



LM agent, language grounding

Yao Teng, Kaiyue Sun

LM agent

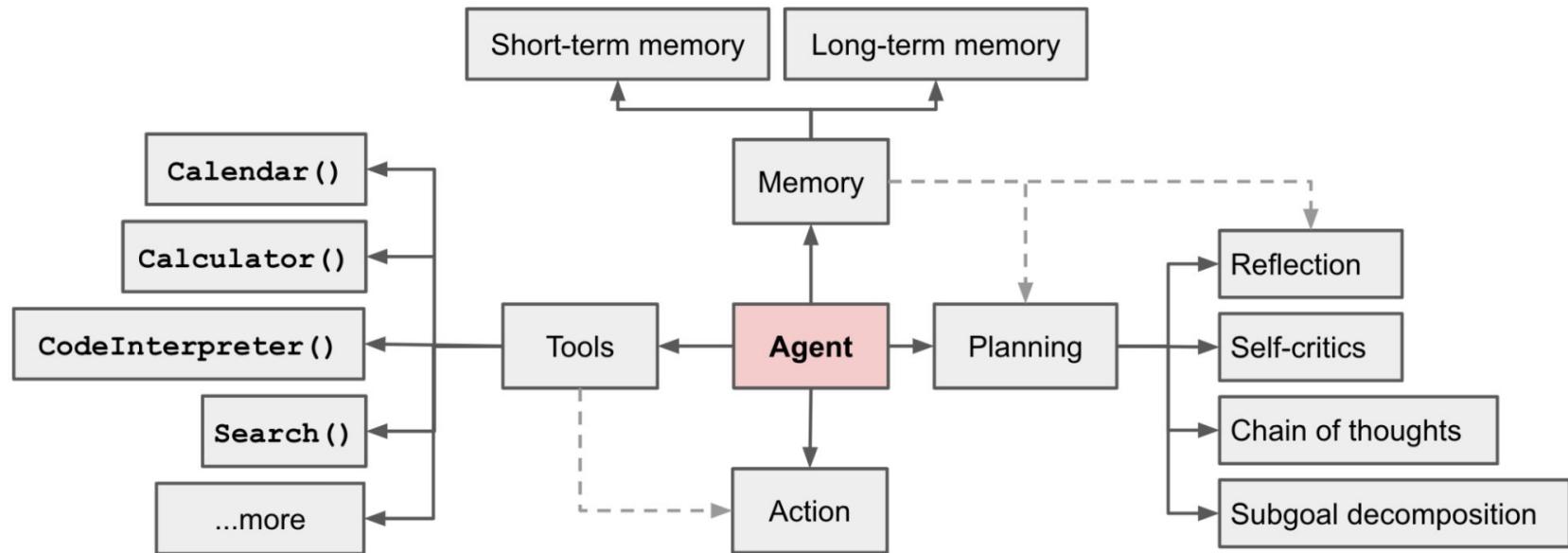
An agent controlled by a large language model (LLM)

An LM agent uses an LLM to produce natural language responses that are relevant to the user query and the context

LM agent

- Tool use:
 - The agent learns to call external APIs for extra information
- Planning:
 - The agent breaks down large tasks into smaller, manageable subgoals
 - The agent can do **self-criticism** and **self-reflection** over past actions
- Memory use:
 - Long term: The agent is able to retain and recall (infinite) information over extended periods
 - Short term: all the in-context learning

LM agent



Planning

Task Decomposition

- Examples:
 - Chain of thought
 - Tree of Thoughts

Planning

Task Decomposition

- Examples:
 - Chain of thought
 - Tree of Thoughts

Implementation of Task Decomposition

- prompting like "Steps for XYZ.\n1.", "What are the subgoals for achieving XYZ?"
- using task-specific instructions; e.g. "Write a story outline." for writing a novel
- human inputs

Planning

LLM+P

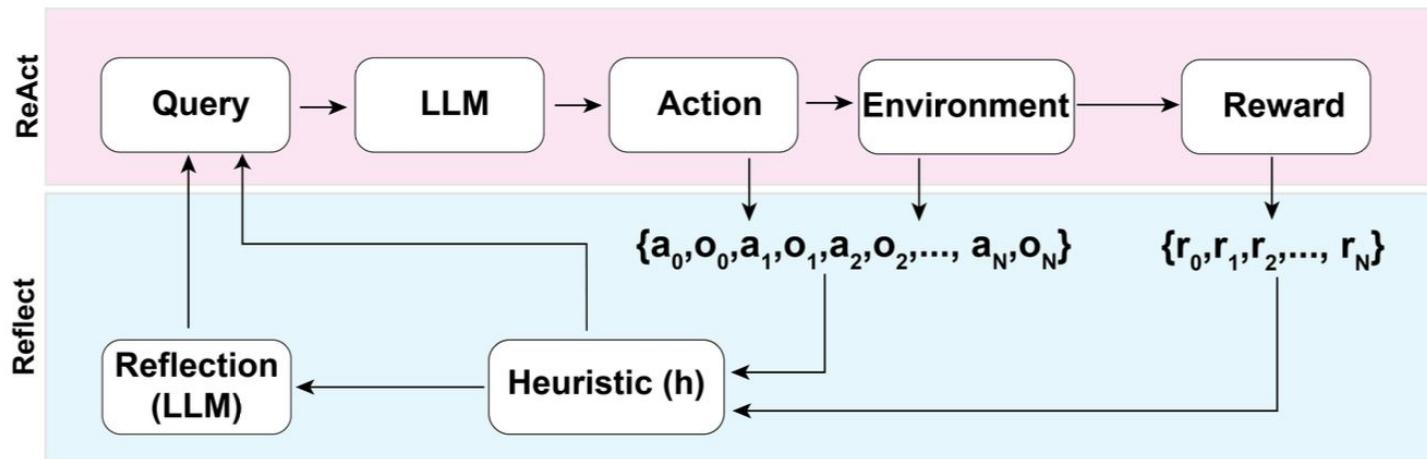
- An external classical planner
- Long-horizon planning

Self-Reflection & -Criticism

- Refining past action decisions
- Correcting previous mistakes
- Examples:
 - ReAct
 - Reflexion

Self-Reflection & -Criticism

- Examples:
 - Reflexion



Memory

- Processes used to acquire, store, retain, and later retrieve information.
- Sensory Memory:
 - Retain impressions of sensory information (visual, auditory, etc) after the original stimuli
 - Last up to a few seconds
 - Subcategories: iconic memory (visual), echoic memory (auditory), and haptic memory (touch).
 - **Learning embedding representations for multi-modalities;**

Memory

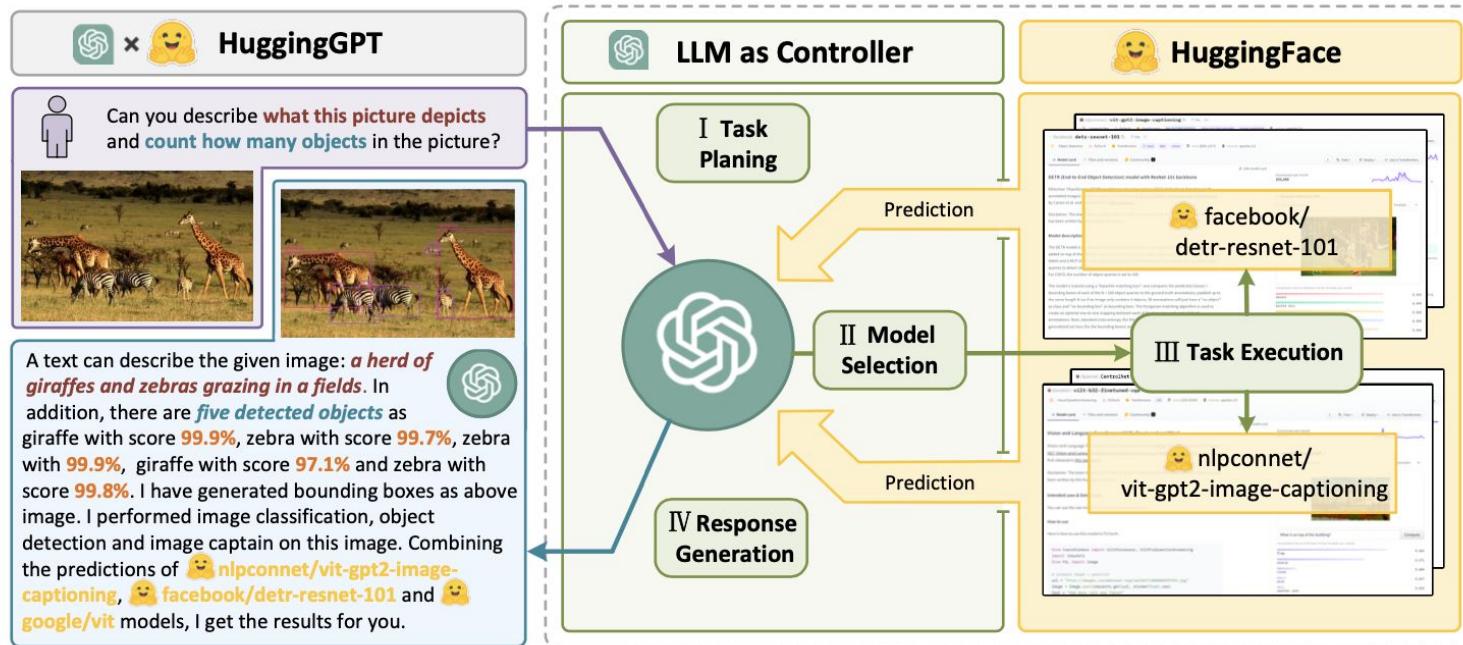
- Short-Term Memory (STM) or Working Memory:
 - For complex cognitive tasks: learning and reasoning.
 - About 7 items and lasts for 20-30 seconds.
 - **In context learning**

Memory

- Long-Term Memory (LTM):
 - A few days to decades, unlimited storage
 - Two subtypes:
 - Explicit / declarative memory
 - Implicit / procedural memory
 - External vector store, fast retrieval.

Tool Use

HuggingGPT



Tool Use

- 4 stages of Tool Use
 - Task planning
 - parsing requests
 - Model selection:
 - Task distribution
 - Task execution
 - Response generation

Experience Grounds Language

EMNLP 2020

Experience Grounds Language

- Where is NLP going?

WS1. Corpus (*our past*)

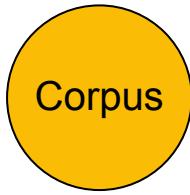
WS2. Internet (*most of current NLP*)

WS3. Perception (*multimodal NLP*)

WS4. Embodiment

WS5. Social

World Scopes



World Scope 1
Corpora & Representations

Text Dataset

```
( (S
    (NP (ADJP Battle-tested industrial)
          managers)
    (? here)
    (? always)
    (VP buck))
    (? (PP up
        (NP nervous newcomers)))
    (? (PP with
        (NP the tale
        (PP of
            (NP the
            (ADJP first))))))
    (? (PP of
        (NP their countrymen)))
    (? (S (NP *)
        to
        (VP visit
        (NP Mexico))))
    (? ,)
    (? (NP a boatload
        (PP of
            (NP warriors)))
        (VP blown
        (? ashore)
        (NP 375 years))))
    (? ago)
    (? .))
```

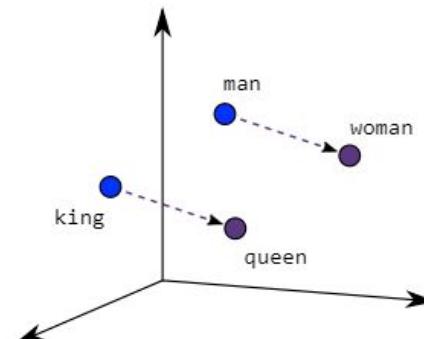
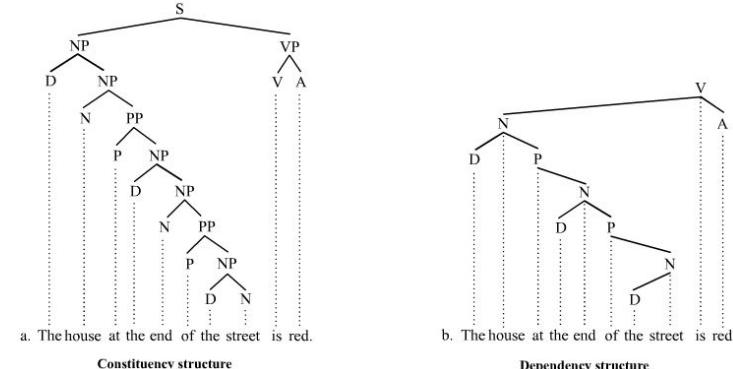
GloVe

word2vec

The Penn Treebank

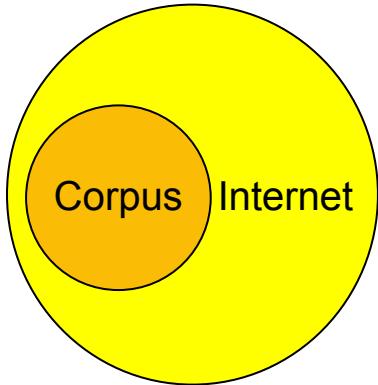
WS1: Corpora & Representations

- Representation study:
 - finding formal linguistic structure
 - e.g. recovering syntax trees:
 - Brown Corpus (1964)
 - Penn Treebank (1993)
 - unstructured fuzzy representations
 - Word2Vec
 - GloVe



Local context and Parameters that saturate on a small corpus

World Scopes



World Scope 2
The Written World

Internet Scale Text Dataset

Fine-tuning

The model is trained via repeated gradient updates using a large corpus of example tasks.

1 sea otter => loutre de mer ← example #1



gradient update



1 peppermint => menthe poivrée ← example #2



gradient update



1 plush giraffe => girafe peluche ← example #N

gradient update

1 cheese => ← prompt

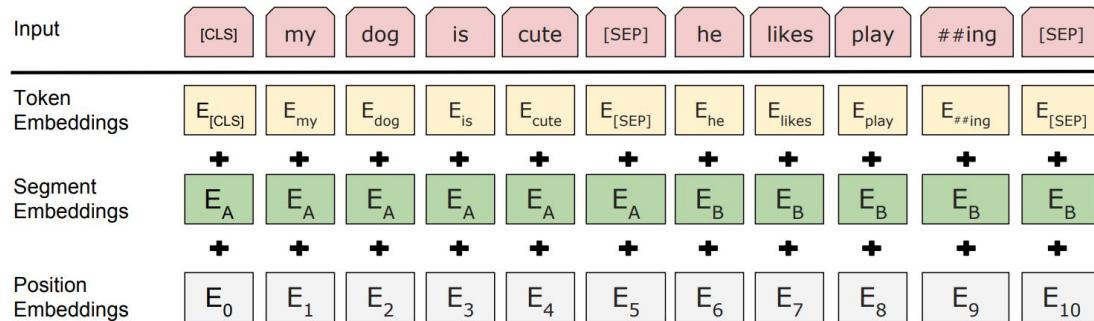
BERT

GPT

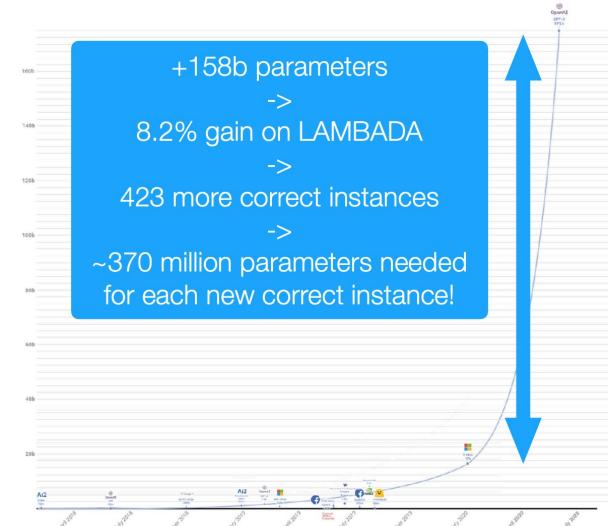
T5

WS2: The Written World

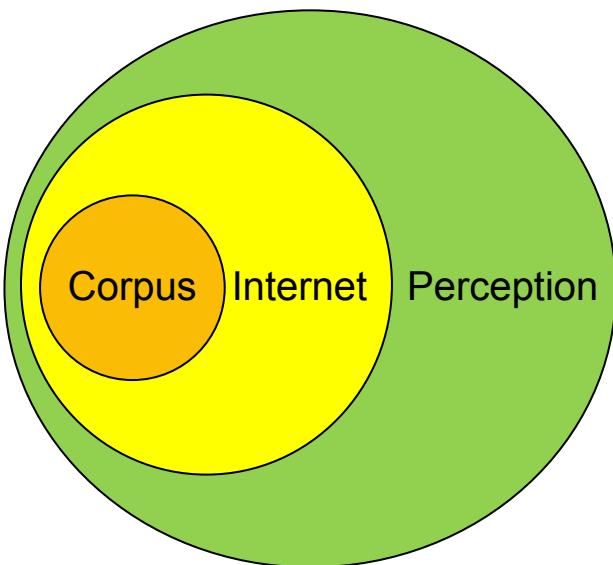
- “large and noisy text corpora”
- Scale in data & parameters:
 - representation ==> rich syntax and semantics



Multi-sentence context and Parameters that saturate eventually?



World Scopes



World Scope 3
The World of Sights and Sound

<https://yonatanbisk.com/papers/2020-EMNLP.pdf>

<https://www.rd.com/list/cutest-cat-breeds/>

https://www.freepik.com/premium-vector/audio-equalizer-line-style-music-tone-effect_29954327.htm

https://www.freepik.com/premium-photo/samples-textiles-upholstery-different-colors-thicknesses-closeup-details-multicolored-fabric-samples-set-isolated-white-background_12212751.htm

Vision, audio, touch
other(perception)



CLIP

SD

Flamingo

GPT4V

Dall-E3

NExT-GPT

WS3: Language beyond text: sight, sound

Gestures



Facial expression

(Even just emojis 😊)



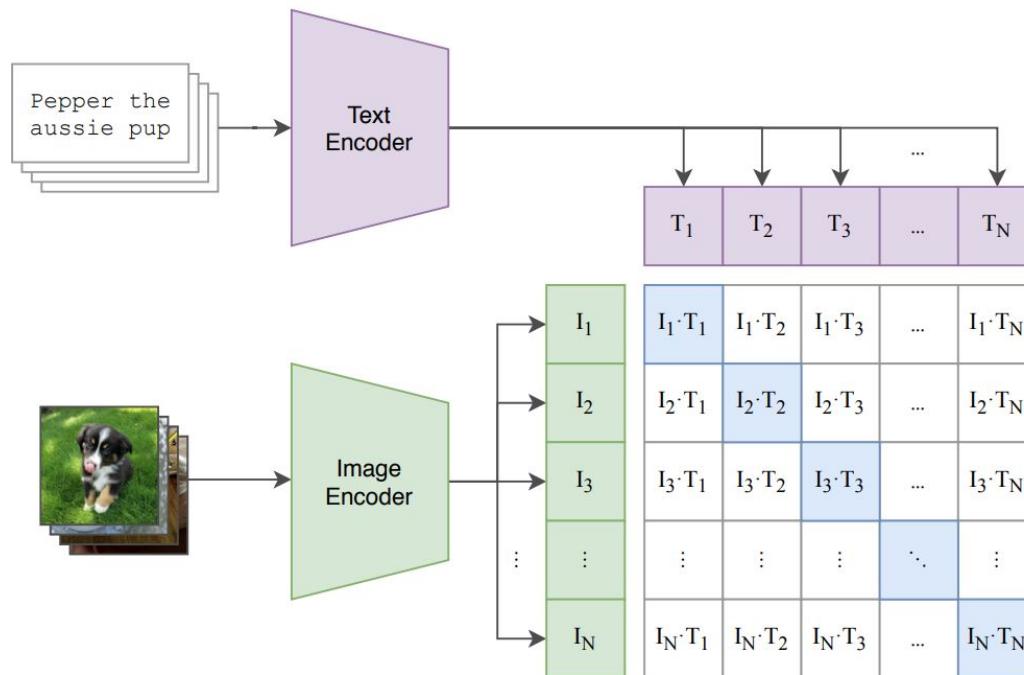
Intonation/Stress

- I didn't take the test yesterday.
(Somebody else did.)
- I didn't take the test yesterday.
(I did not take it.)
- I didn't take the test yesterday.
(I did something else with it.)
- I didn't take the test yesterday.
- I didn't take the test yesterday.
- I didn't take the test yesterday.

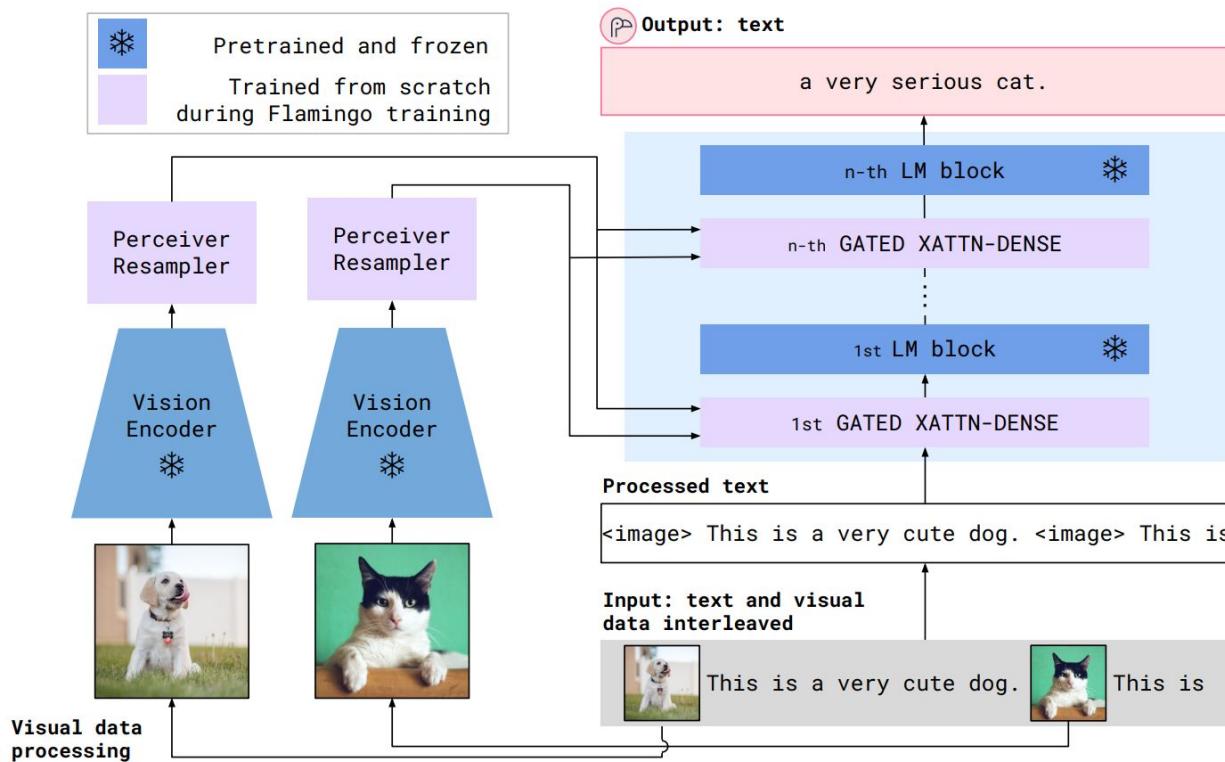
[https://en.wikipedia.org/wiki/Stress_\(linguistics\)](https://en.wikipedia.org/wiki/Stress_(linguistics))

WS3: Language beyond text: sight, sound

(1) Contrastive pre-training



WS3: Language beyond text: sight, sound



WS3: Language beyond text: sight, sound

Sec. 2.3 Interleaved Image-text Inputs

Interleaved Image-text Pairs

Prompt:

How much did I pay for tax?
Where is the information?



GPT-4V:

In the first receipt, you paid \$3.72 for tax. The information is located at the bottom of the receipt. In the second receipt, you paid \$42.23 for tax. The information is located near the bottom of the receipt, under the "Sales Tax" line. In the third receipt, you paid \$4.58 for tax. The information is located at the bottom of the receipt, under the "TAX" line. In total, you paid \$50.53 in tax across all three receipts.

Prompt Continued

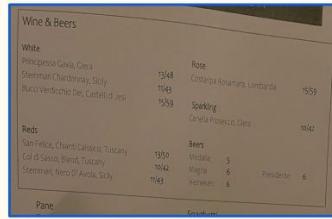


Prompt:

How much should I pay for the beer on the table according to the price on the menu?



Here is a photo of the menu.

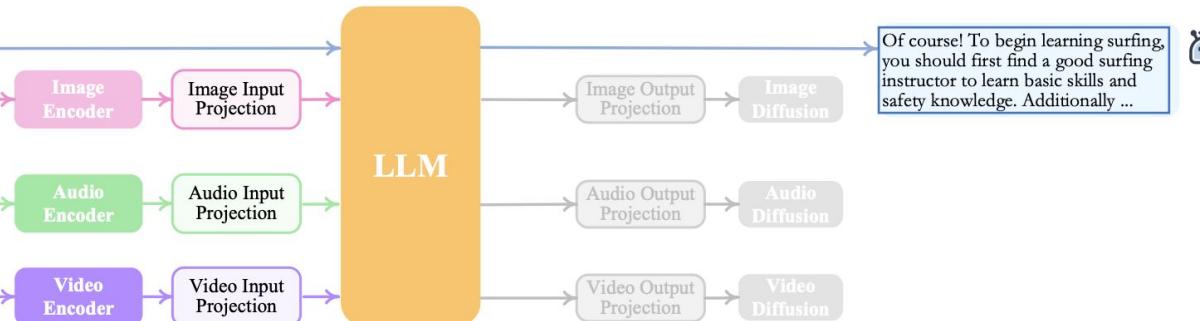


GPT-4V:

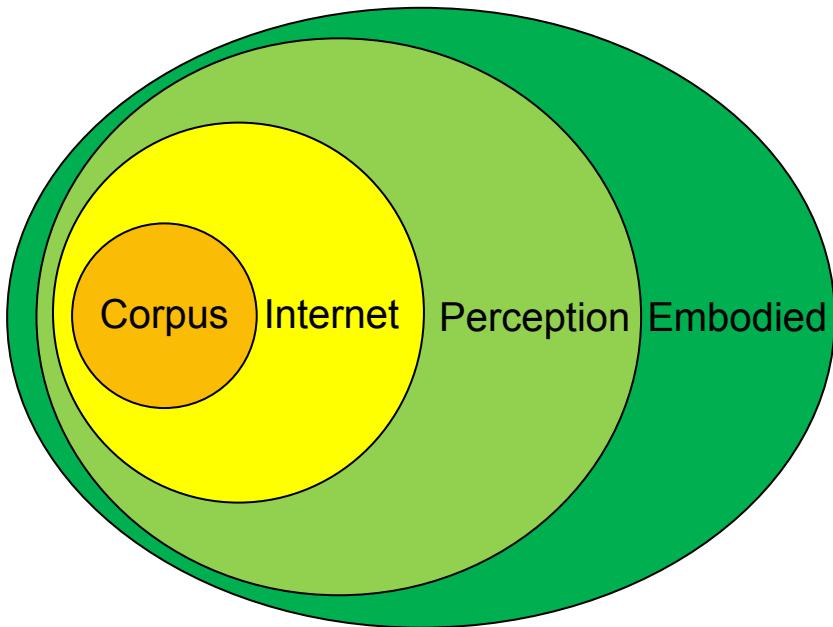
According to the menu, the price for a Magna beer is \$6. Since there are two Magna beers on the table, the total cost for the beers would be \$12.

Figure 1: GPT-4V can work with multi-image and interleaved image-text inputs. Check Section 2.3 for detailed discussions.

WS3: Language beyond text: sight, sound

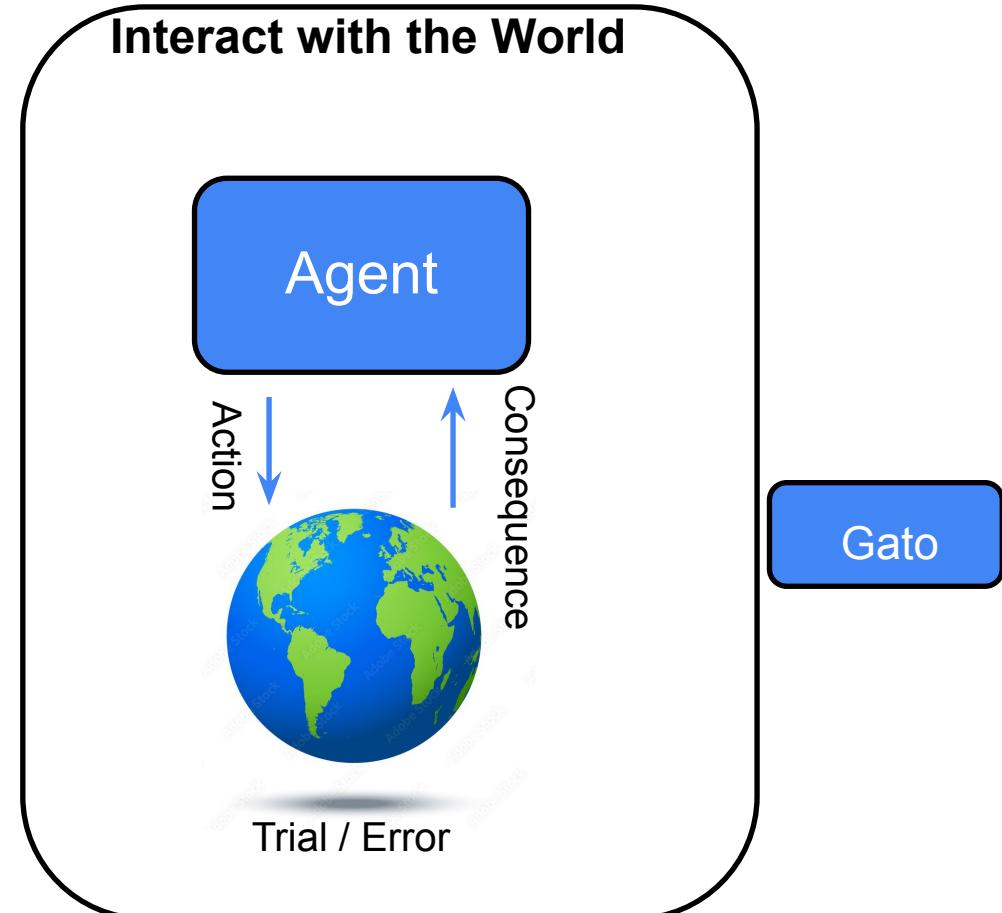


World Scopes



World Scope 4
Embodiment and Action

Interact with the World



WS4: How much knowledge is in an environment?



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

Is it every object affordance?

Blocks can stack.

Well, some blocks can stack.

Wheels can roll.

Wheels experience friction.

...

Is it every possible state or transition?

Car can have 3 pink block on top.

Car can have one green on top.

Car can slide forward or backward.

Top pink block can be removed.

...

It starts to seem silly to even imagining using language to enumerate the possibilities of an environment, even a tabletop with blocks. The affordances, states, and transitions we learn for planning are strikingly mundane to express verbally, even on something like WikiHow.

WS4: How much knowledge is in an environment?

Sec. 9.7 Embodied Agent

Navigation

Prompt:

Imagine that you are a home robot, and is asked to go to the kitchen to fetch something from the fridge. At the previous turn, you were at the position shown in the below image:



Note: Same image from the previous turn, resized to save space

The action taken in previous turn is: turn right and move forward towards the kitchen.

Now you are at the kitchen, this is what you can see from there. Can you see the fridge from here? What is your next action?



GPT-4V:

From this position, I can see the fridge on the right side of the kitchen. My next action would be to move forward and turn slightly to the right to approach the fridge.

Prompt:

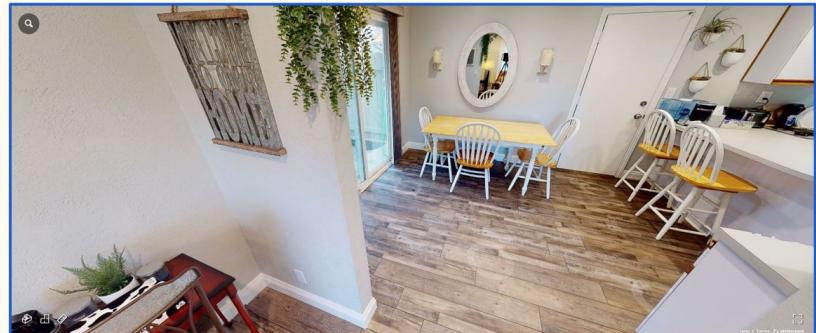
Imagine that you are a home robot, and is asked to go to the kitchen to fetch something from the fridge. At the previous turn, you were at the position shown in the below image:



Note: Same image from the previous turn, resized to save space

The action taken in previous turn is: turn right and move forward towards the hallway.

Now you are at the hallway, this is what you can see from there. Can you see the kitchen from here? What is your next action?



GPT-4V:

From the current position, I can see the kitchen on the right side. My next action would be to turn right and move forward towards the kitchen.

WS4: How much knowledge is in an environment?

Sec. 9.7 Embodied Agent

Navigation

Prompt:

Imagine that you are a home robot, and is asked to go to the kitchen to fetch something from the fridge. At the previous turn, you were at the position shown in the below image:



Note: Same image from the previous turn, resized to save space

The action taken in previous turn is: turn right and move forward towards the kitchen.

Now you are at the kitchen, this is what you can see from there. Can you see the fridge from here? What is your next action?



GPT-4V:

From this position, I can see the fridge on the right side of the kitchen. My next action would be to move forward and turn slightly to the right to approach the fridge.

Prompt:

Imagine that you are a home robot, and is asked to go to the kitchen to fetch something from the fridge. At the previous turn, you were at the position shown in the below image:



Note: Same image from the previous turn, resized to save space

The action taken in previous turn is: move forward and turn slightly to the right to approach the fridge.

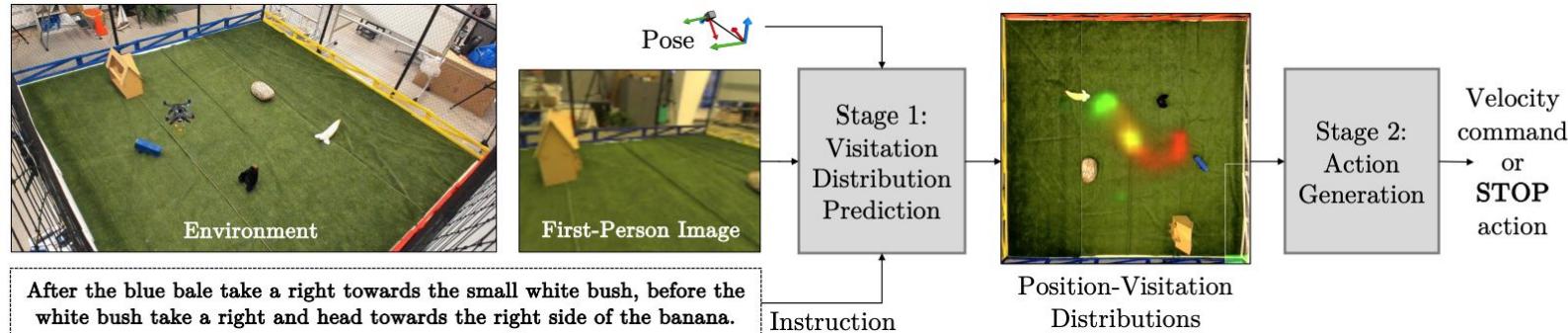
Now this is what you can see from there. What is your next action?



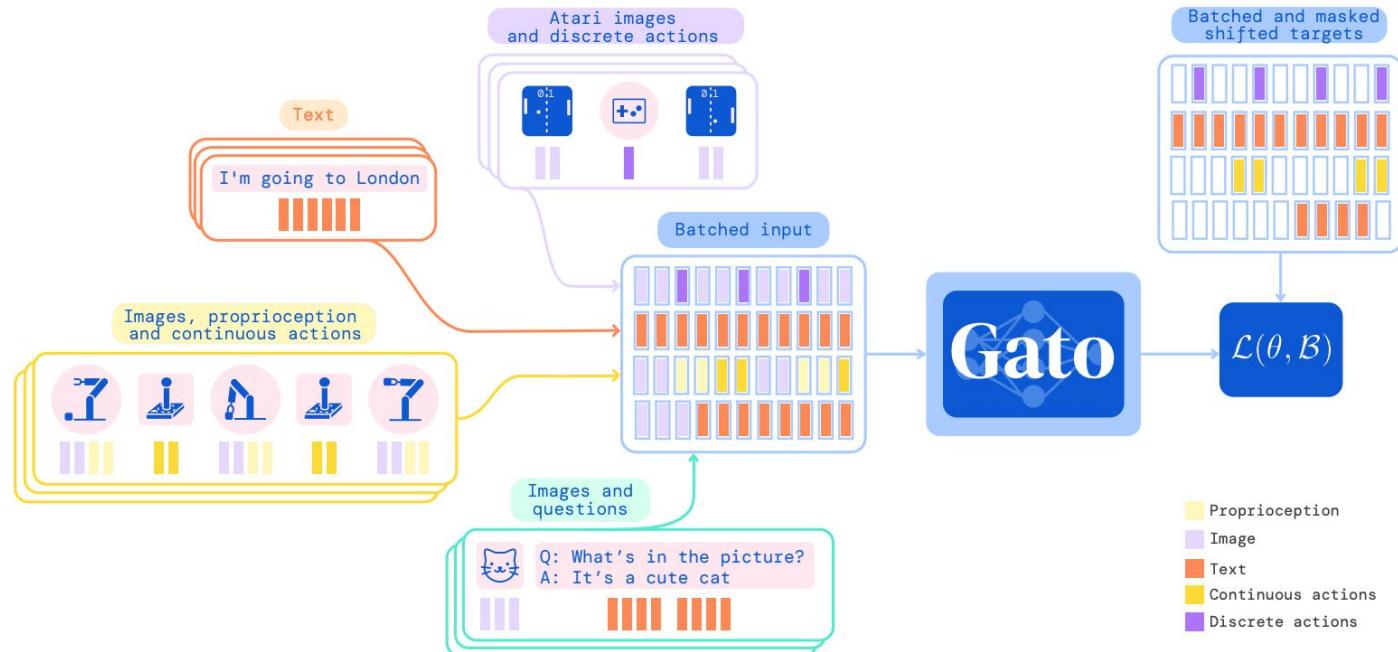
GPT-4V:

My next action would be to move forward and slightly to the left to align myself with the fridge door. Then, I would use my robotic arm to open the fridge door and retrieve the requested item.

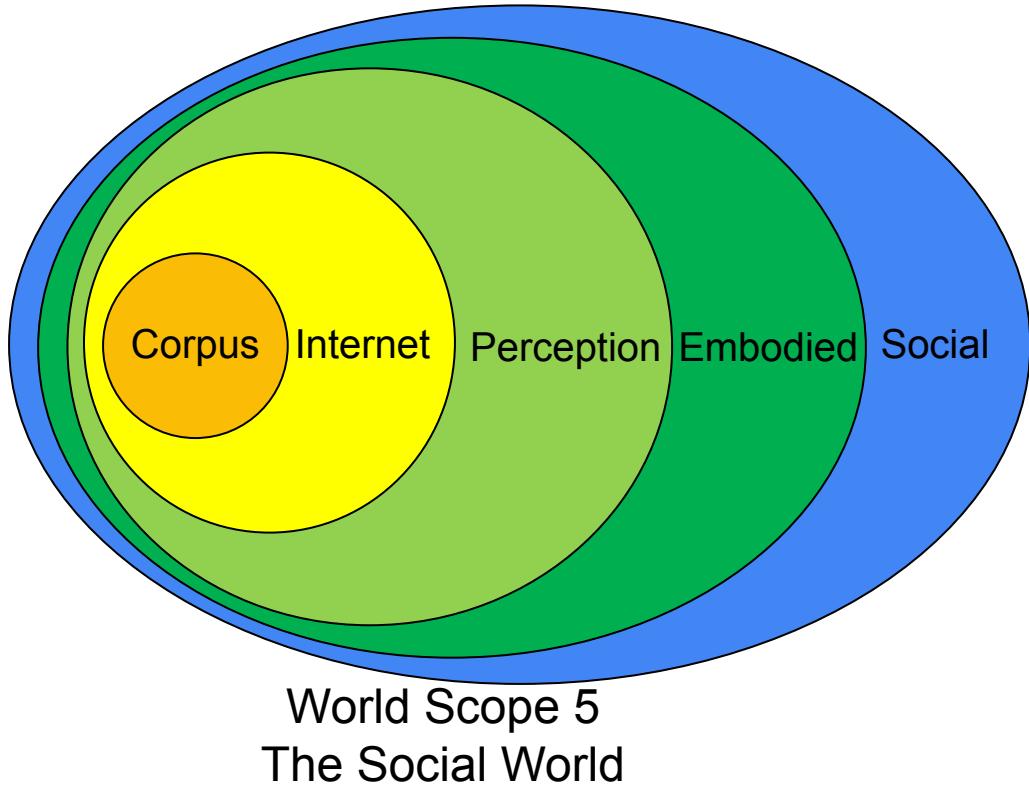
WS4: How much knowledge is in an environment?



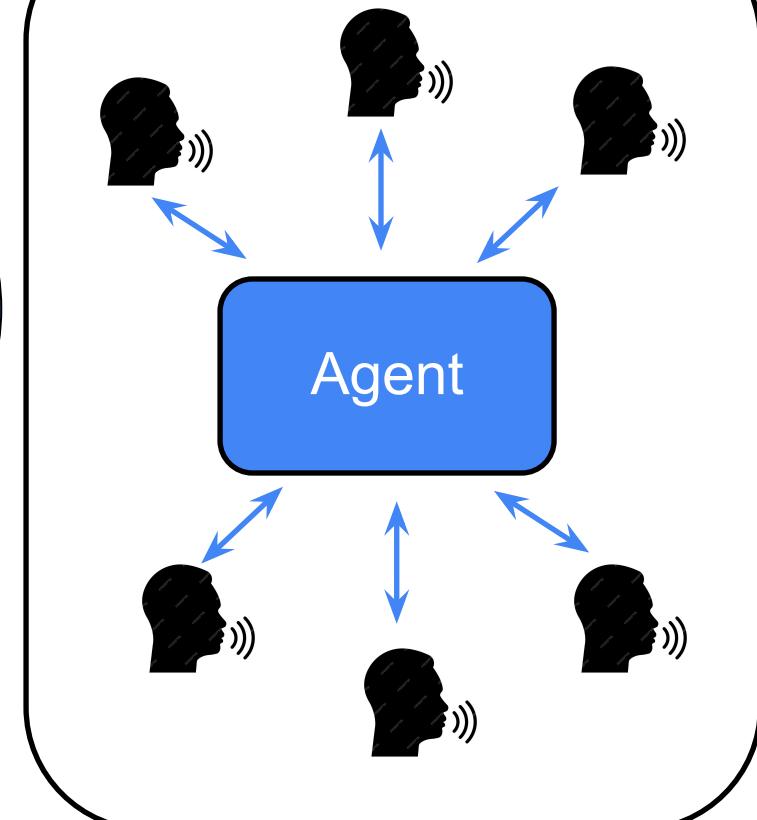
WS4: How much knowledge is in an environment?



World Scopes



Social Interaction



Paper 2:



REACT: Synergizing Reasoning and Acting in Language Models

Yao Teng

Concepts

ACTING

Physical or behavioral actions for a goal.

REASONING

Thinking, analyzing. Logic.

Concepts

ACTING

Physical or behavioral actions for a goal.

Task-specific actions

REASONING

Thinking, analyzing. Logic.

Verbal reasoning traces to support the actions and handle exceptions

Why SYNERGIZING ?

COT

- Reasons only.
- Static, black box, no update, hallucination and error propagation

Why SYNERGIZING ?

(1) Hotspot QA

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

(1a) Standard

Answer: iPod



(1b) CoT (Reason Only)

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

Answer: iPhone, iPad, iPod Touch

(1c) Act-Only

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control ...

Act 2: Search[Front Row]

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

Act 3: Search[Front Row

(software)]

Obs 3: Front Row is a discontinued media center software ...

Act 4: Finish[yes]

Reason/Act-only get wrong answer

Why SYNERGIZING ?

COT

- Reasons only.
- Static, black box, no update, hallucination and error propagation

Traditional approaches:

- No high-level goals or maintain a working memory

REACT

Combining reasoning and acting in a synergistic manner

REACT

- Augment the agent's action space with the space of language
 - LLM to interact with the environment (e.g. use Wikipedia search API)
 - Prompting LLM to generate reasoning traces in natural language
- A frozen PaLM-540B, few-shot in-context examples to generate both domain-specific actions and free-form language thoughts for task solving
- Each in-context example is a human trajectory of actions, thoughts, and environment observations to solve a task instance

REACT

- For the tasks where reasoning is of primary importance,
 - Alternate the generation of thoughts and actions so that the task-solving trajectory consists of multiple thought-action-observation steps.
- For decision making tasks that potentially involve a large number of actions,
 - Thoughts only need to appear sparsely in the most relevant positions of a trajectory, so let the language model decide the asynchronous occurrence of thoughts and actions for itself.

REACT

- Advantages
 - Intuitive and easy to design prompts
 - General and flexible
 - Performant and robust
 - Human aligned and controllable

REACT

- Prompt Template

Thought: ...

Action: ...

Observation: ...

... (Repeated many times)

REACT

(1) Hotspot QA

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

(1a) Standard

Answer: iPod



(1c) Act-Only

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control ...

Act 2: Search[Front Row]

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

Act 3: Search[Front Row (software)]

Obs 3: Front Row is a discontinued media center software ...

Act 4: Finish[yes]

(1d) ReAct (Reason + Act)

Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with.

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control introduced in October 2005 by Apple ... originally designed to control the Front Row media center program ...

Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search Front Row next and find what other device can control it.

Act 2: Search[Front Row]

Obs 2: Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports', 'Front Row (software)', ...]

Thought 3: Front Row is not found. I need to search Front Row (software) .

Act 3: Search[Front Row (software)]

Obs 3: Front Row is a discontinued media center software ...

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

Act 4: Finish[keyboard function keys]



REACT

- decomposing task goals and create action plans

(1) Hotspot QA

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

(1d) ReAct (Reason + Act)

Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with.

Act 1: `Search[Apple Remote]`

Obs 1: The Apple Remote is a remote control introduced in October 2005 by Apple ... originally designed to control the `Front Row media center` program ...

Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search `Front Row` next and find what other device can control it.

Act 2: `Search[Front Row]`

Obs 2: Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports', 'Front Row (software)', ...]

Thought 3: Front Row is not found. I need to search `Front Row (software)` .

Act 3: `Search[Front Row (software)]`

Obs 3: Front Row is a discontinued media center software ...

Thought 4: `Front Row (software)` is controlled by an Apple Remote or the keyboard function keys. So the answer is keyboard function keys.

Act 4: `Finish[keyboard function keys]`



REACT

(1) Hotspot QA

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

(1d) ReAct (Reason + Act)

Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with.

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control introduced in October 2005 by Apple ... originally designed to control the Front Row media center program ...

Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search Front Row next and find what other device can control it.

Act 2: Search[Front Row]

Obs 2: Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports', 'Front Row (software)', ...]

Thought 3: Front Row is not found. I need to search Front Row (software) .

Act 3: Search[Front Row (software)]

Obs 3: Front Row is a discontinued media center software ...

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

Act 4: Finish[keyboard function keys]

- extracting important parts from observations



REACT

(1) Hotspot QA

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

(1d) ReAct (Reason + Act)

Thought 1: I need to search Apple Remote and find the program it was originally designed to interact with.

Act 1: Search[Apple Remote]

Obs 1: The Apple Remote is a remote control introduced in October 2005 by Apple ... originally designed to control the Front Row media center program ...

Thought 2: Apple Remote was originally designed to control the Front Row media center program. I need to search Front Row next and find what other device can control it.

Act 2: Search[Front Row]

Obs 2: Could not find [Front Row]. Similar: ['Front Row Seat to Earth', 'Front Row Motorsports', 'Front Row (software)', ...]

Thought 3: Front Row is not found. I need to search Front Row (software) .

Act 3: Search[Front Row (software)]

Obs 3: Front Row is a discontinued media center software ...

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

Act 4: Finish[keyboard function keys]

- handle exceptions and adjust action plans



REACT

(2) AlfWorld

You are in the middle of a room. Looking quickly around you, you see a cabinet 6, a cabinet 1, a coffee machine 1, a countertop 3, a stove burner 1, and a toaster 1.

Your task is to: Put some pepper shaker on a drawer.

(2a) Act-Only

Act 1: Go to drawer 1

Obs 1: The drawer 1 is closed.

Act 2: Open drawer 1

Obs 2: You open the drawer 1. The drawer 1 is open. In it, you see a dishponge 2, and a spoon 1.

Act 3: Go to sinkbasin 1

Obs 3: On the sinkbasin 1, you see a dishponge 3, a spatula 1, and a spoon 2.

Act 4: Take peppershaker 1 from sinkbasin 1

Obs 4: Nothing happens.

Act 5: Take peppershaker 1 from sinkbasin 1

Obs 5: Nothing happens.



(2b) ReAct (Reason + Act)

Act 1: Think[First I need to find a pepper shaker...more likely to appear in cabinets (1-6), countertops (1-3), ...]

Act 2: Go to cabinet 1

Obs 2: On the cabinet 1, you see a vase 2.

(Here the agent go to cabinet 1, then cabinet 2, 3, then countertop 1 and 2)

Act 6: Go to countertop 3

Obs 6: On the countertop 3, you see a apple 1, a bread 1, a pepper shaker 1, and a vase 3.

Act 7: Take pepper shaker 1 from countertop 3

Obs 7: You pick up the pepper shaker 1 from the countertop 3.

Act 8: Think[Now I find a pepper shaker 1. Next, I need to put it in/on drawer 1.]

Act 9: Go to drawer 1

Obs 9: Drawer 1 is closed.

Act 10: Open drawer 1

Obs 10: You open Drawer 1 ...

Act 11: Put pepper shaker 1 in/on drawer 1

Obs 11: You put pepper shaker 1 in/on the drawer 1.



REACT

(2) AlfWorld

You are in the middle of a room. Looking quickly around you, you see a cabinet 6, a cabinet 1, a coffee machine 1, a countertop 3, a stove burner 1, and a toaster 1.

Your task is to: Put some pepper shaker on a drawer.

(2a) Act-Only

Act 1: Go to drawer 1

Obs 1: The drawer 1 is closed.

Act 2: Open drawer 1

Obs 2: You open the drawer 1. The drawer 1 is open. In it, you see a dishponge 2, and a spoon 1.

Act 3: Go to sinkbasin 1

Obs 3: On the sinkbasin 1, you see a dishponge 3, a spatula 1, and a spoon 2.

Act 4: Take peppershaker 1 from sinkbasin 1

Obs 4: Nothing happens.

Act 5: Take peppershaker 1 from sinkbasin 1

Obs 5: Nothing happens.



(2b) ReAct (Reason + Act)

Act 1: Think[First I need to find a pepper shaker...more likely to appear in cabinets (1-6), countertops (1-3), ...]

Act 2: Go to cabinet 1

Obs 2: On the cabinet 1, you see a vase 2.

(Here the agent go to cabinet 1, then cabinet 2, 3, then countertop 1 and 2)

Act 6: Go to countertop 3

Obs 6: On the countertop 3, you see a apple 1, a bread 1, a pepper shaker 1, and a vase 3.

Act 7: Take pepper shaker 1 from countertop 3

Obs 7: You pick up the pepper shaker 1 from the countertop 3.

Act 8: Think[Now I find a pepper shaker 1. Next, I need to put it in/on drawer 1.]

Act 9: Go to drawer 1

Obs 9: Drawer 1 is closed.

Act 10: Open drawer 1

Obs 10: You open Drawer 1 ...

Act 11: Put pepper shaker 1 in/on drawer 1

Obs 11: You put pepper shaker 1 in/on the drawer 1.



REACT

- decomposing task goals and create action plans
- commonsense injection

(2) AlfWorld

You are in the middle of a room. Looking quickly around you, you see a cabinet 6, a cabinet 1, a coffee machine 1, a countertop 3, a stove burner 1, and a toaster 1.
Your task is to: Put some pepper shaker on a drawer.

(2b) ReAct (Reason + Act)

Act 1: Think[First I need to find a pepper shaker...more likely to appear in cabinets (1-6), countertops (1-3), ...]

Act 2: Go to cabinet 1

Obs 2: On the cabinet 1, you see a vase 2.

(Here the agent go to cabinet 1, then cabinet 2, 3, then countertop 1 and 2)

Act 6: Go to countertop 3

Obs 6: On the countertop 3, you see a apple 1, a bread 1, a pepper shaker 1, and a vase 3.

Act 7: Take pepper shaker 1 from countertop 3

Obs 7: You pick up the pepper shaker 1 from the countertop 3.

Act 8: Think[Now I find a pepper shaker 1. Next, I need to put it in/on drawer 1.]

Act 9: Go to drawer 1

Obs 9: Drawer 1 is closed.

Act 10: Open drawer 1

Obs 10: You open Drawer 1 ...

Act 11: Put pepper shaker 1 in/on drawer 1

Obs 11: You put pepper shaker 1 in/on the drawer 1.



REACT

(2) AlfWorld

You are in the middle of a room. Looking quickly around you, you see a cabinet 6, a cabinet 1, a coffee machine 1, a countertop 3, a stove burner 1, and a toaster 1.

Your task is to: Put some pepper shaker on a drawer.

(2b) ReAct (Reason + Act)

Act 1: Think[First I need to find a pepper shaker...more likely to appear in cabinets (1-6), countertops (1-3), ...]

Act 2: Go to cabinet 1

Obs 2: On the cabinet 1, you see a vase 2.

(Here the agent go to cabinet 1, then cabinet 2, 3, then countertop 1 and 2)

Act 6: Go to countertop 3

Obs 6: On the countertop 3, you see a apple 1, a bread 1, a pepper shaker 1, and a vase 3.

Act 7: Take pepper shaker 1 from countertop 3

Obs 7: You pick up the pepper shaker 1 from the countertop 3.

Act 8: Think[Now I find a pepper shaker 1. Next, I need to put it in/on drawer 1.]

Act 9: Go to drawer 1

Obs 9: Drawer 1 is closed.

Act 10: Open drawer 1

Obs 10: You open Drawer 1 ...

Act 11: Put pepper shaker 1 in/on drawer 1

Obs 11: You put pepper shaker 1 in/on the drawer 1.



Experiments

Prompt Method ^a	HotpotQA (EM)	Fever (Acc)
Standard	28.7	57.1
CoT (Wei et al., 2022)	29.4	56.3
CoT-SC (Wang et al., 2022a)	33.4	60.4
Act	25.7	58.9
ReAct	27.4	60.9
CoT-SC → ReAct	34.2	64.6
ReAct → CoT-SC	35.1	62.0
Supervised SoTA^b	67.5	89.5

Table 1: PaLM-540B prompting results on HotpotQA and Fever.

^a HotpotQA EM is 27.1, 28.9, 33.8 for Standard, CoT, CoT-SC in Wang et al. (2022b).

^b (Zhu et al., 2021; Lewis et al., 2020)

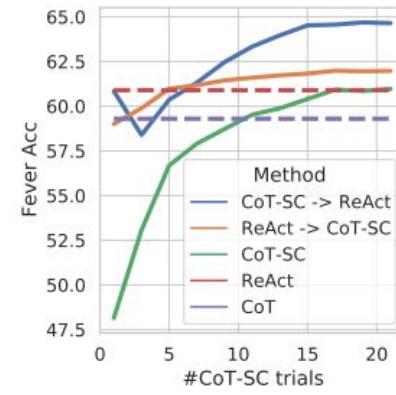
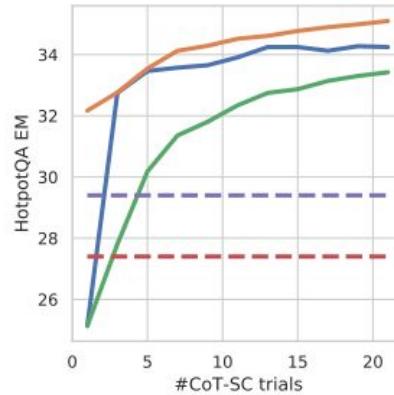


Figure 2: PaLM-540B prompting results with respect to number of CoT-SC samples used.

Experiments

	Type	Definition	ReAct	CoT
Success	True positive	Correct reasoning trace and facts	94%	86%
	False positive	Hallucinated reasoning trace or facts	6%	14%
Failure	Reasoning error	Wrong reasoning trace (including failing to recover from repetitive steps)	47%	16%
	Search result error	Search return empty or does not contain useful information	23%	-
	Hallucination	Hallucinated reasoning trace or facts	0%	56%
	Label ambiguity	Right prediction but did not match the label precisely	29%	28%

Table 2: Types of success and failure modes of ReAct and CoT on HotpotQA, as well as their percentages in randomly selected examples studied by human.

Experiments

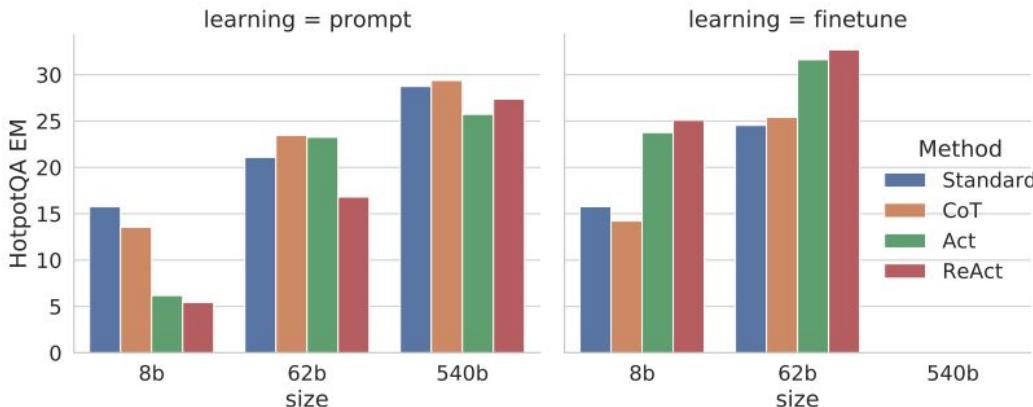


Figure 3: Scaling results for prompting and finetuning on HotPotQA with ReAct (ours) and baselines.

Experiments

Method	Pick	Clean	Heat	Cool	Look	Pick 2	All	Method	Score	SR
Act (best of 6)	88	42	74	67	72	41	45	Act	62.3	30.1
ReAct (avg)	65	39	83	76	55	24	57	ReAct	66.6	40.0
ReAct (best of 6)	92	58	96	86	78	41	71			
ReAct-IM (avg)	55	59	60	55	23	24	48	IL	59.9	29.1
ReAct-IM (best of 6)	62	68	87	57	39	33	53	IL+RL	62.4	28.7
BUTLER _g (best of 8)	33	26	70	76	17	12	22			
BUTLER (best of 8)	46	39	74	100	22	24	37	Human Expert	82.1	59.6

Table 3: AlfWorld task-specific success rates (%). BUTLER and BUTLER_g results are from Table 4 of Shridhar et al. (2020b). All methods use greedy decoding, except that BUTLER uses beam search.

Table 4: Score and success rate (SR) on Webshop. IL/IL+RL taken from Yao et al. (2022).

Questions to Discuss

- How to empower LLM agents with fine-grained perception capabilities? For example, distinguish between 1 meter and 2 meters.
- Can LLM agents create another agent?
- Will something new happen if multimodal generative models are integrated into the agent system?

Thank you.