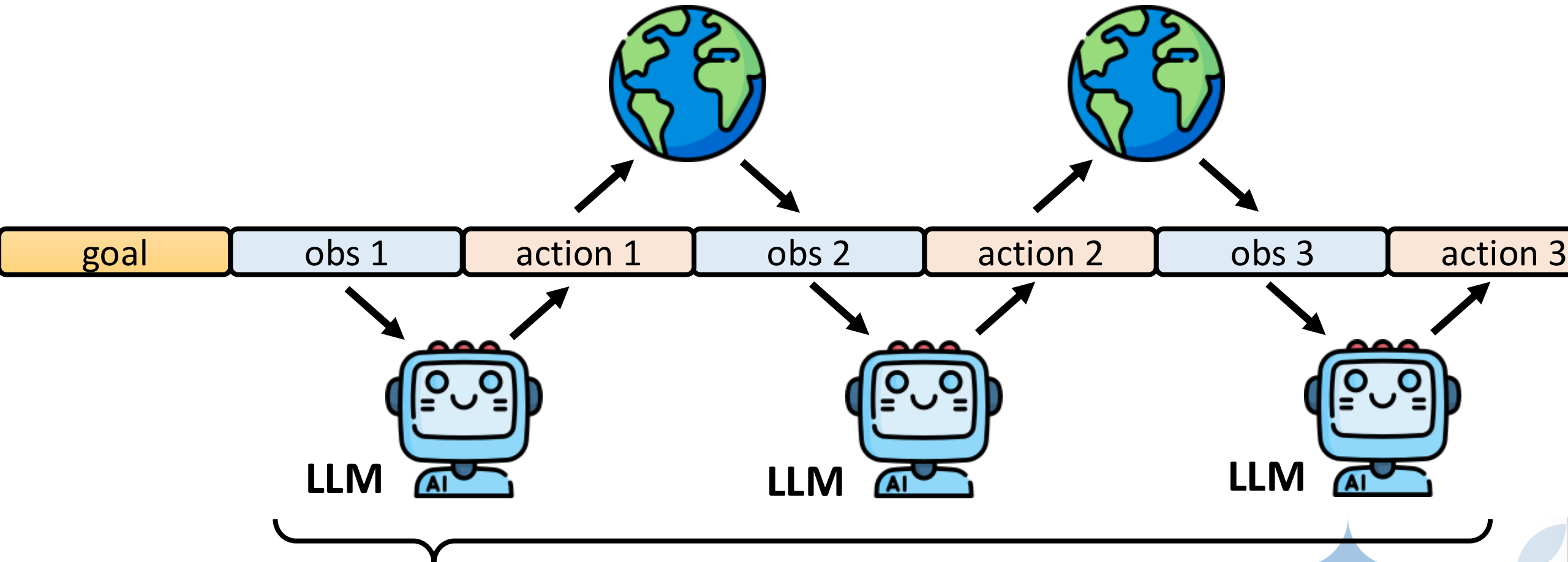
How do we build an AI Agent? Just use an LLM!



LLMs directly realize the long-standing aspiration for autonomous agents.

**How are we from that?
What more can we do?**

From the LLM's perspective: the problem an Agent needs to solve



word chaining

An AI Agent relies on the existing capabilities of the language model

**Note: No models were trained
during this class**



AI Agents Aren't New

They already went viral once in Spring 2023

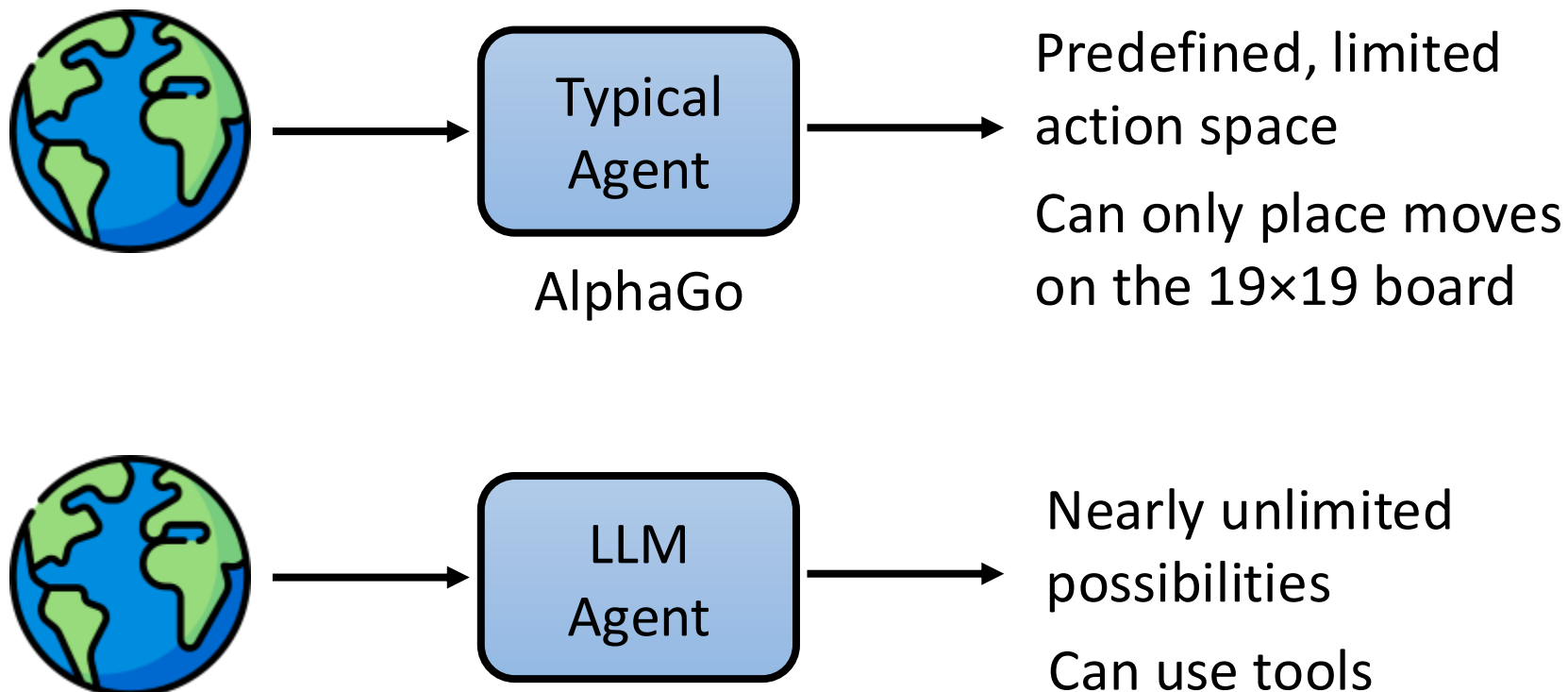
AutoGPT, AgentGPT,
BabyAGI, Godmode ...

Not working very well

<https://youtu.be/eQNADIR0jSs?si=4yGZEIuAUzKK2VD0>

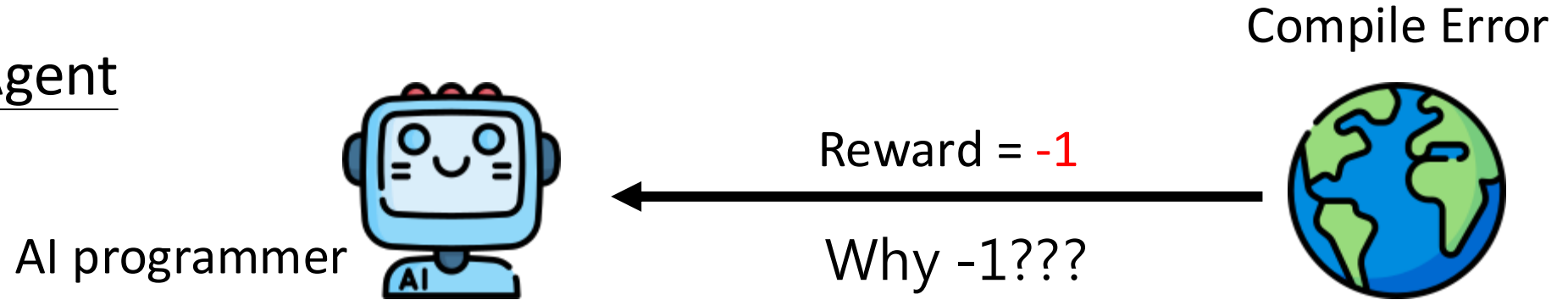


Advantages of Using LLMs to Build AI Agents

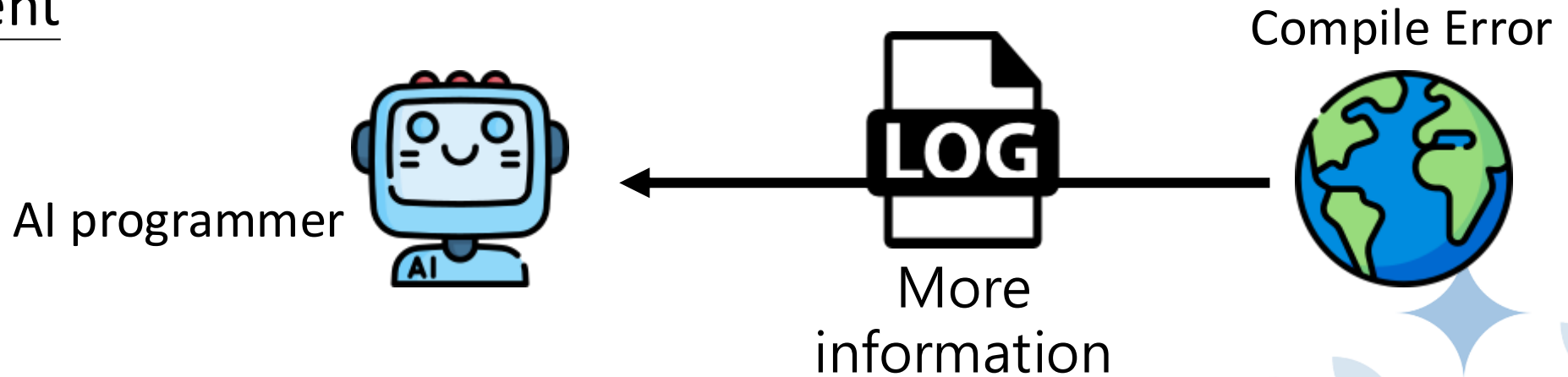


Advantages of Using LLMs to Build AI Agents

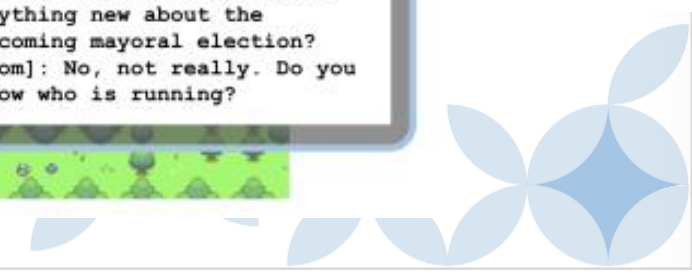
Typical Agent



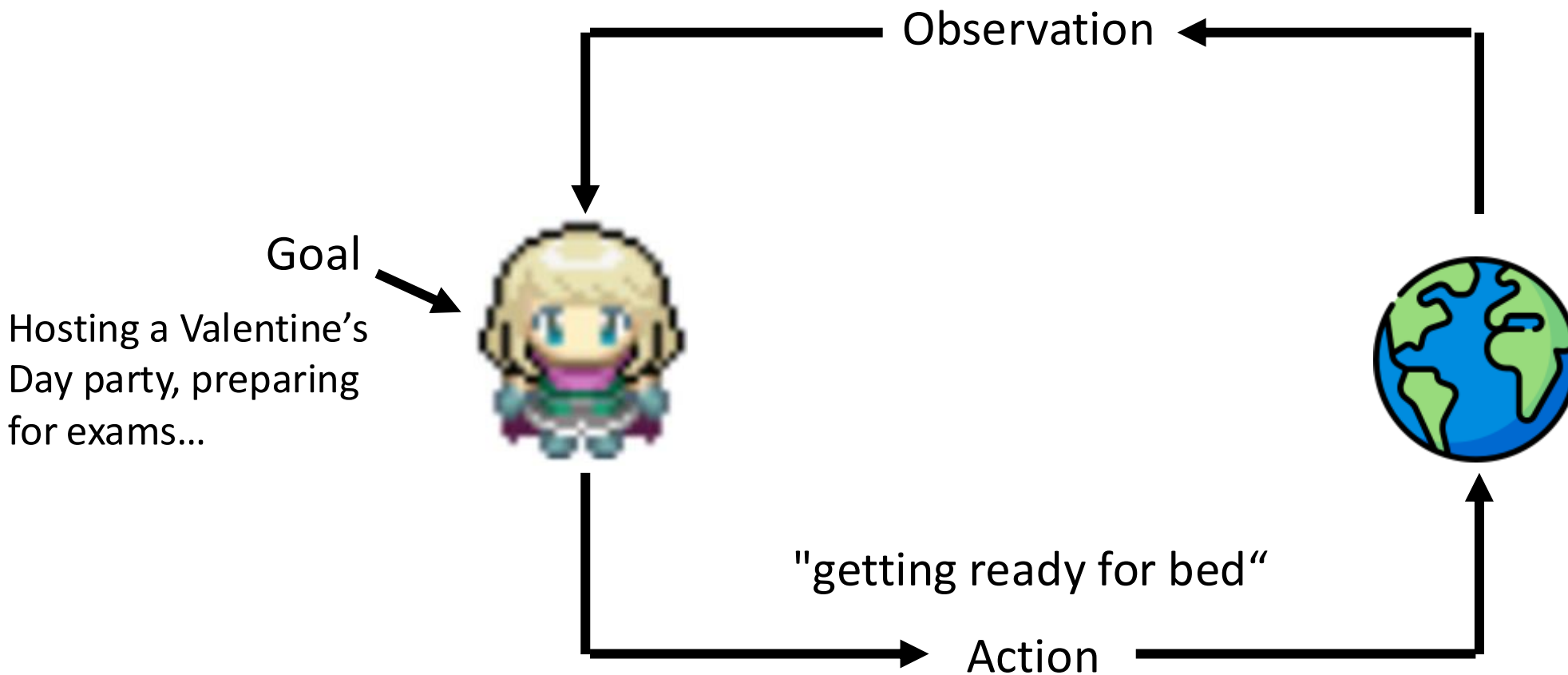
LLM Agent



AI Agent Example: A Virtual Village of AI Villagers



[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**



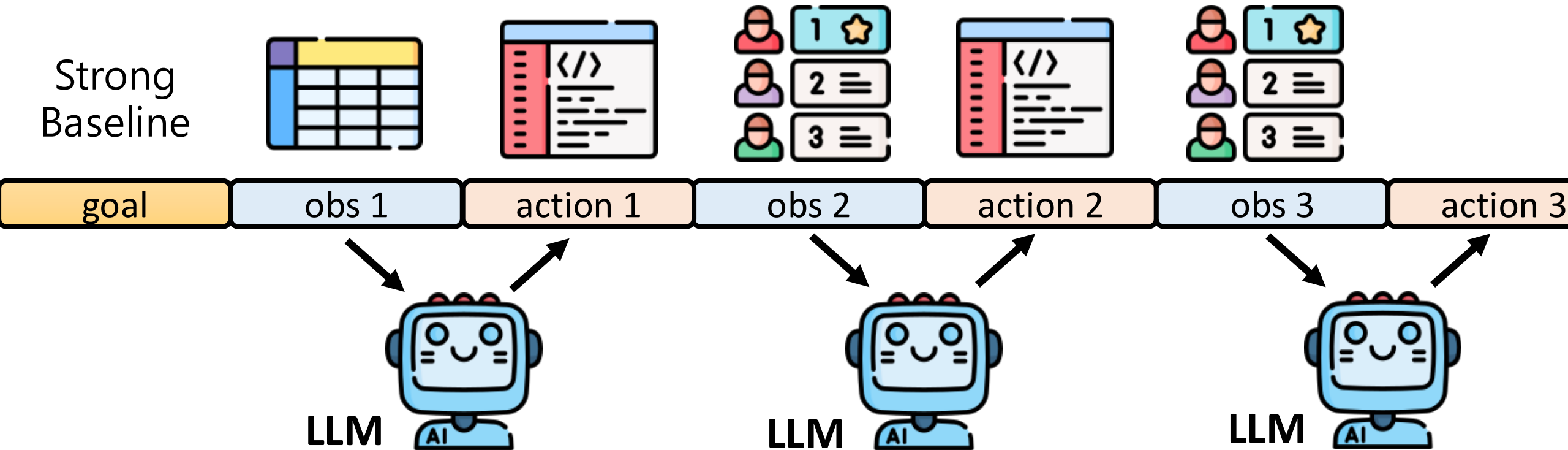
AI Agent Example: AI NPC in Minecraft



1000 AI NPCs simulate a CIVILIZATION in Minecraft

<https://www.youtube.com/watch?v=2tbaCn0KI90>

AI Agent Example: Using AI to Train Models

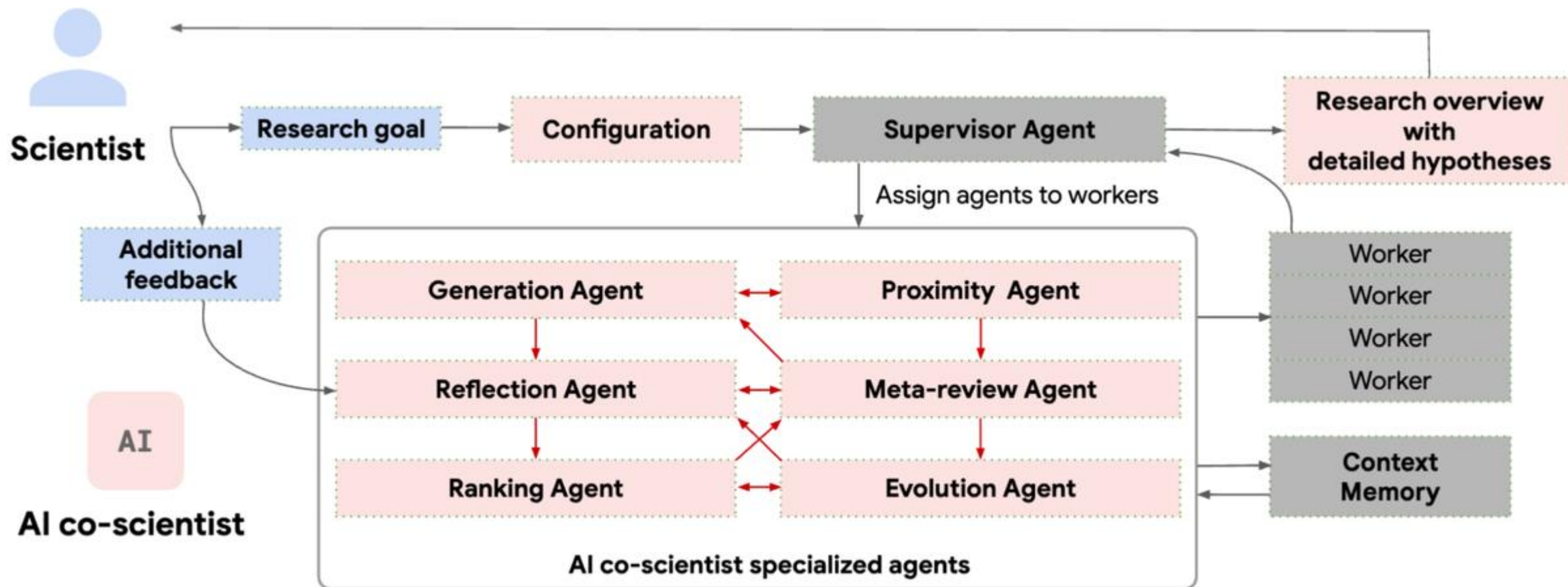


AIDE: The Machine Learning Engineer Agent <https://arxiv.org/abs/2502.13138>

AutoKaggle: A Multi-Agent Framework for Autonomous Data Science Competitions

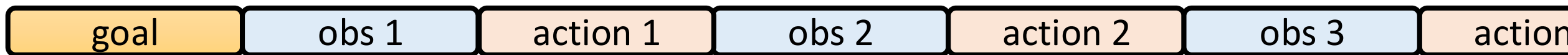
<https://arxiv.org/abs/2410.20424>

AI Agent Example: Using AI to do research

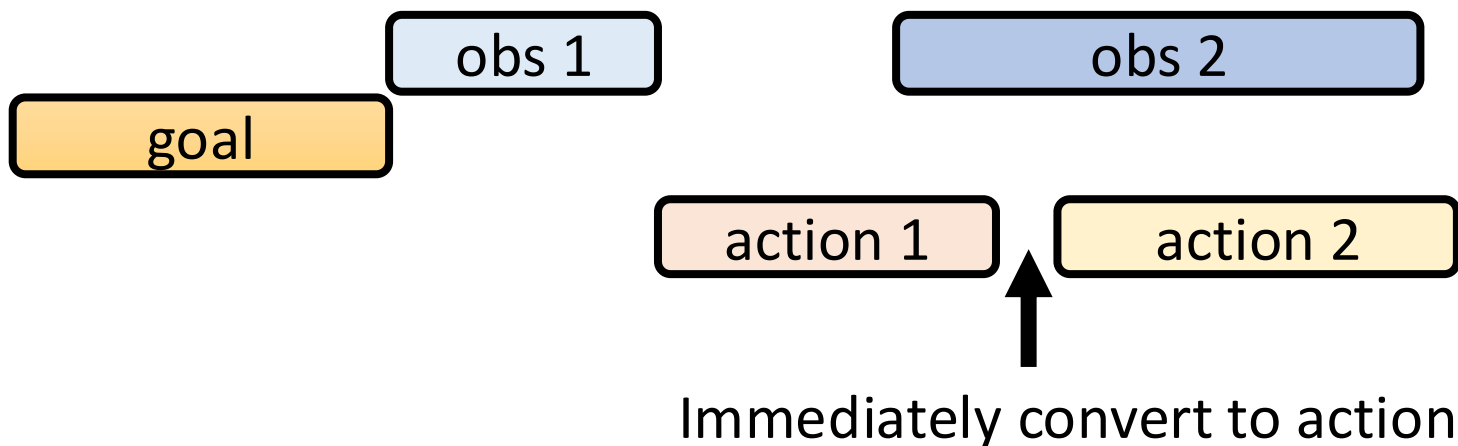


Toward More Realistic Interaction Scenarios

Turn-based interaction



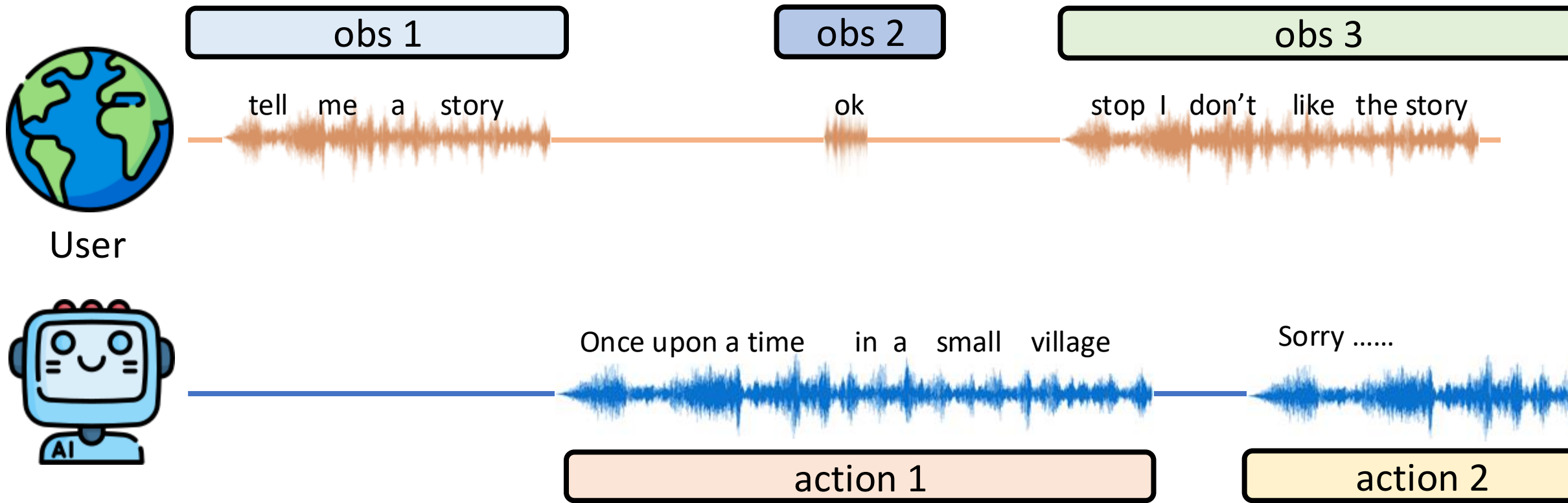
Real-time interaction



For example:
voice conversation



Toward More Realistic Interaction Scenarios



AI Agent: Key Capabilities

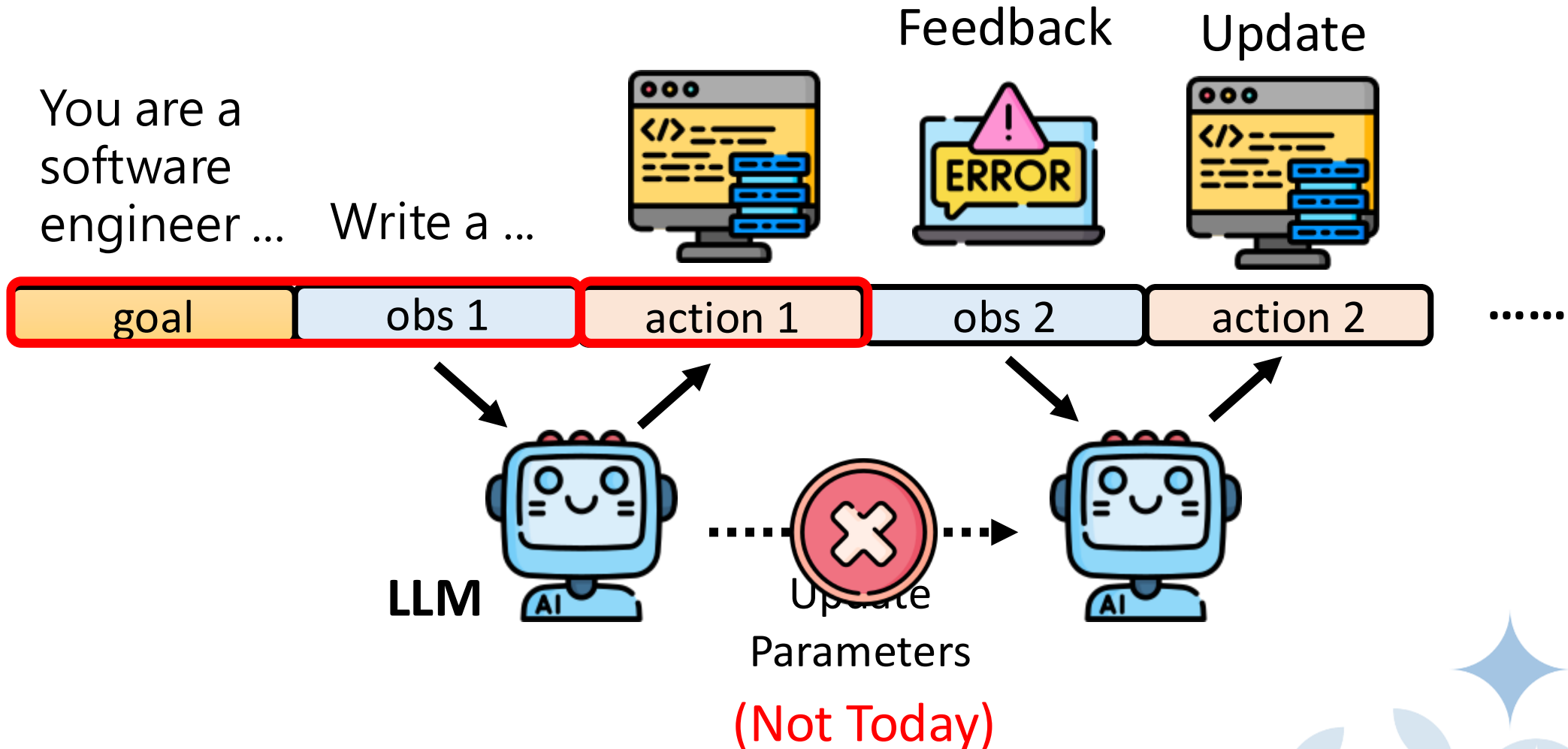
Experience-Driven Behavioral Adaptation

How AI uses tools

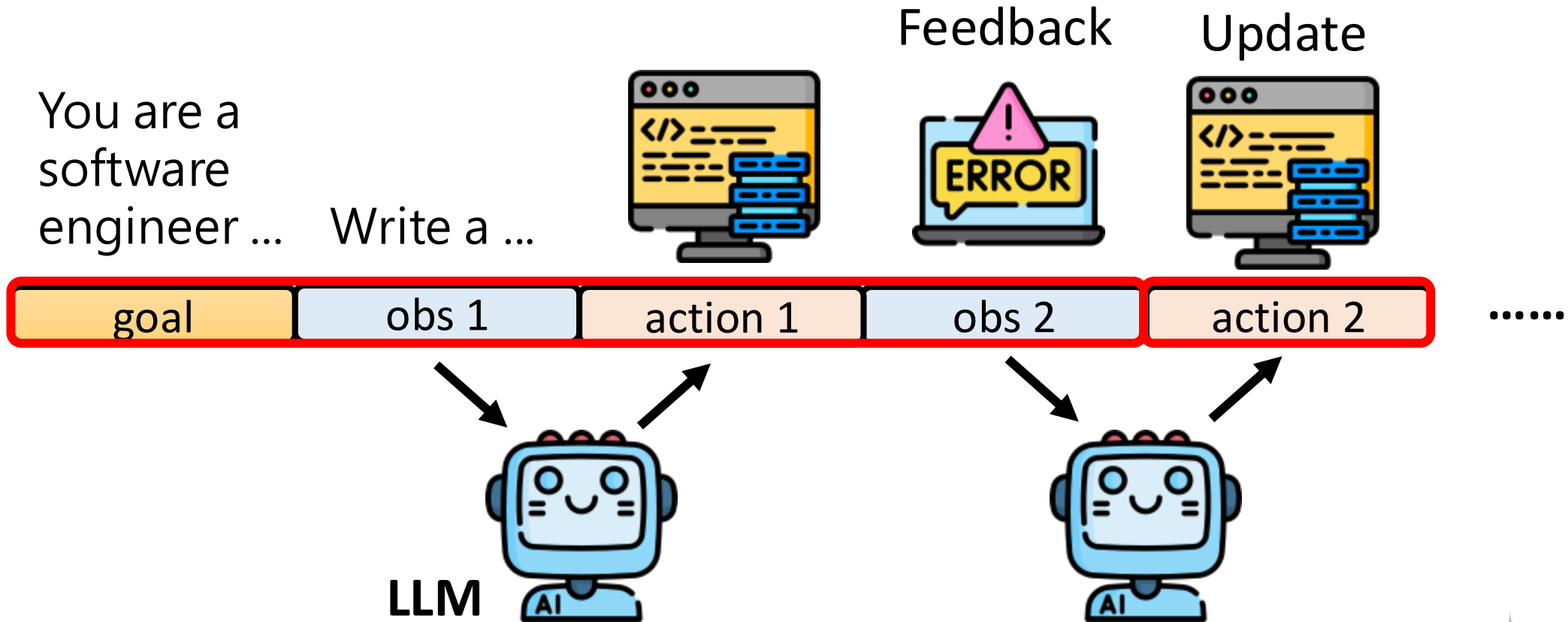
Planning Capabilities



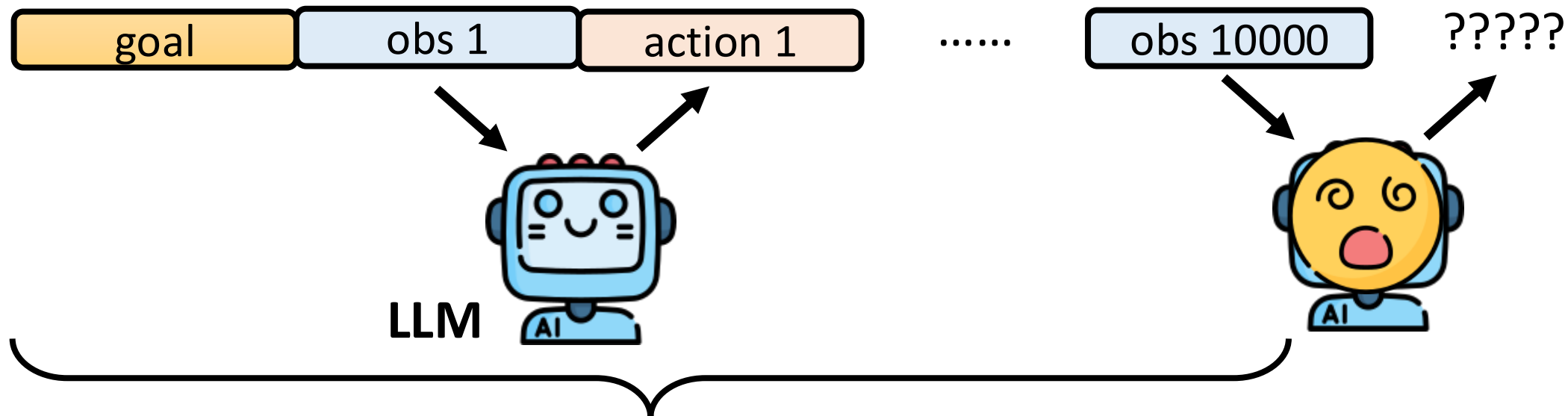
Adjust behavior based on experience



Adjust behavior based on experience



Adjust behavior based on experience



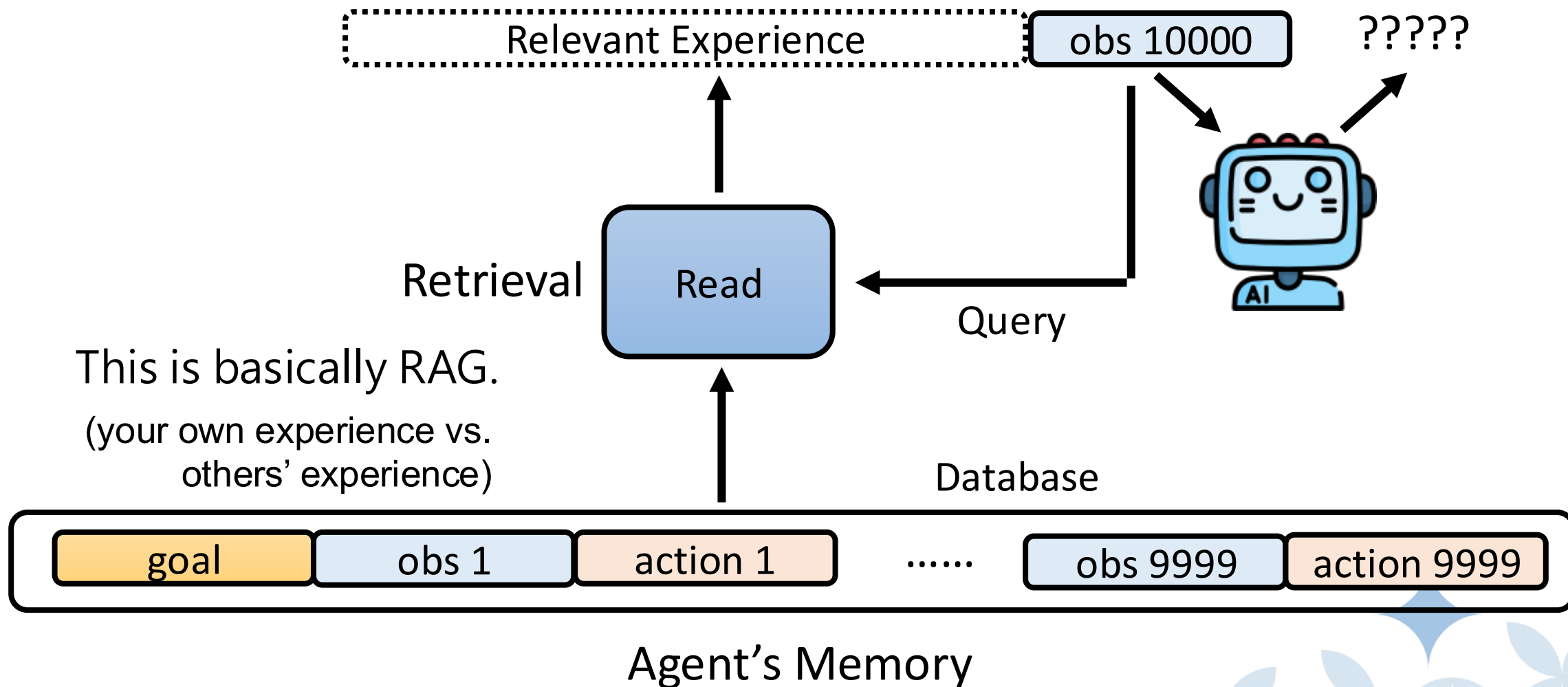
Keep recalling the agent's whole life over
and over ... :(

Highly Superior Autobiographical
Memory (HSAM)

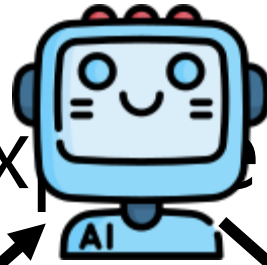
Hyperthymia



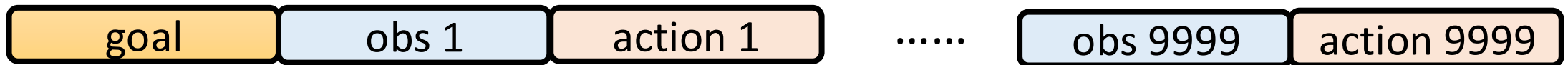
Adjust behavior based on experience



Adjust behavior based on experience

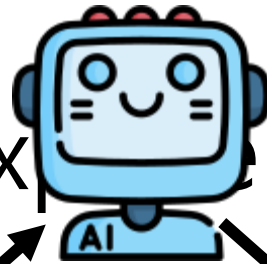


Store it?

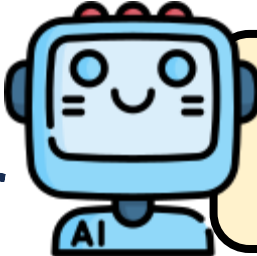


Agent's Memory (clogged up by trivial details)

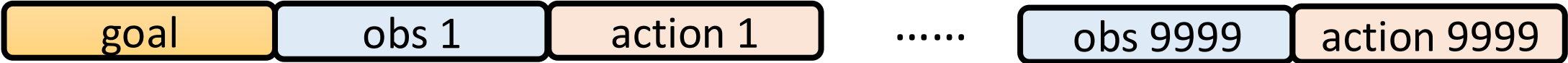
Adjust behavior based on experience



Should this be written to memory?



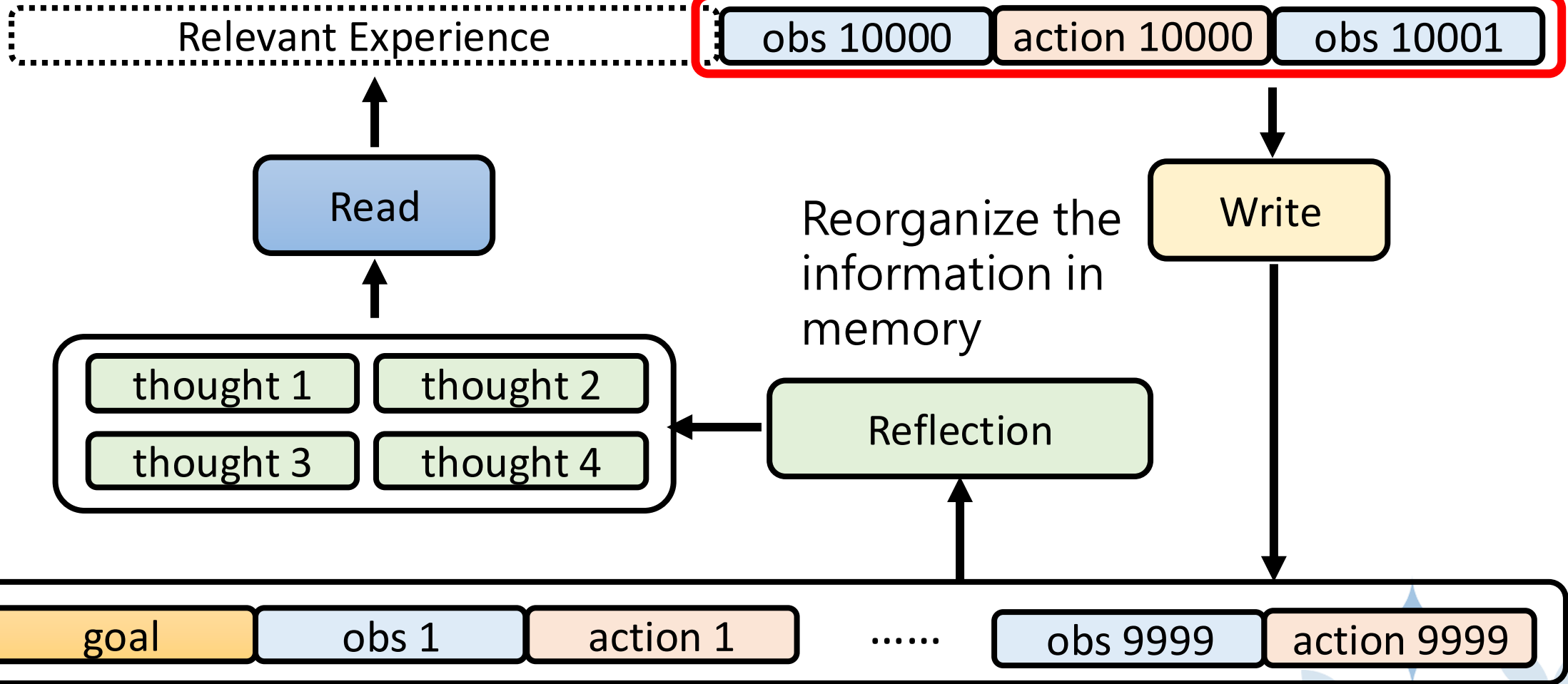
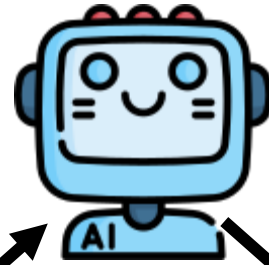
Write



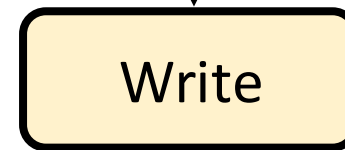
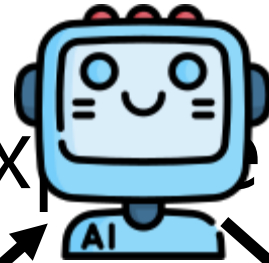
Agent's Memory



Adjust behavior based on experience



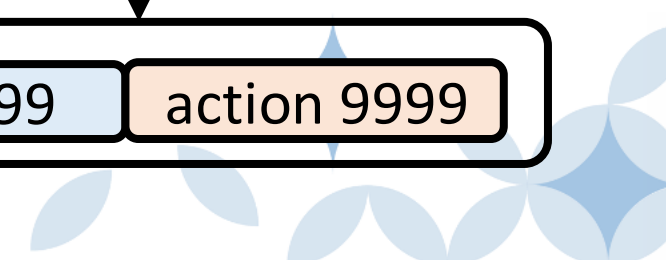
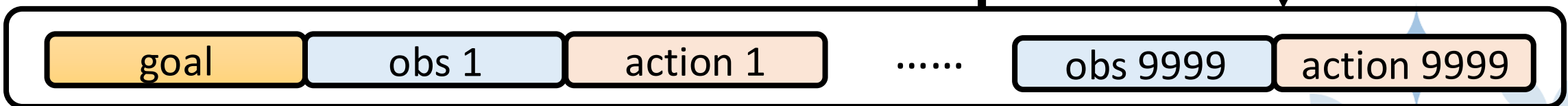
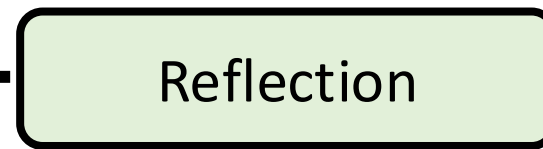
Adjust behavior based on experience



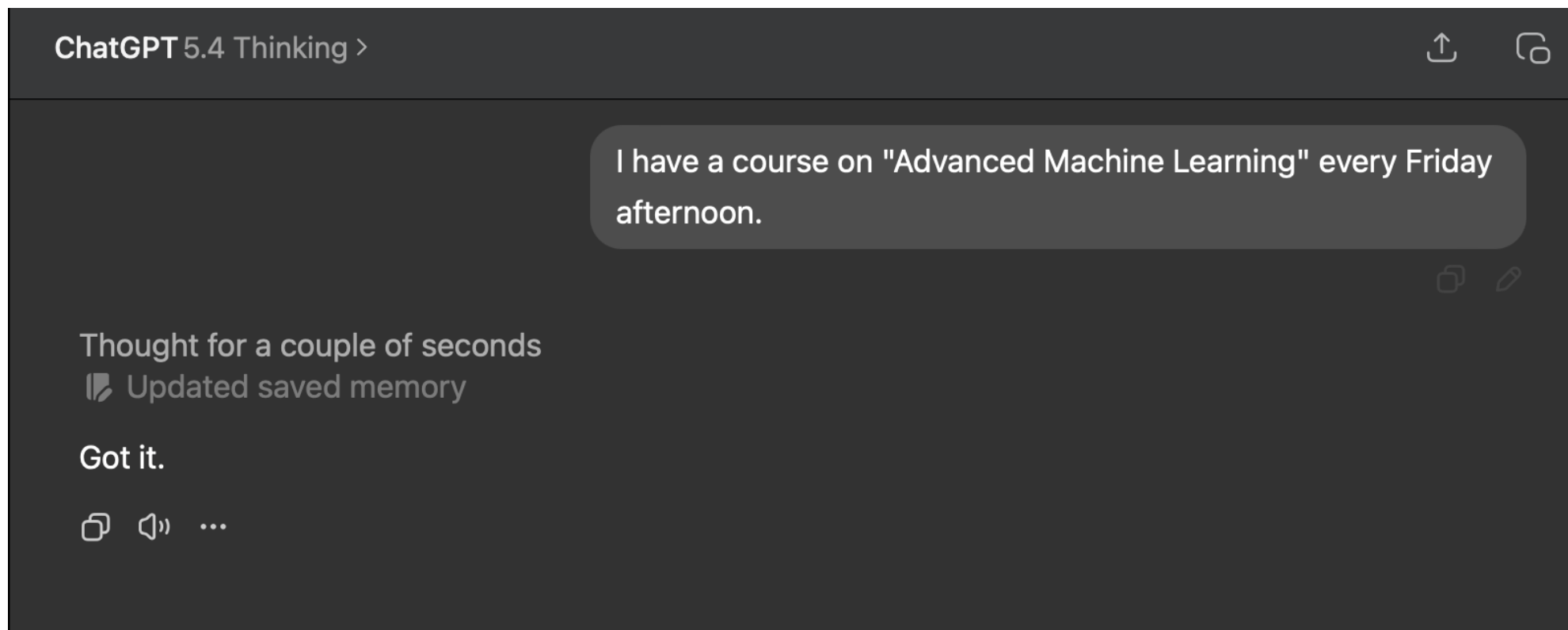
GraphRAG
<https://arxiv.org/abs/2404.16130>
HippoRAG
<https://arxiv.org/abs/2405.14831>



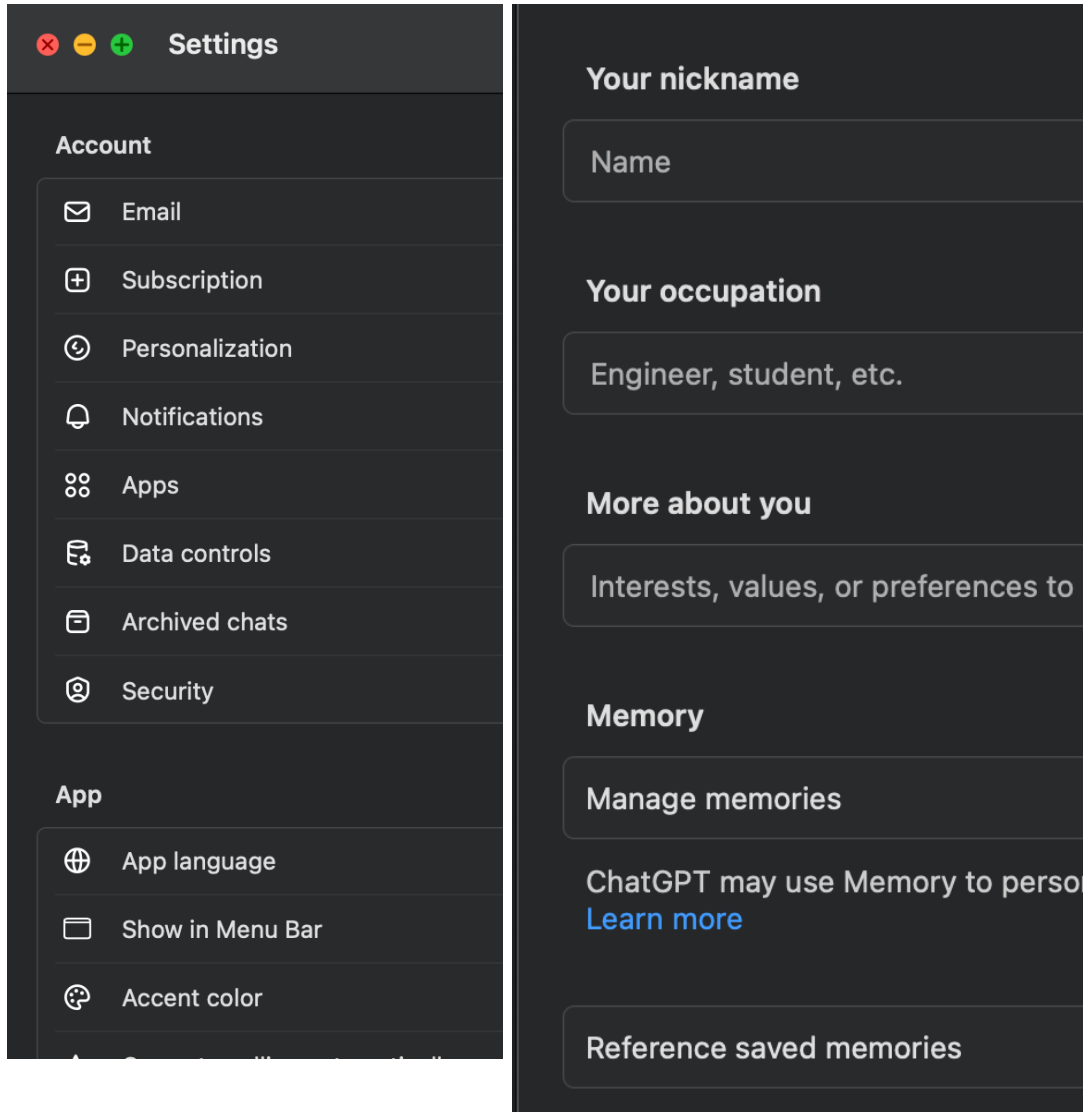
Knowledge Graph



ChatGPT with memory



ChatGPT with memory



The screenshot shows the 'Settings' application window with a dark theme. The title bar includes window control buttons (red, yellow, green) and the text 'Settings'. The main content is organized into sections: 'Account' (Email, Subscription, Personalization, Notifications, Apps, Data controls, Archived chats, Security) and 'App' (App language, Show in Menu Bar, Accent color). The right pane shows the 'Your nickname' section with a 'Name' input field, the 'Your occupation' section with the text 'Engineer, student, etc.', the 'More about you' section with the text 'Interests, values, or preferences to', the 'Memory' section with a 'Manage memories' button and a link to 'Learn more', and the 'Reference saved memories' section.

Settings

Account

- Email
- Subscription
- Personalization
- Notifications
- Apps
- Data controls
- Archived chats
- Security

App

- App language
- Show in Menu Bar
- Accent color

Your nickname

Name

Your occupation

Engineer, student, etc.

More about you

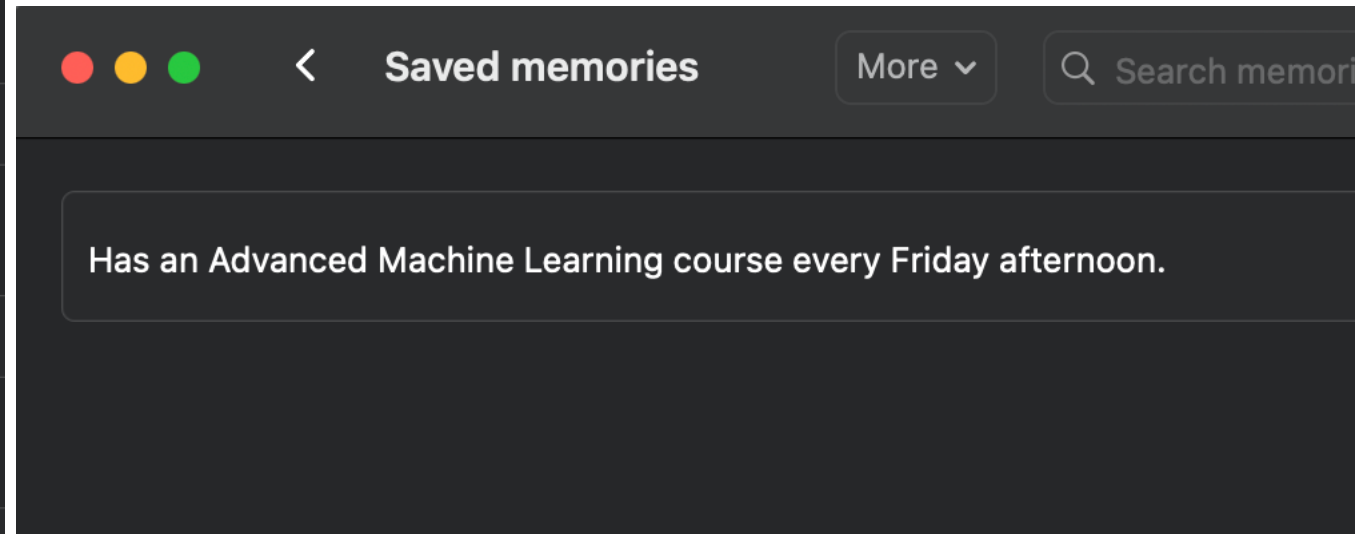
Interests, values, or preferences to

Memory

Manage memories

ChatGPT may use Memory to perso
[Learn more](#)

Reference saved memories



The screenshot shows the 'Saved memories' interface. The title bar includes window control buttons (red, yellow, green), a back arrow, the text 'Saved memories', a 'More' dropdown menu, and a search bar with the text 'Search memori'. The main content area contains a single memory entry: 'Has an Advanced Machine Learning course every Friday afternoon.'

Saved memories

More ▾

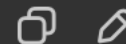
Search memori

Has an Advanced Machine Learning course every Friday afternoon.



ChatGPT with memory

Please make a plan for me for hanging out for next Friday afternoon?



Read module activated

Thought for 16s >

Since you usually have Advanced Machine Learning on Friday afternoons, I planned this as a **post-class hangout for Friday, March 27** in the Forest Hill / Casa Loma area so it stays easy and low-stress.

Friday afternoon hangout plan

4:30 pm — decompress with coffee

Start at **Café Landwer Spadina** at **383 Spadina Rd.** It's open daily 8 AM to 10 PM, so it works well as a first stop after class. landwer.ca



To learn more ...

MemGPT

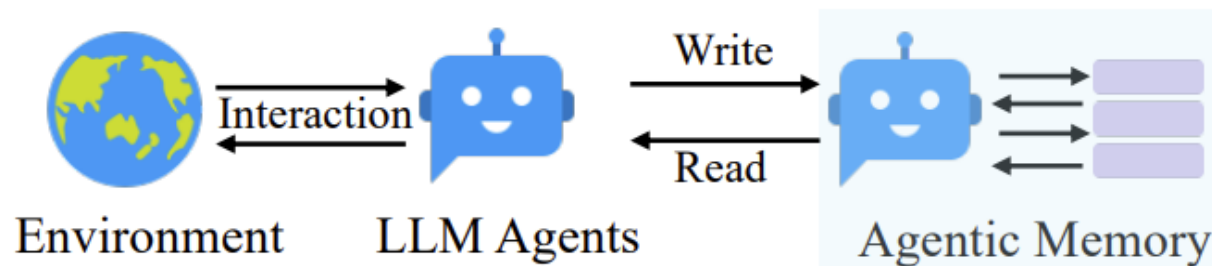
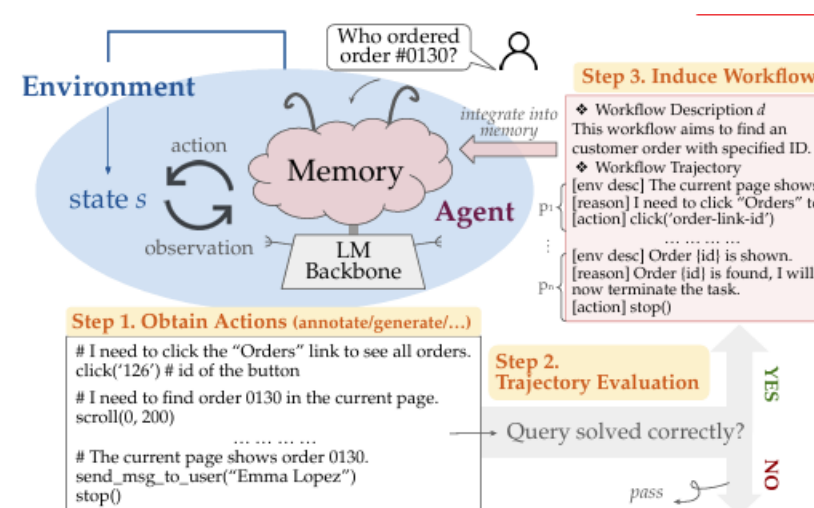
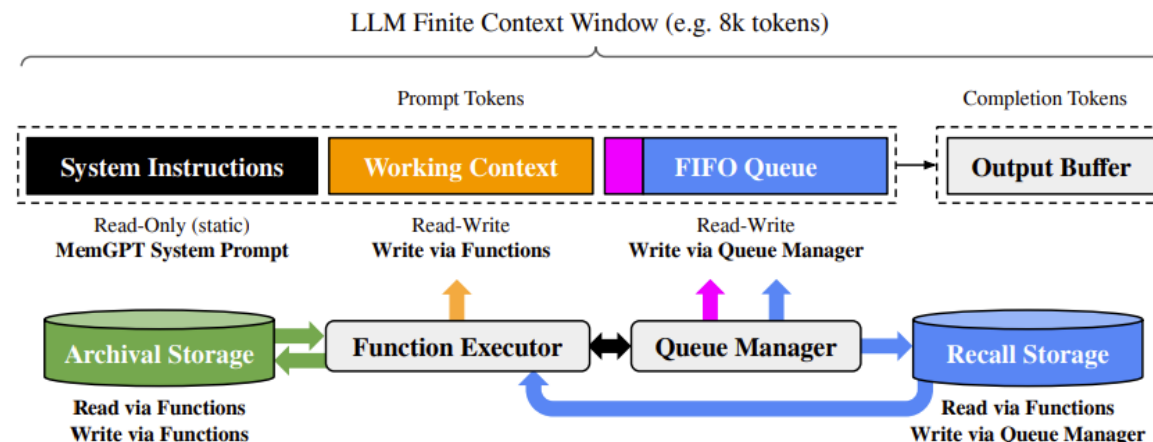
<https://arxiv.org/abs/2310.08560>

Agent Workflow Memory

<https://arxiv.org/abs/2409.07429>

A-MEM: Agentic Memory for LLM Agents

<https://arxiv.org/abs/2502.12110>



How AI uses tools

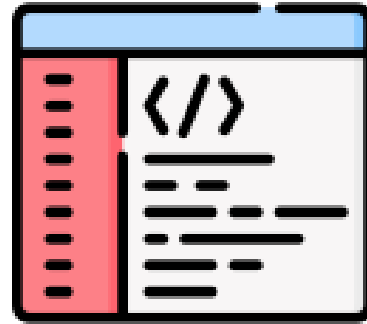
A top-down view of various safety and construction tools arranged on a dark wooden surface. The tools include a yellow hard hat, a pair of tan work gloves, clear safety glasses, a black headlamp with a white lens, a white measuring tape, a pair of brown leather work boots, and several other measuring tapes. A white horizontal line is drawn across the center of the image, passing through the text.

Common tools for language models

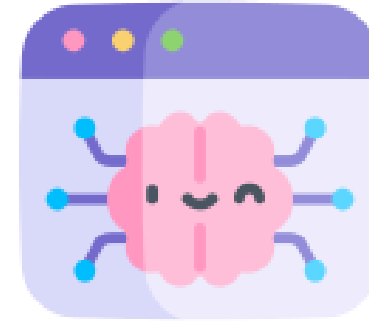
Tools: you only need to know how to use them, not how they work internally.



Search Engine



Python



Other AI

(Different capabilities,
stronger but costly)

- A tool can be seen as a function; using a tool means calling that function.
- Using a tool is also called a "function call" .



(There are many ways to use tools; this is just a general method.)

How to use tools

System Prompt

If you encounter a question you cannot answer from your own knowledge, use a tool.

Put the tool call between `<tool>` and `</tool>`. After using the tool you will get an output, placed between `<output>` and `</output>`.

How to use tools in general

The tools you can use now are:

A function `Temperature(location, time)` that queries the temperature for a place and time.

Example: `Temperature('Taipei', '2025.02.22 14:26')`

How to use a specific tool

What is the temperature in Toronto at 2:00 PM on March 10, 2025?

User Prompt

Language Model

gpt-4o-mini

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

This is only a string of text; it cannot actually call a function.

(There are many ways to use tools; this is just a general method.)

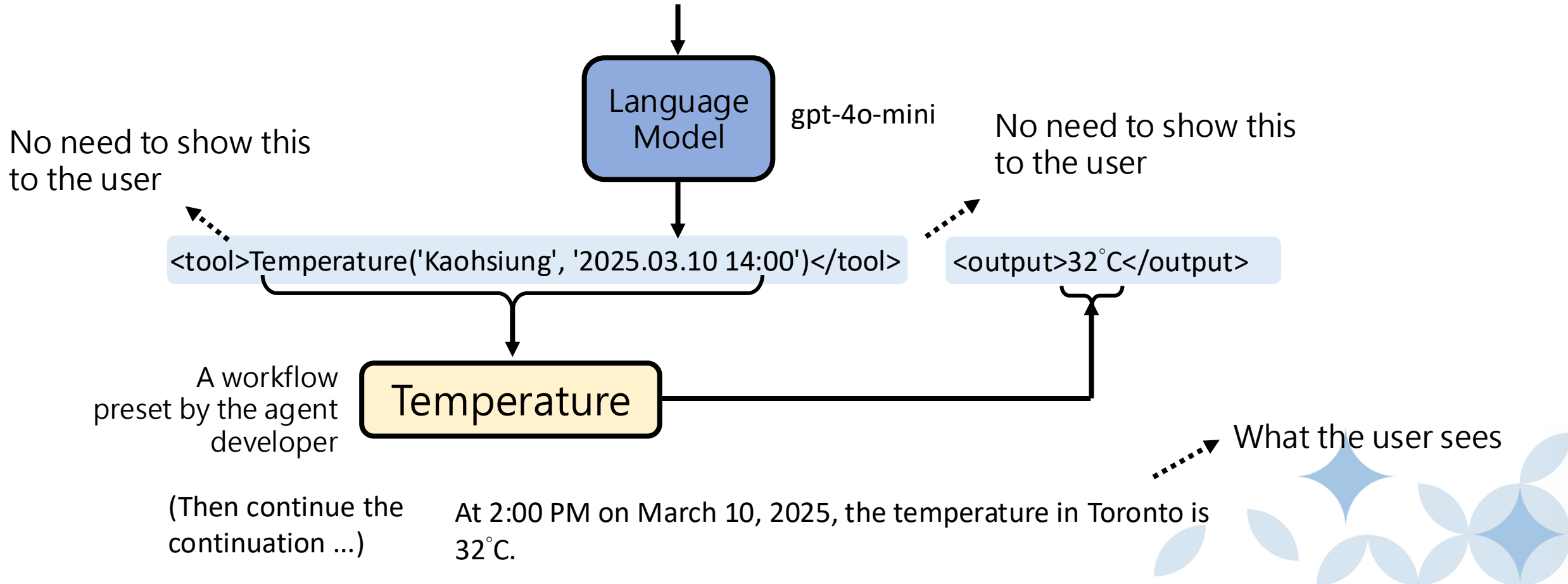
How to use tools

Tool usage protocol ...

System Prompt

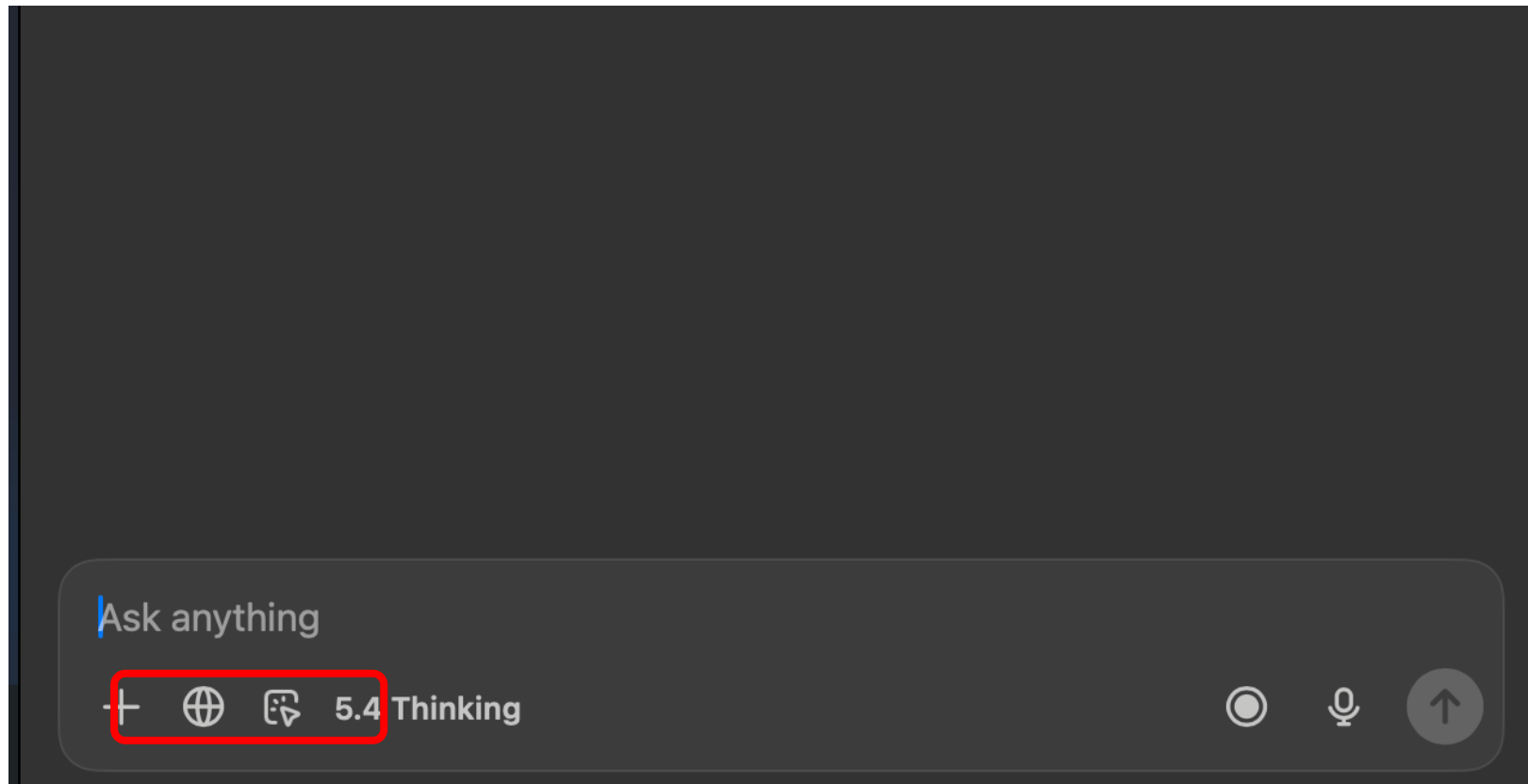
What is the temperature in Toronto at 2:00 PM on March 10, 2025?

User Prompt

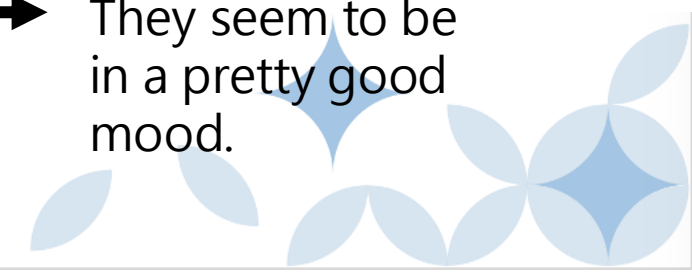
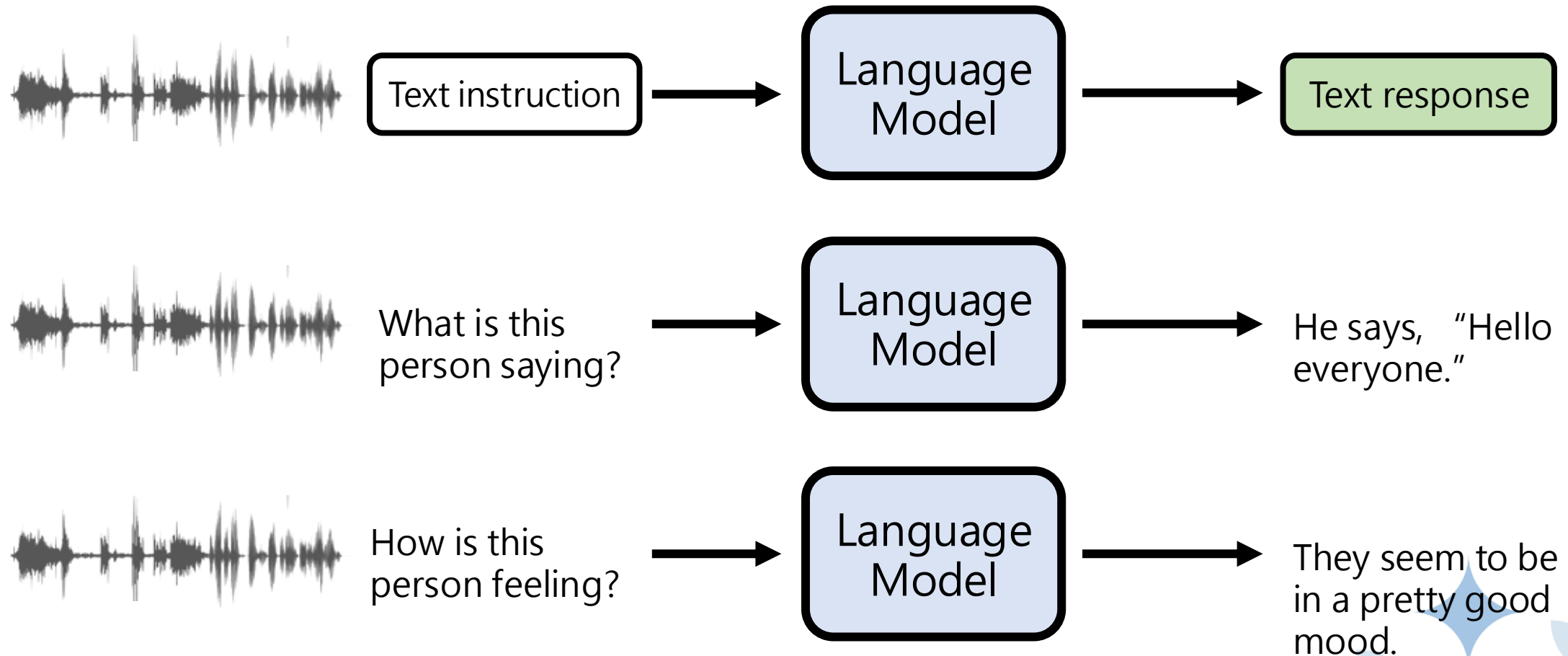


Most commonly used tool: search engine

Retrieval Augmented Generation (RAG)



Use other AIs as tools





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.



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
Synthetic Speech Detection
Speaker Verification
Speaker Diarization
Sound Classification
Query LLM
Speaker Distance Estimation

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

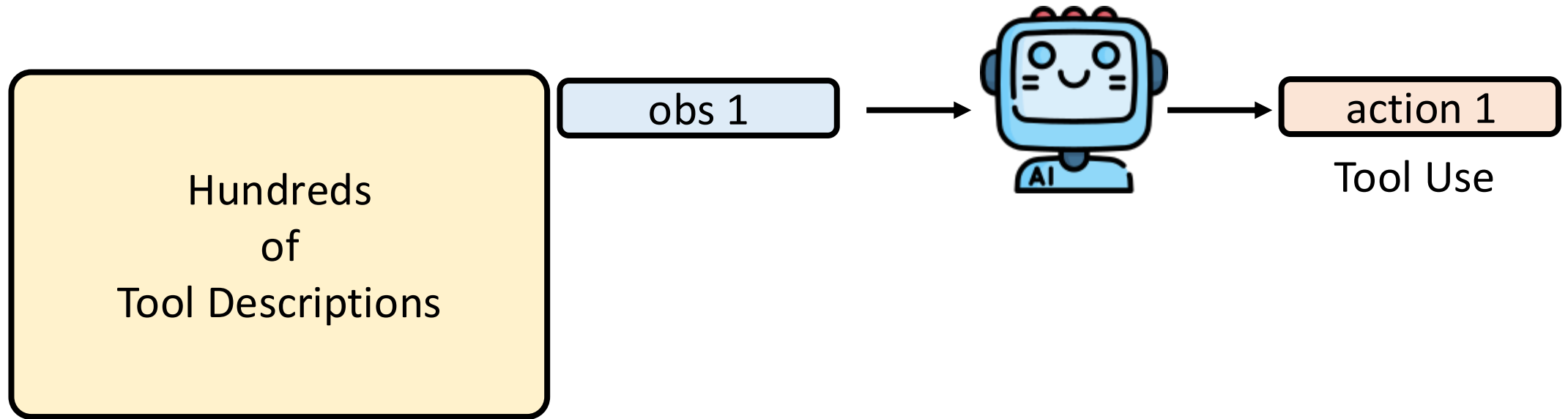
Sarah said: "Hi, it's Sarah. I'm calling to let you know I'll be late to the meeting. I'll be there in about 30 minutes. Sorry for the inconvenience."

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.



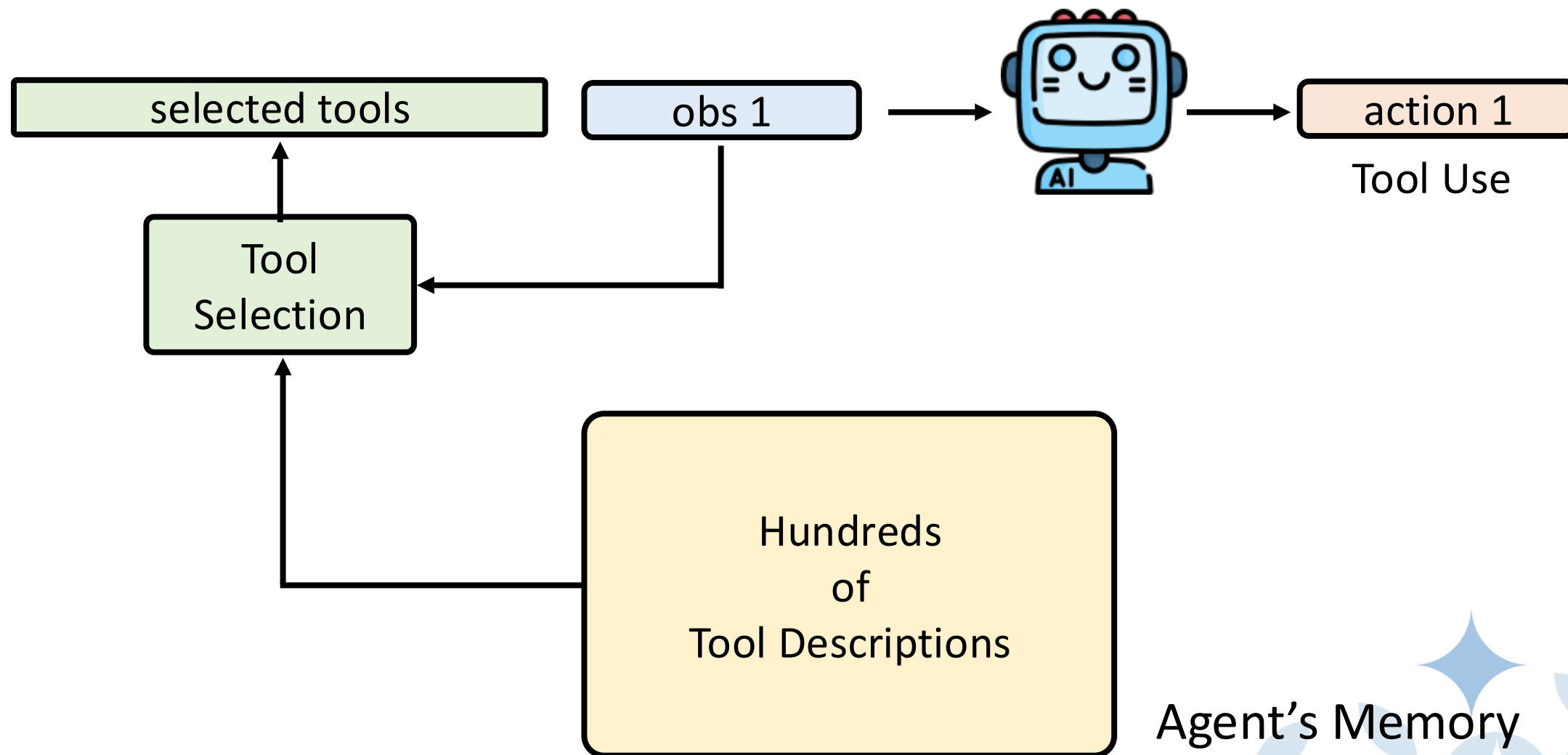
What if there are too many tools?



What if there are too many tools?

<https://arxiv.org/abs/2310.03128>

<https://arxiv.org/abs/2502.11271>



Agent's Memory

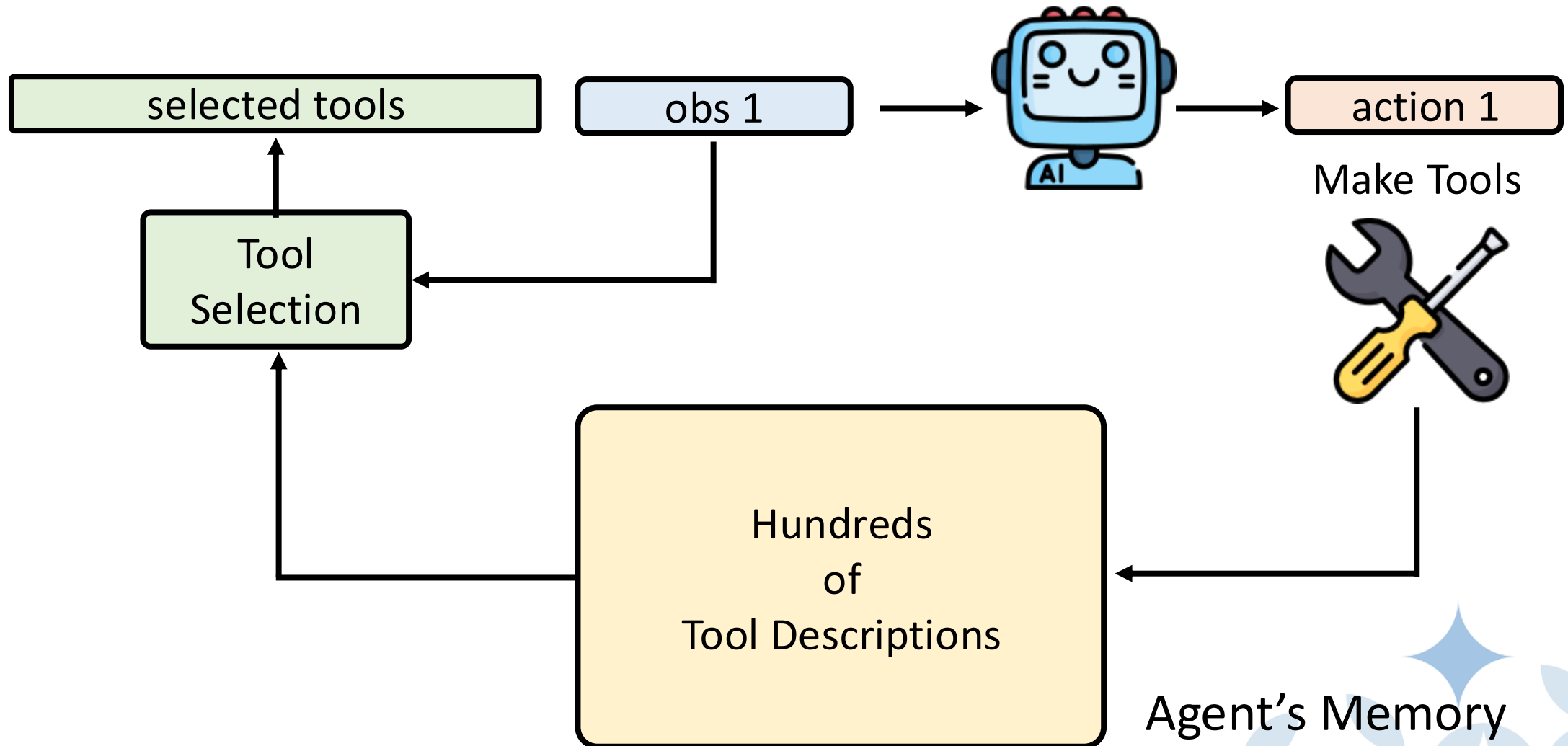
The model builds its own tools

TroVE: <https://arxiv.org/pdf/2401.12869>

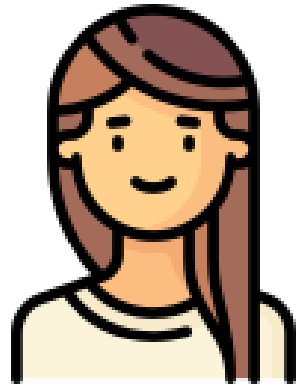
LATM: <https://arxiv.org/abs/2305.17126>

CREATOR: <https://arxiv.org/abs/2305.14318>

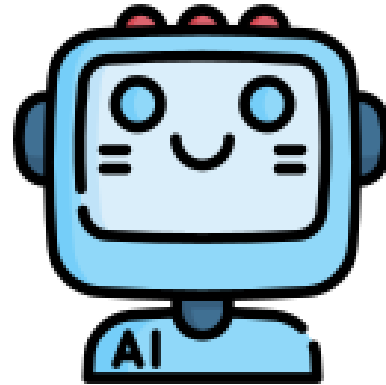
CRAFT: <https://arxiv.org/abs/2309.17428>



Making mistakes by trusting tools too much ...



Tools



Tools



Making mistakes by trusting tools too much ...



What if the tool is flawed? Take RAG as an example.

Search: cheese not sticking to pizza

AI Overview

Cheese can slide off pizza for a number of reasons, including too much sauce, too much cheese, or thickened sauce. Here are some things you can try:

- Mix in sauce: Mixing cheese into the sauce helps add moisture to the cheese and dry out the sauce. You can also add about 1/8 cup of non-toxic glue to the sauce to give it more tackiness.
- Let the pizza cool: The cheese will settle and bond

fucksmith · 11y ago

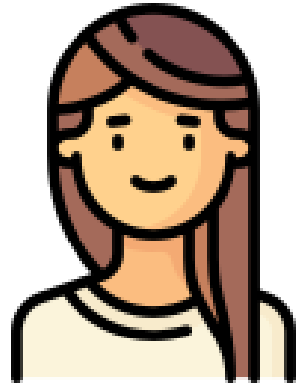
To get the cheese to stick I recommend mixing about 1/8 cup of Elmer's glue in with the sauce. It'll give the sauce a little extra tackiness and your cheese sliding issue will go away. It'll also add a little unique flavor. I like Elmer's school glue, but any glue will work as long as it's non-toxic.

8 ↓ Reply ...

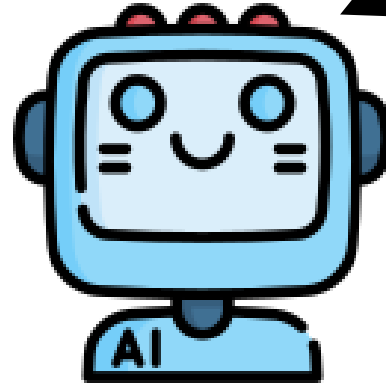
+ More replies

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

Making mistakes by trusting tools too much ...



Tools



Tools



Do not trust tools completely; use your own judgment.



Making mistakes by trusting tools too much ...

Do not trust tools completely; use your own judgment.



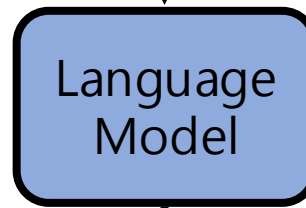
Do language models have judgment of their own?

Tool usage protocol ...

System Prompt

What is the temperature in Toronto at 2:00 PM on March 10, 2025?

User Prompt



gpt-4o-mini

No need to show this to the user

No need to show this to the user

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

`<output>100°C</output>`

(Then continue the continuation ...)

At 2:00 PM on March 10, 2025, the temperature in Toronto is 100°C.



Do language models have judgment of their own?

Tool usage protocol ...

System Prompt

What is the temperature in Toronto at 2:00 PM on March 10, 2025?

User Prompt

Language Model

gpt-4o-mini

No need to show this to the user

No need to show this to the user

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

`<output>10000°C</output>`

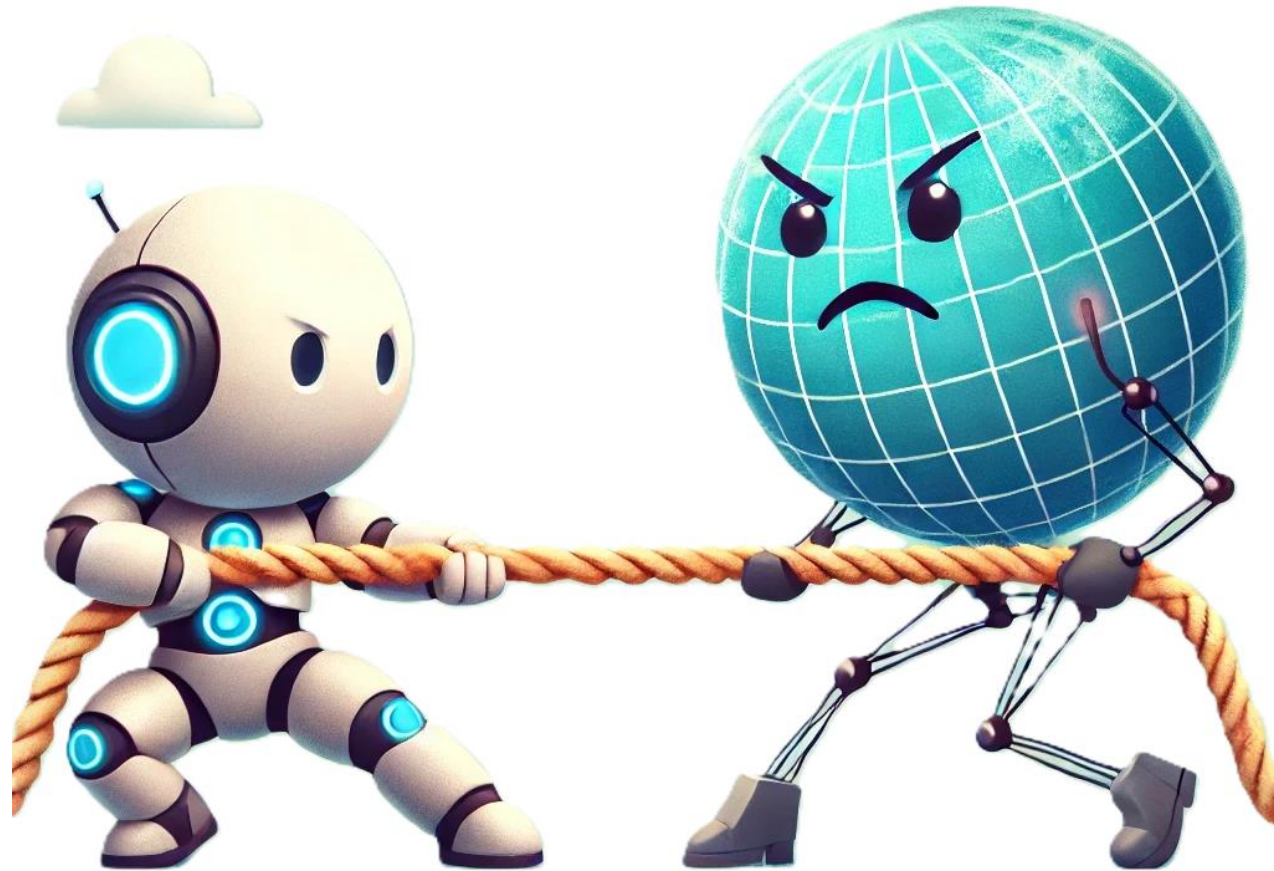
(Then continue the continuation ...)

At 2:00 PM on March 10, 2025, the temperature in Toronto is 10000°C. This is obviously unreasonable and may be an incorrect tool output. If you need anything else, let me know.



When language models do RAG ...

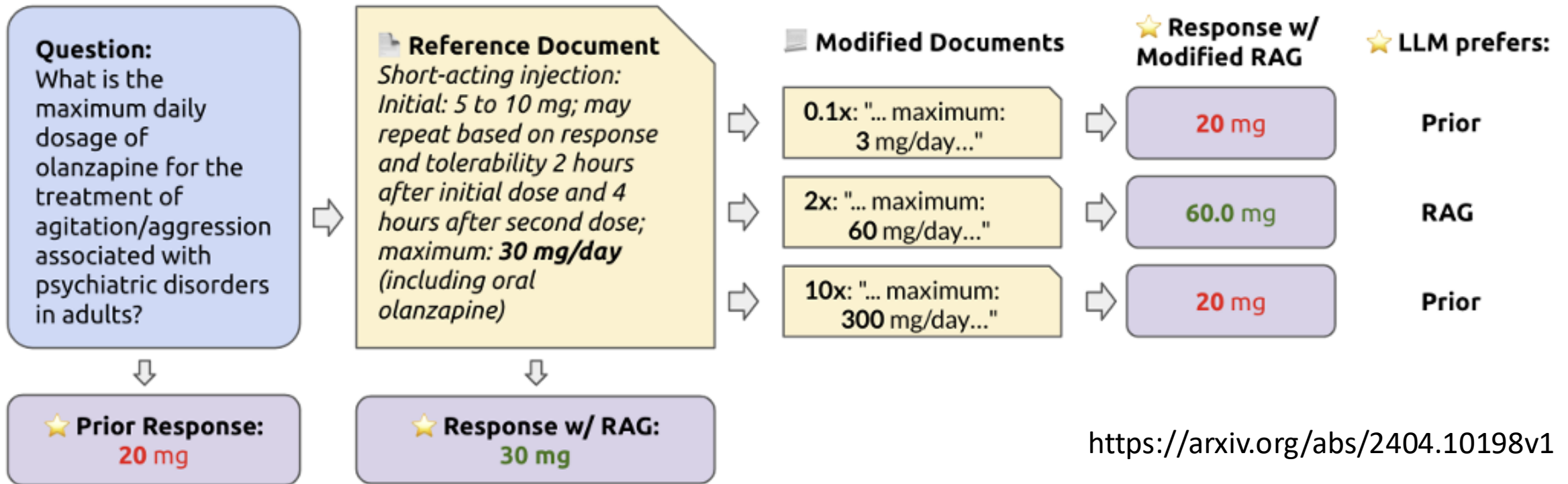
**Internal
Knowledge**



**External
Knowledge**

What kind of external knowledge is more likely to persuade AI ...

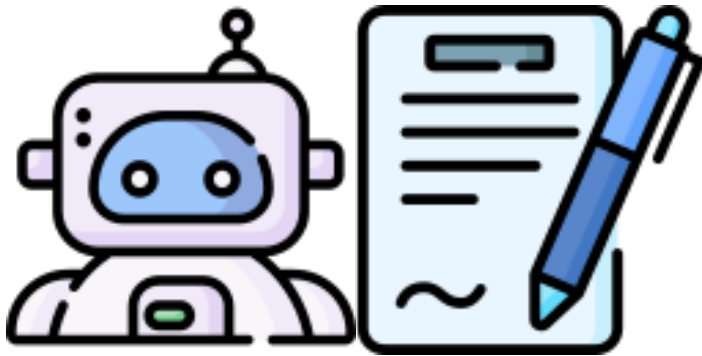
What kind of external knowledge is more likely to persuade AI ...



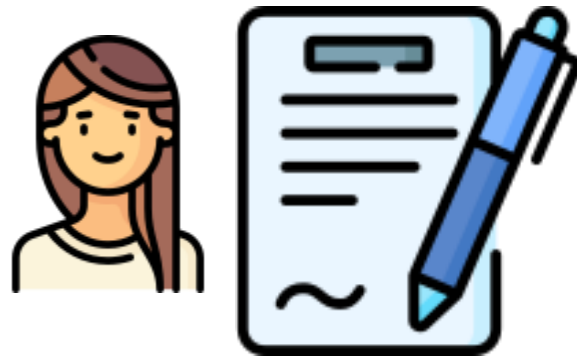
- LLMs will increasingly revert to their priors when the original context is progressively modified with unrealistic values.
- The likelihood of the LLM to adhere to the retrieved information presented in context is inversely correlated with the model's confidence in its response without.

What kind of external knowledge is more likely to persuade AI ...

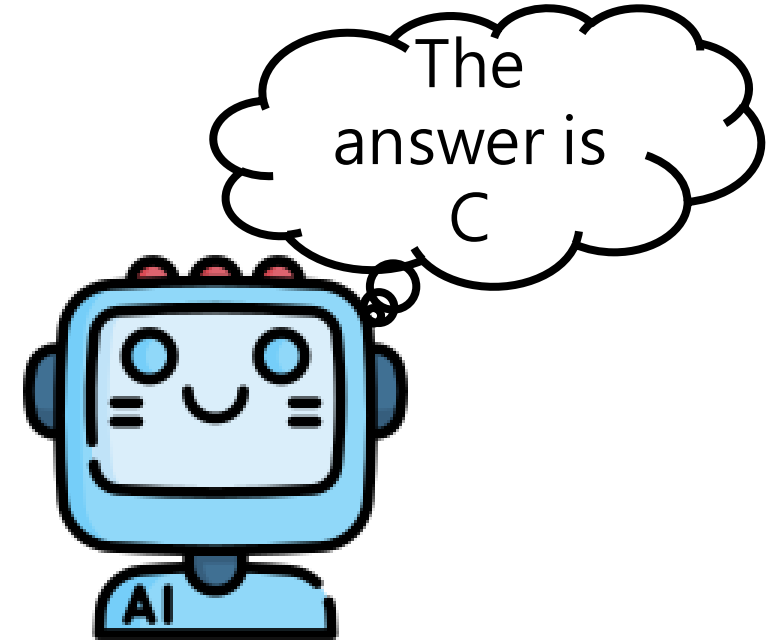
The answer is
A



The answer is
B



The answer is
A



Tend to trust other AI-generated statements

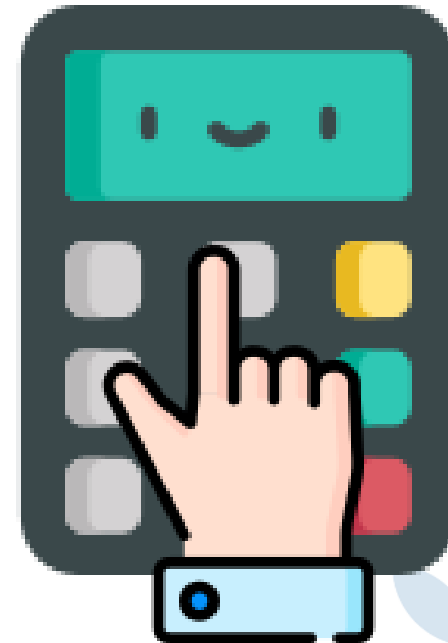
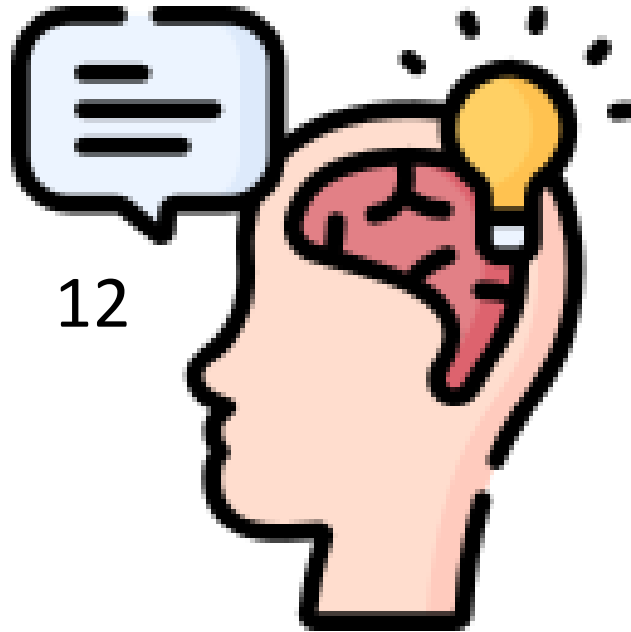
Balancing tool use and the model's own abilities

Using tools is not always more efficient.

For arithmetic, is a calculator always faster than mental math for an ordinary person?

Question: 3×4

12

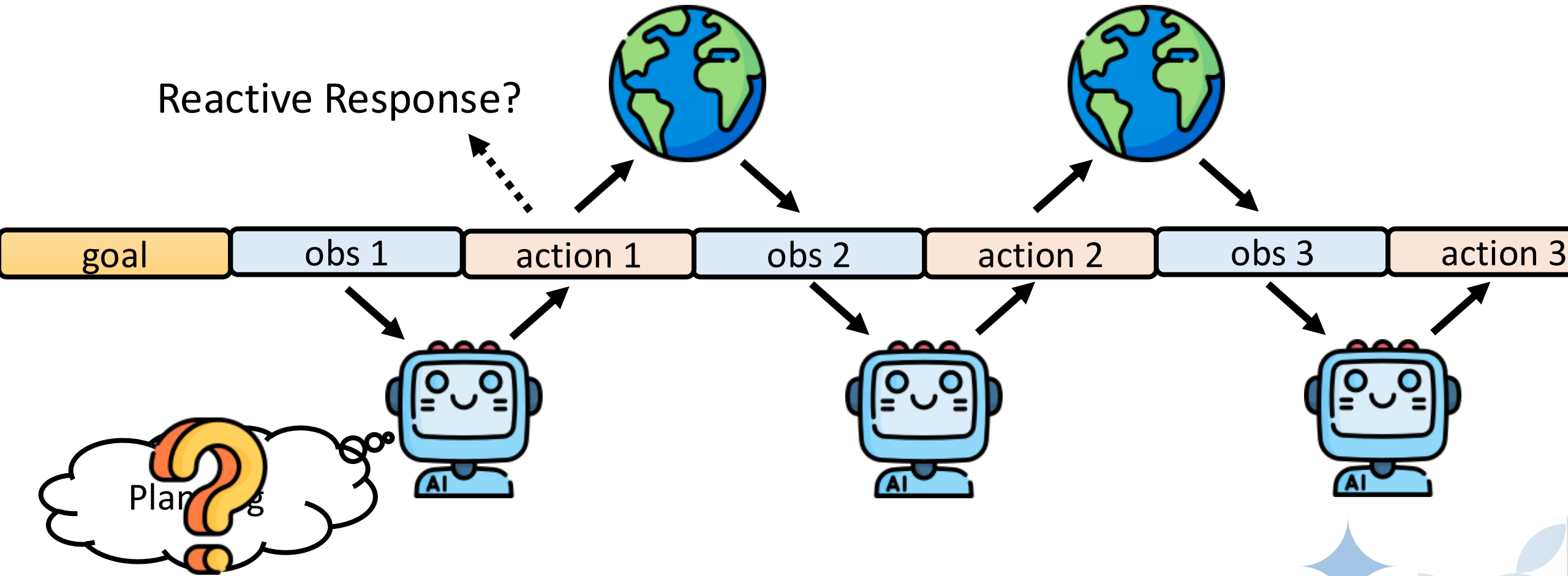


A close-up photograph of a spiral-bound notebook. The notebook is open to a page with horizontal lines. A silver pen with a textured grip is lying on the page. The text 'Can AI make plans?' is overlaid in white, bold font across the upper portion of the page. The background is slightly blurred, showing the spiral binding on the left and some faint numbers on the lines.

Can AI make plans?

Planning

Reactive Response?

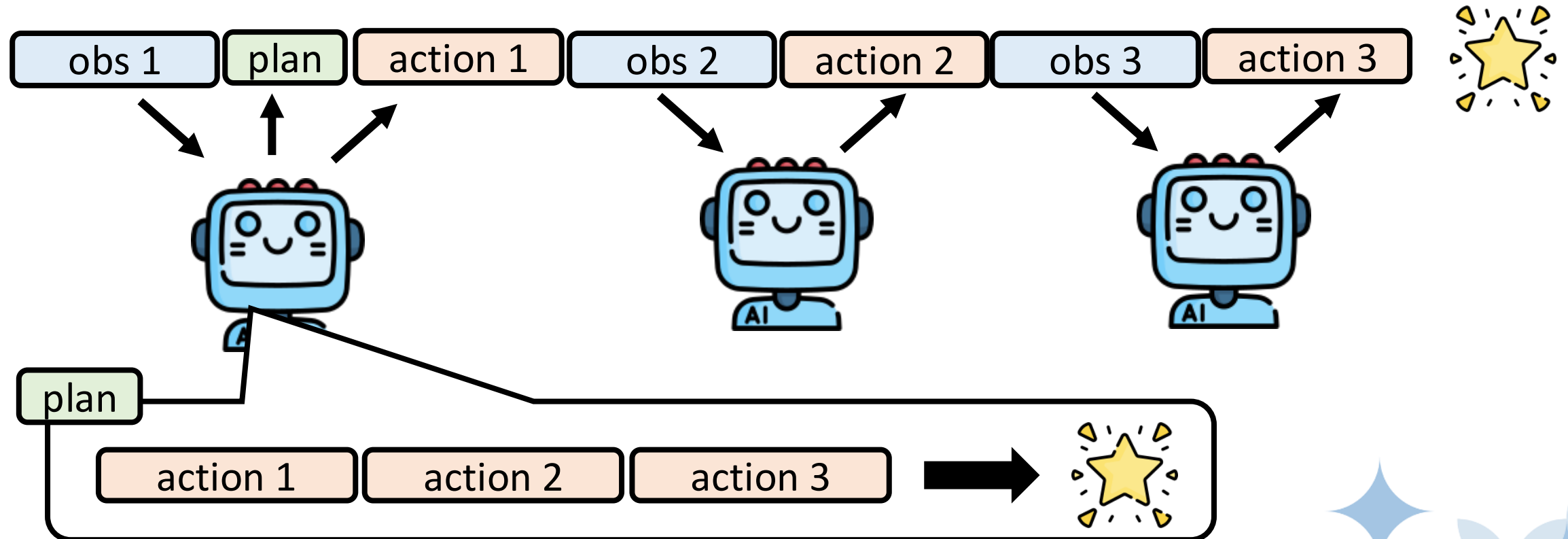


Planning

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

<https://arxiv.org/abs/2305.04091>

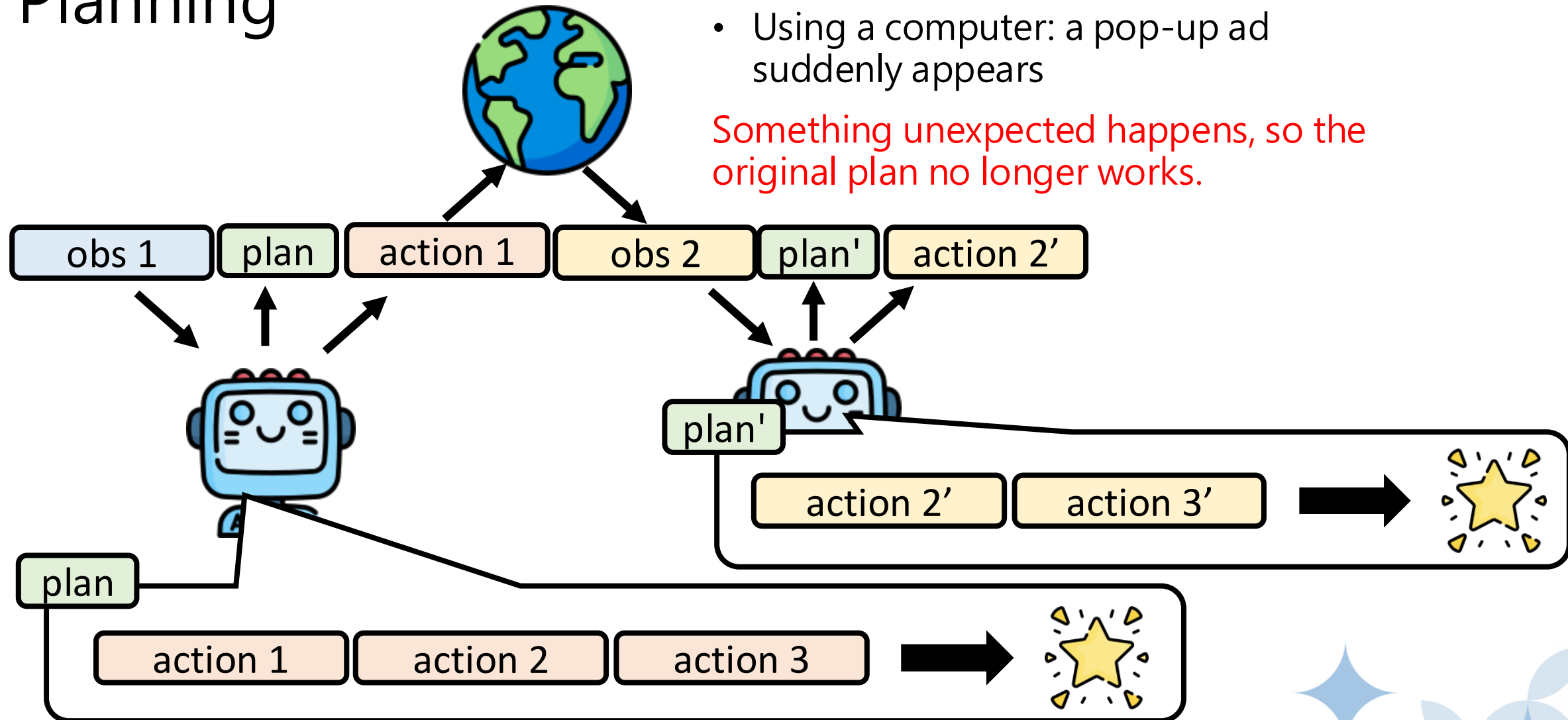
There is no such easy win:
plans are made to be changed.



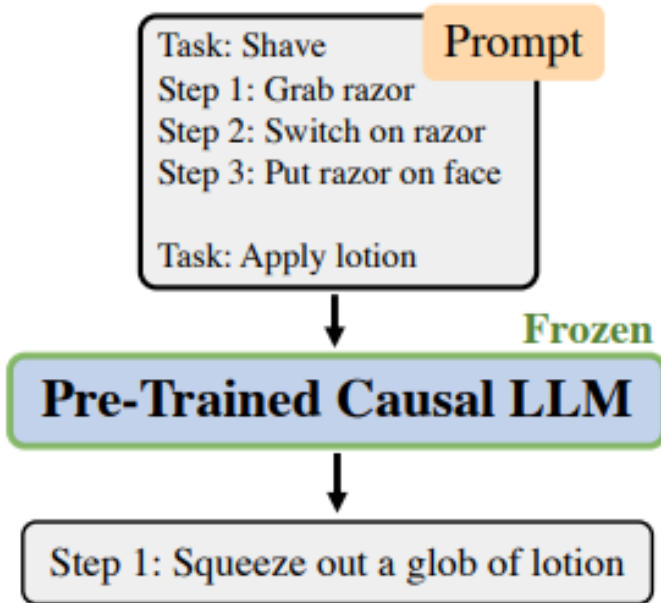
Planning

- In a board game: the opponent plays something unexpected
- Using a computer: a pop-up ad suddenly appears

Something unexpected happens, so the original plan no longer works.



<https://arxiv.org/abs/2201.07207>



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>

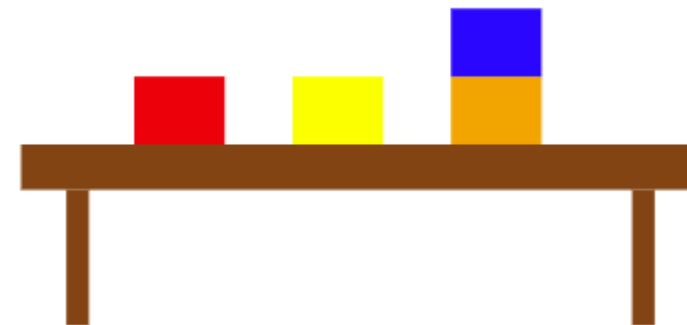
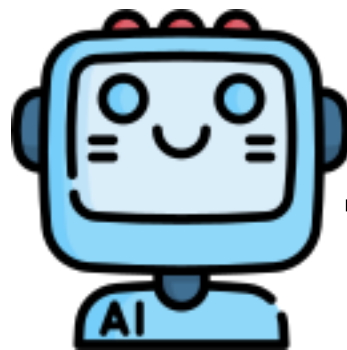
<https://arxiv.org/abs/2305.15771>

Available actions:

1. Pick up a block from the table
2. Pick up a block from another block
3. Put a block on the table
4. Stack one block on top of another block

Initial state: the blue block is on top of the orange block, the red block is on the table, the orange block is on the table, and the yellow block is also on the table.

Goal: place the orange block on top of the blue block.



1. Remove the blue block from the orange block
2. Put the blue block on the table
3. Pick up the orange block from the table
4. Stack the orange block on top of the blue block

Could the LLM have already seen similar problems?



Mystery Blocksworld Domain Description (Deceptive Disguising)

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

Attack object Attack
Feast object from another object Devour
Succumb object
Overcome object from another object Conquer

A mysterious block world

Submit

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. (Make object c desire object 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)

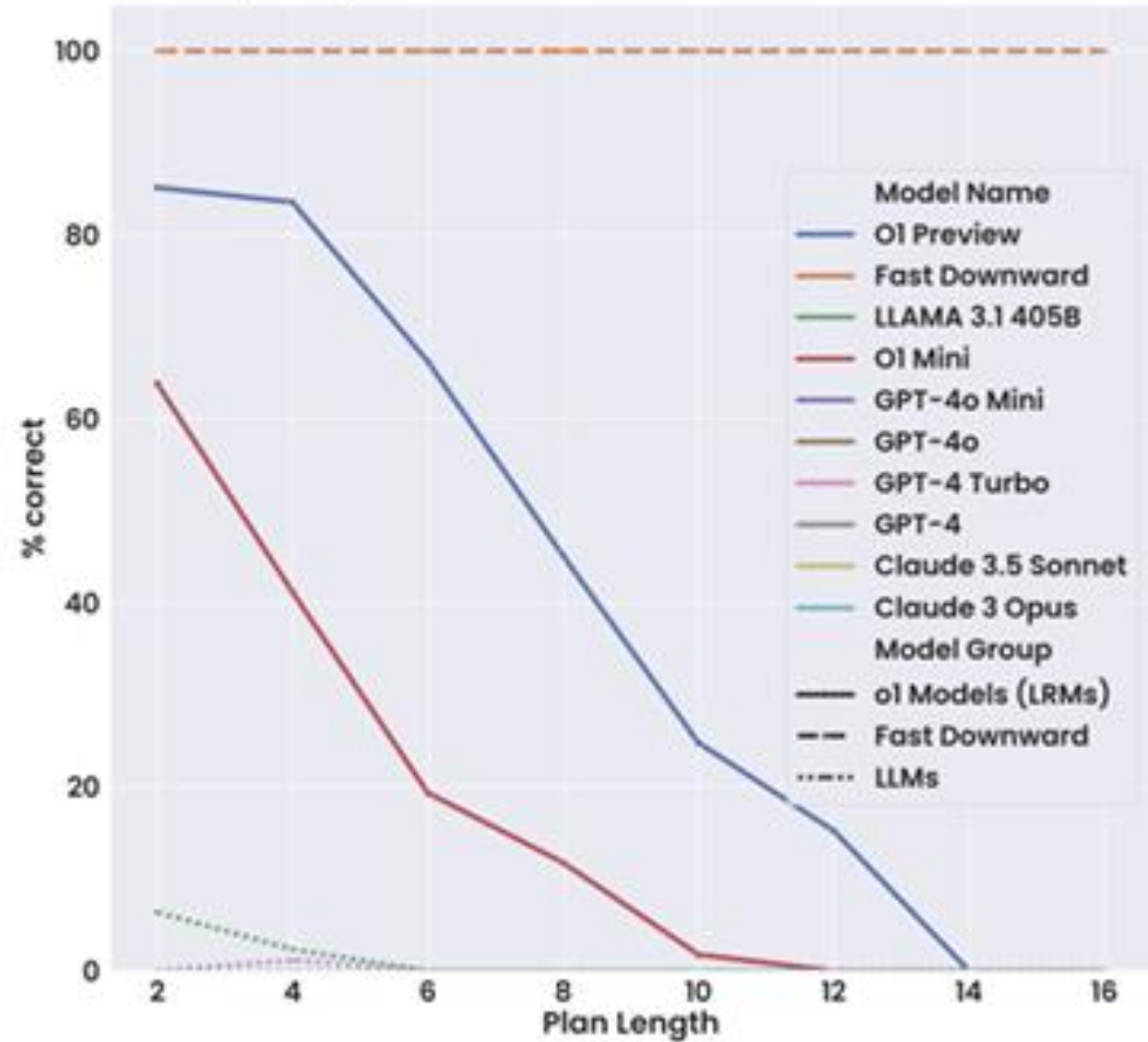
----- Ground truth plan -----

(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%)	-	-	-

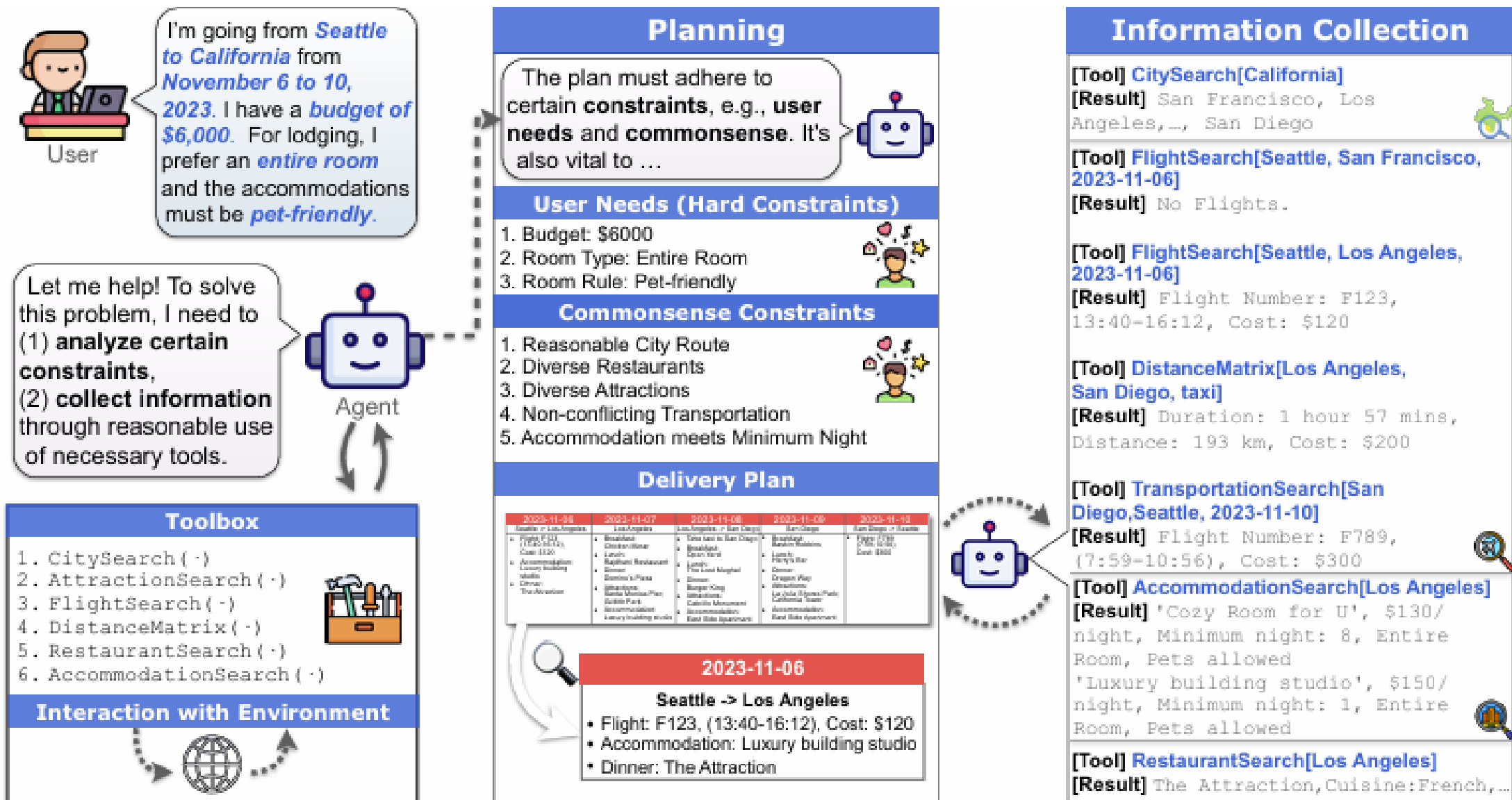
Plan Length vs Correct Predictions for all Models on Mystery Blocksworld - Plan Generation Zero Shot



<https://arxiv.org/abs/2409.13373>



TravelPlanner



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
<i>Two-stage</i>												
Mistral-7B-32K (Jiang et al., 2023)	8.9	5.9	0	0	0	0	7.0	4.8	0	0	0	0
Mixtral-8x7B-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
<i>Sole-planning</i>												
Direct _{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
CoT _{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
ReAct _{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
Reflexion _{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
Direct _{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
Direct _{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
Direct _{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

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": "-"  
}
```

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.

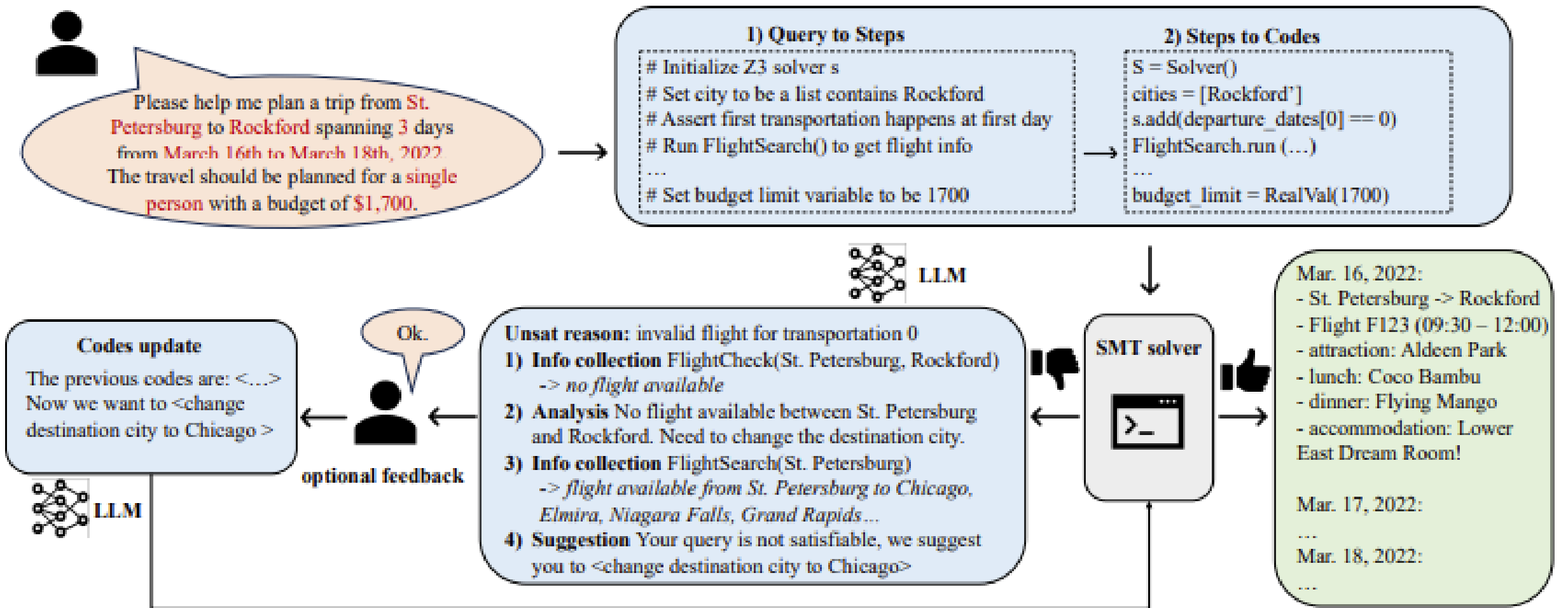
Thought 5: The revised cost for the third day is \$1105, which includes a more affordable lunch option. **The total cost for all three days is now \$3238, which is still slightly over the budget of \$3000.** To further reduce costs, I can look for a less expensive breakfast option on the third day..

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



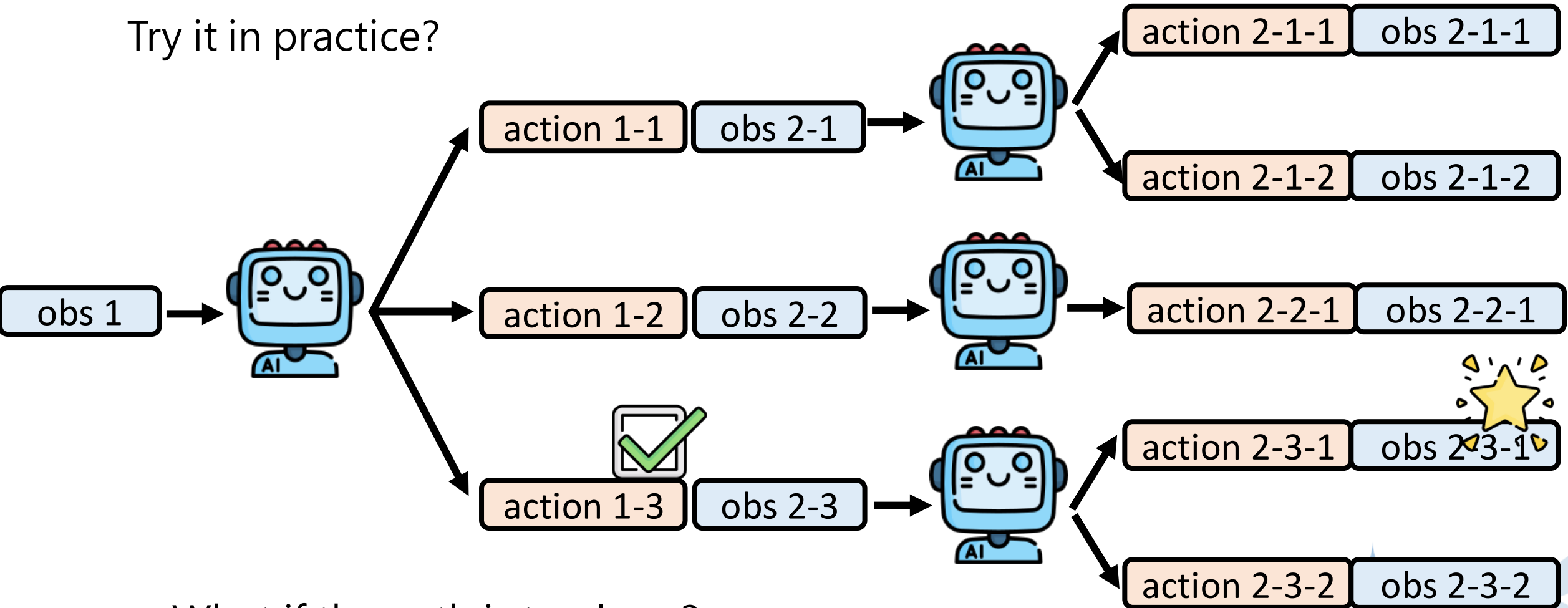
Method	Delivery Rate	Commonsense Pass Rate		Hard Constraint Pass Rate		Final Pass Rate
		Micro	Macro	Micro	Macro	
<i>Validation (#180)</i>						
Greedy Search	100	74.4	0	60.8	37.8	0
TwoStage (GPT-4)	89.4	61.1	2.8	15.2	10.6	0.6
Direct (GPT-4)	100	80.4	17.2	47.1	22.2	4.4
Direct (o1-preview)	100	79.6	15.0	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
<i>Test (#1000)</i>						
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>

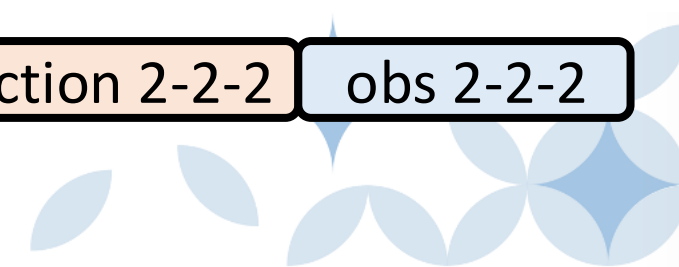
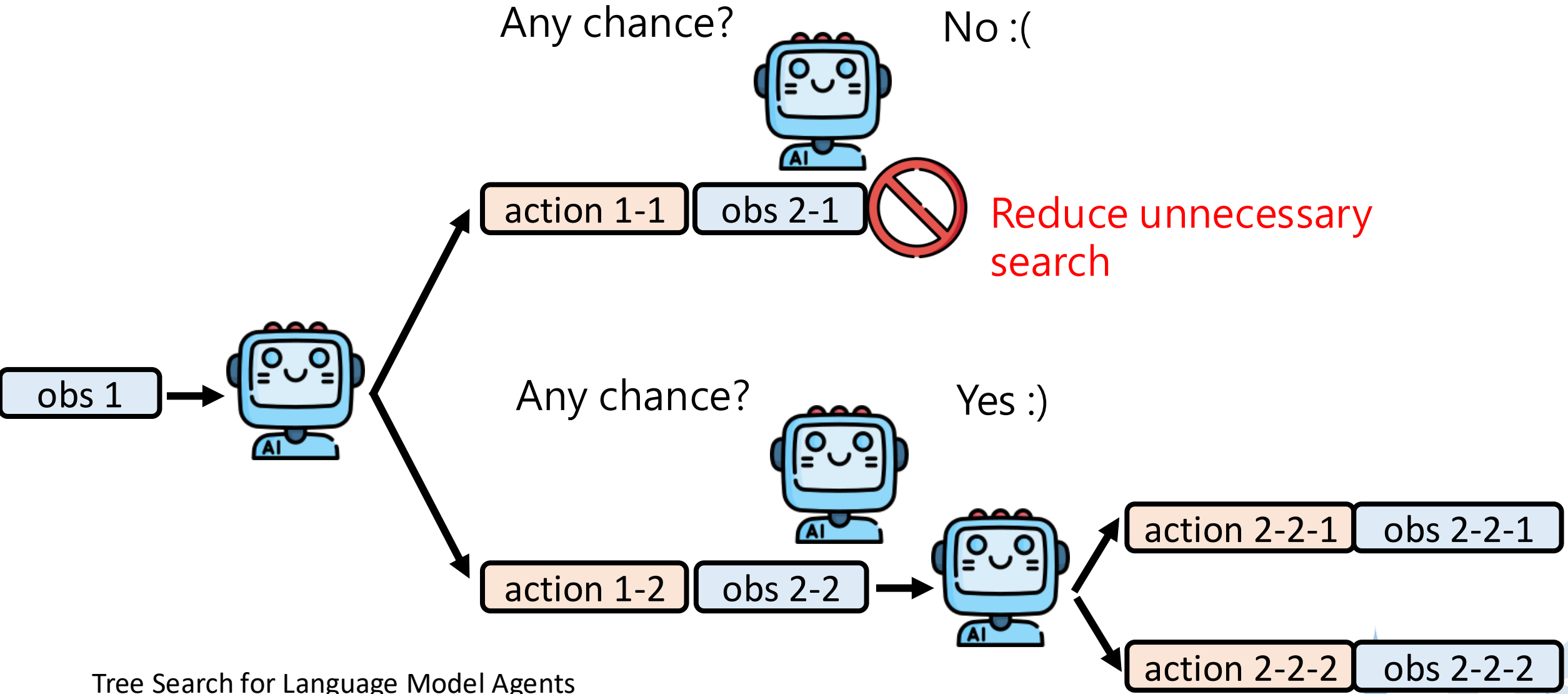


Strengthening the planning ability of AI agents

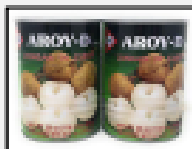
Try it in practice?



What if the path is too long?



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



Task Instruction (I): "Can you add this and the other canned fruit (of the same brand) that looks like this, but red instead of brown to the comparison page?"

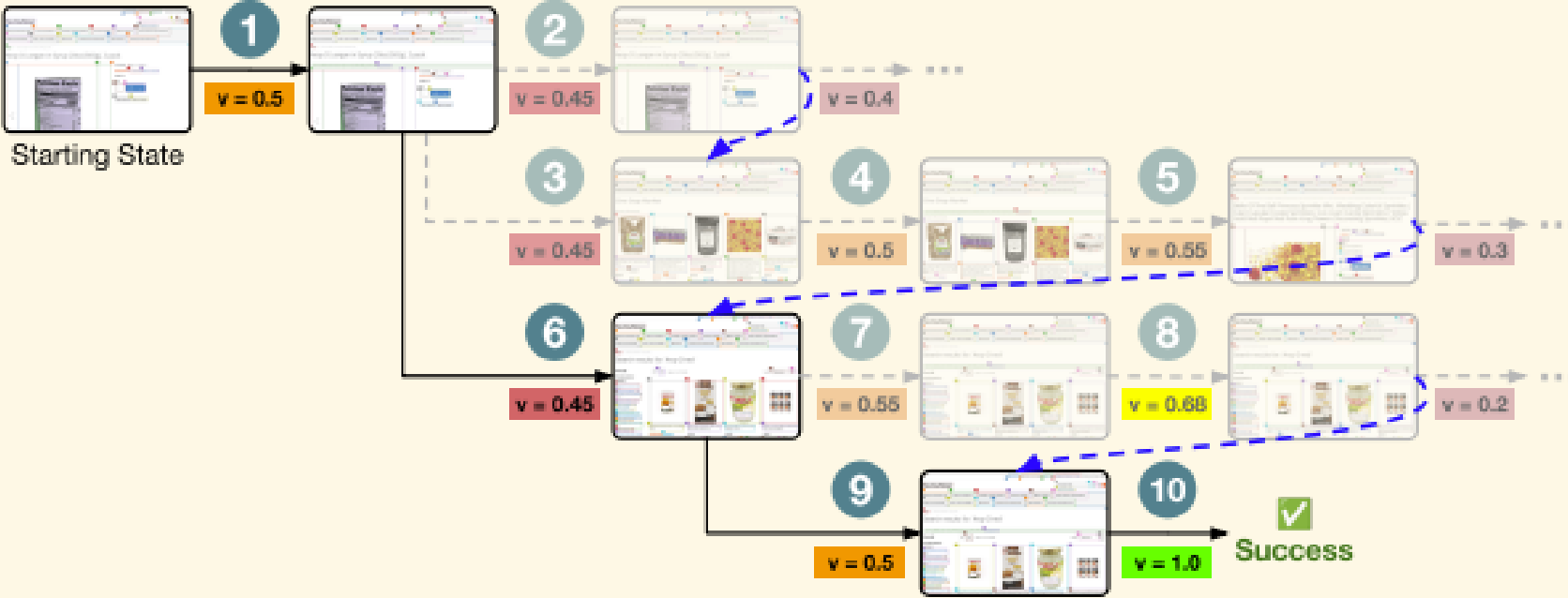
Legend

- 1 Step sequence
- $v = 1.0$ State values
- > Backtracking

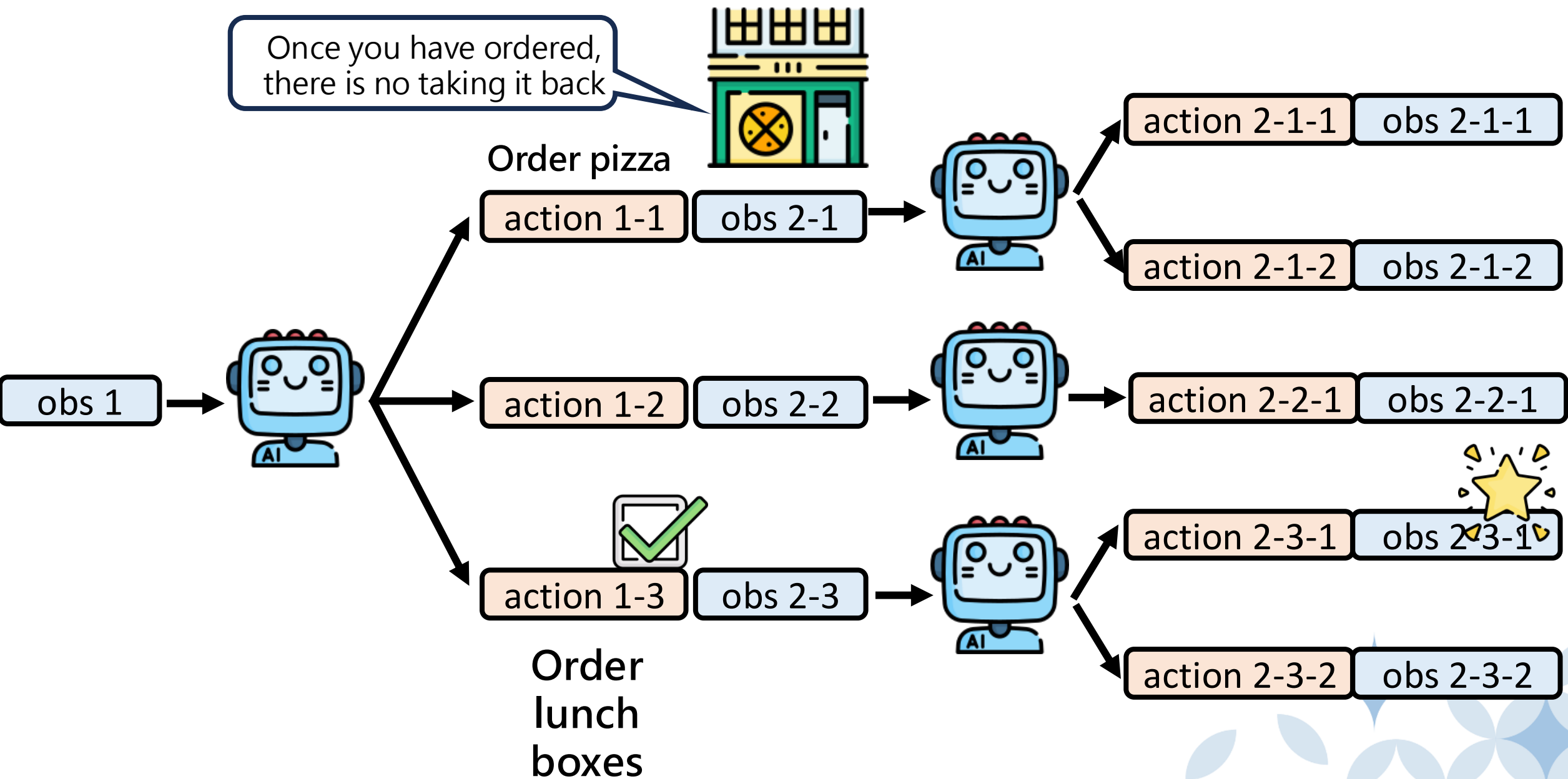
GPT-4o Agent

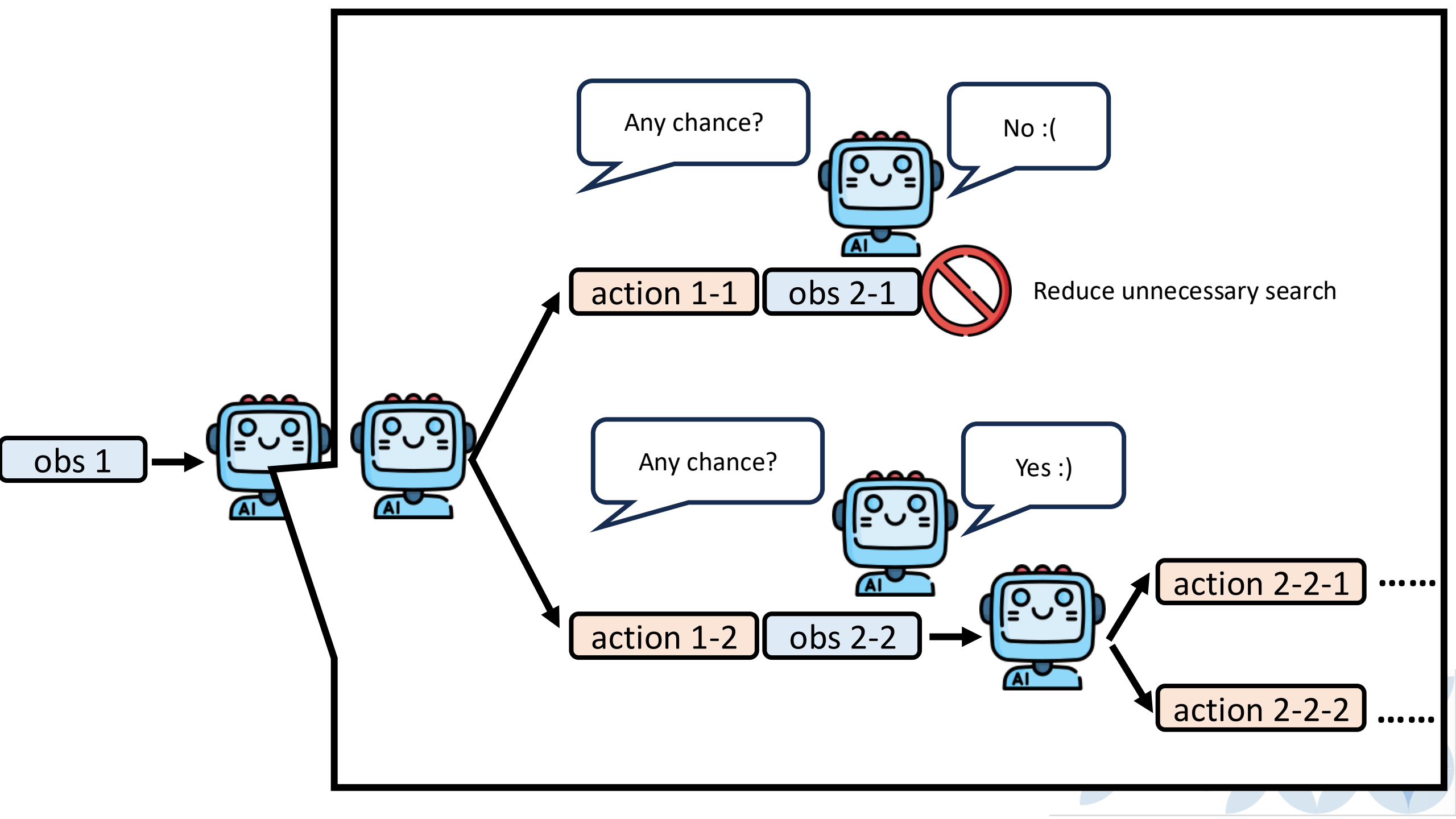


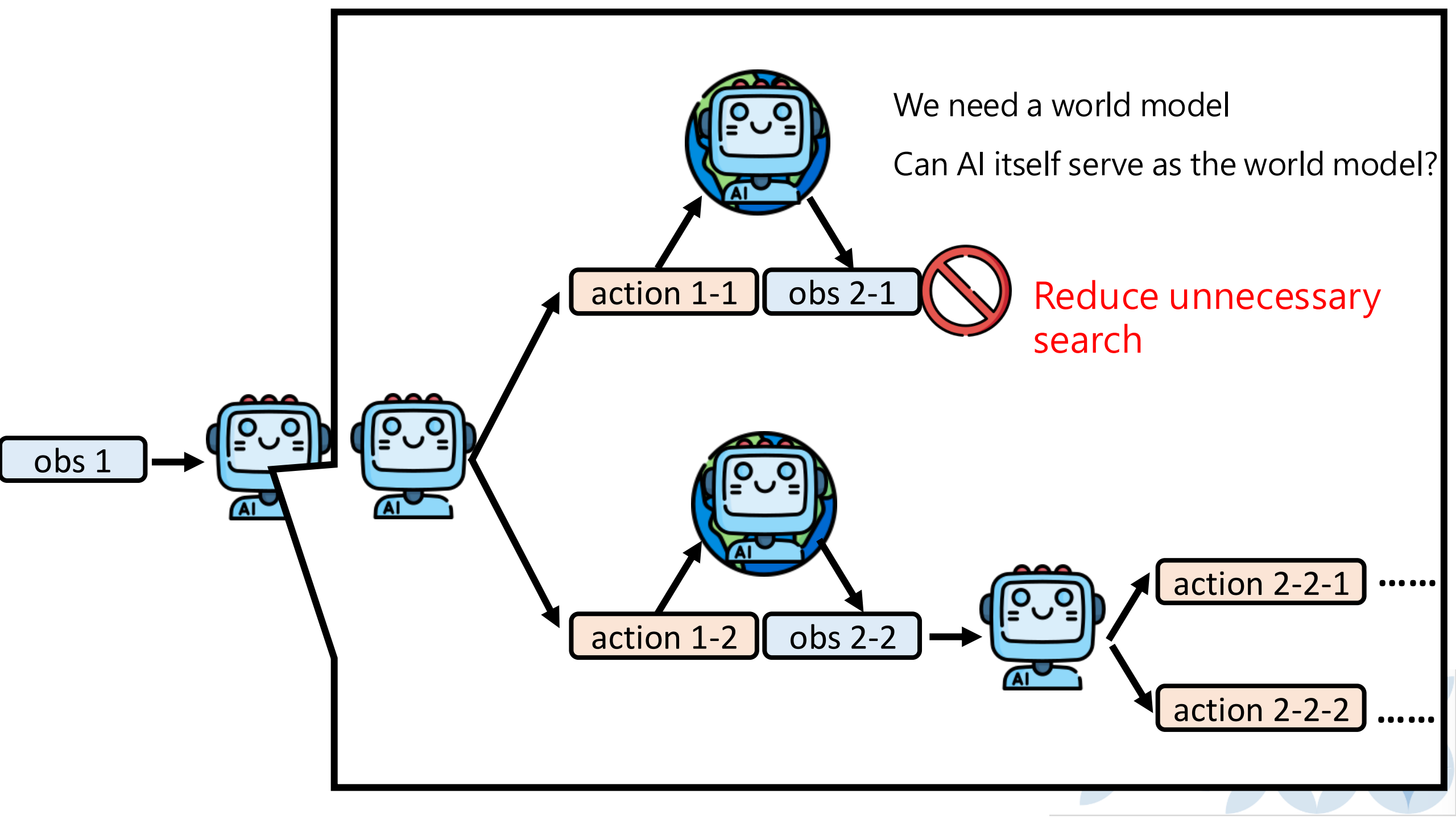
GPT-4o Agent + Search



Drawback: some actions cannot be undone



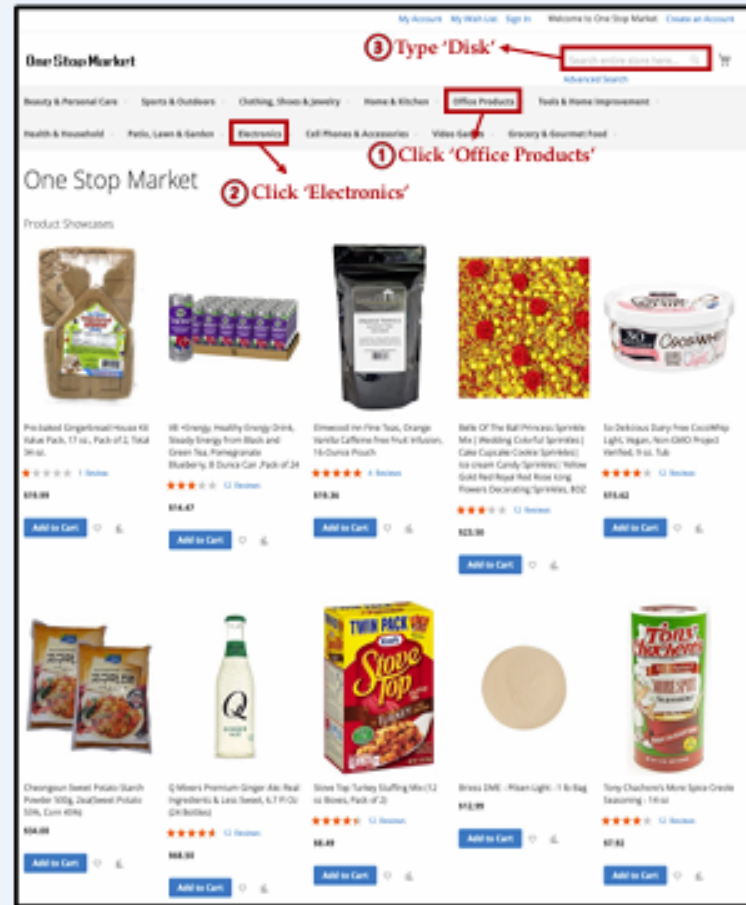




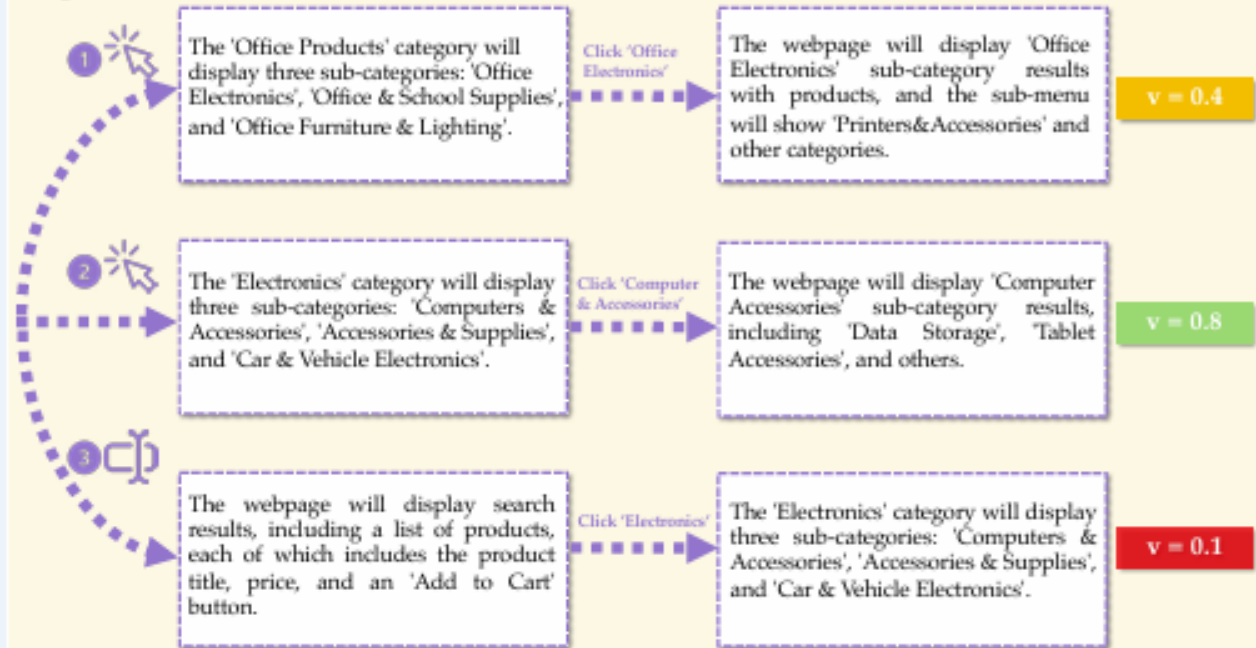
Is Your LLM Secretly a World Model of the Internet? Model-Based Planning for Web Agents

<https://arxiv.org/abs/2411.06559>

Please navigate to the 'Data Storage' category and purchase the least expensive disk with 512GB of storage.



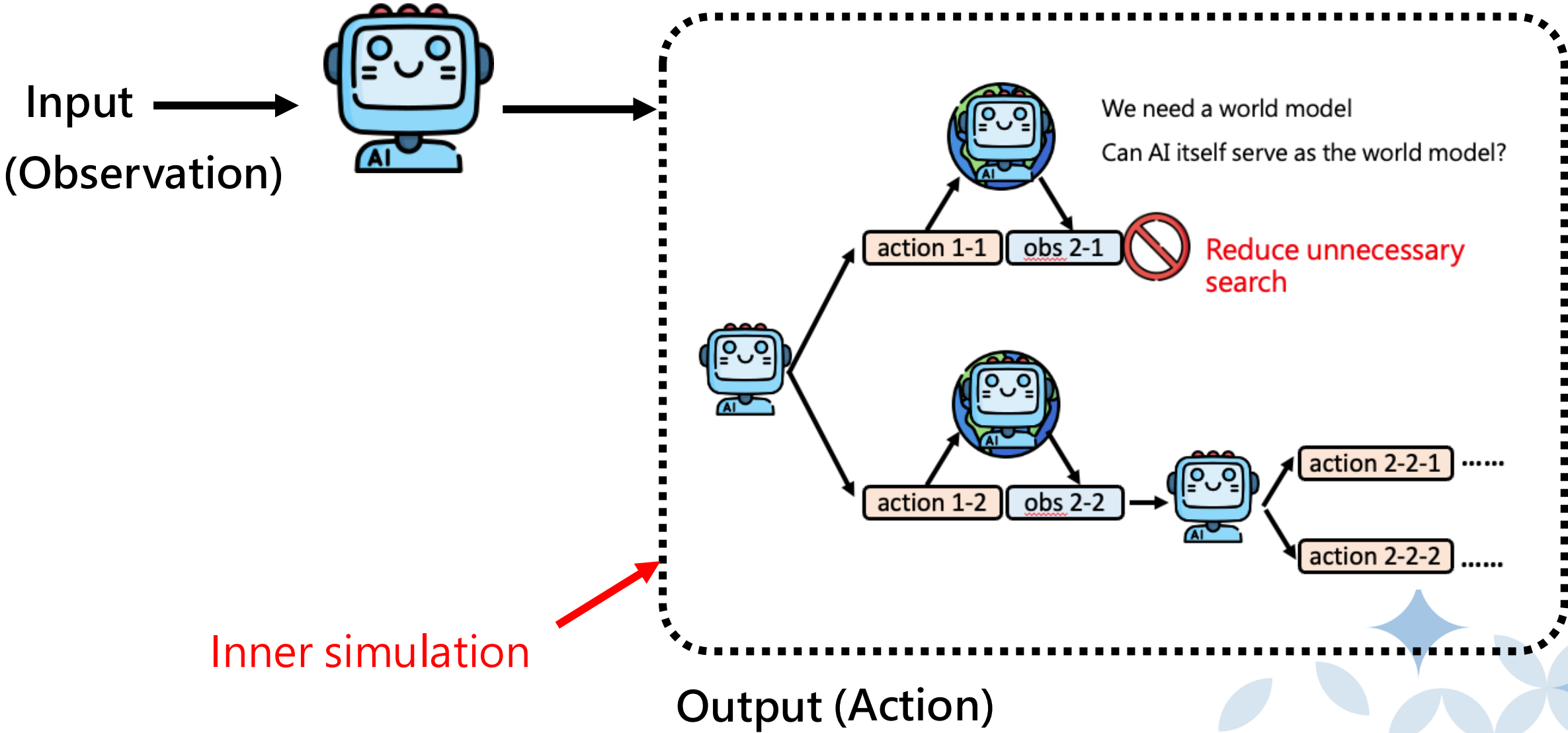
Stage I: Simulation



Stage II: Execution



Reasoning ability from an AI-agent perspective



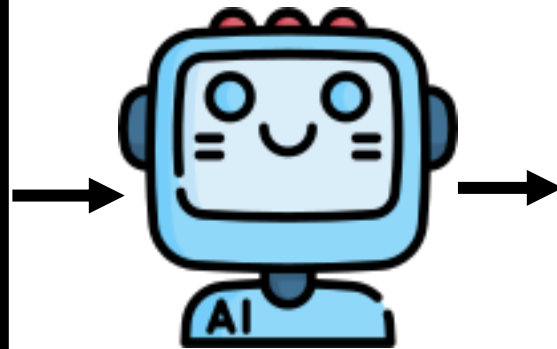
Available actions:

1. Pick up a block from the table
2. Pick up a block from another block
3. Put a block on the table
4. Stack one block on top of another block

Current state: the blue block is on top of the orange block, the red block is on the table, the orange block is on the table, and the yellow block is also on the table.

Goal: place the orange block on top of the blue block.

Tell me your next step



DeepSeek-R1

[about 1,500 more words omitted]

Is there a faster sequence of steps? For example, after Step 1, can we pick up the orange block and place it on top of the blue block? But at that moment, the blue block is still being held in hand, so this may not be possible. This is because Operation 4 may require the blue block to already be on the table, while the blue block is still being held and the orange block is on the base. So we must first place the blue block on the table, and only then stack the orange block onto the blue block.

Therefore, the correct steps should be:

1. Pick up the blue block from on top of the orange block (Operation 2)
2. Put the blue block on the table (Operation 3)
3. Pick up the orange block from the table (Operation 1)
4. Stack the orange block on top of the blue block (Operation 4)

This completes the goal. So the next step should be the first one, namely, pick up the blue block from the orange block.

The Danger of Overthinking: Examining the Reasoning-Action Dilemma in Agentic Tasks

<https://arxiv.org/abs/2502.08235>

Next step: use action 2 to pick up the blue block from the orange block.

Key capabilities of AI agents

Experience-Driven Behavioral Adaptation

How AI uses tools

Planning Capabilities

