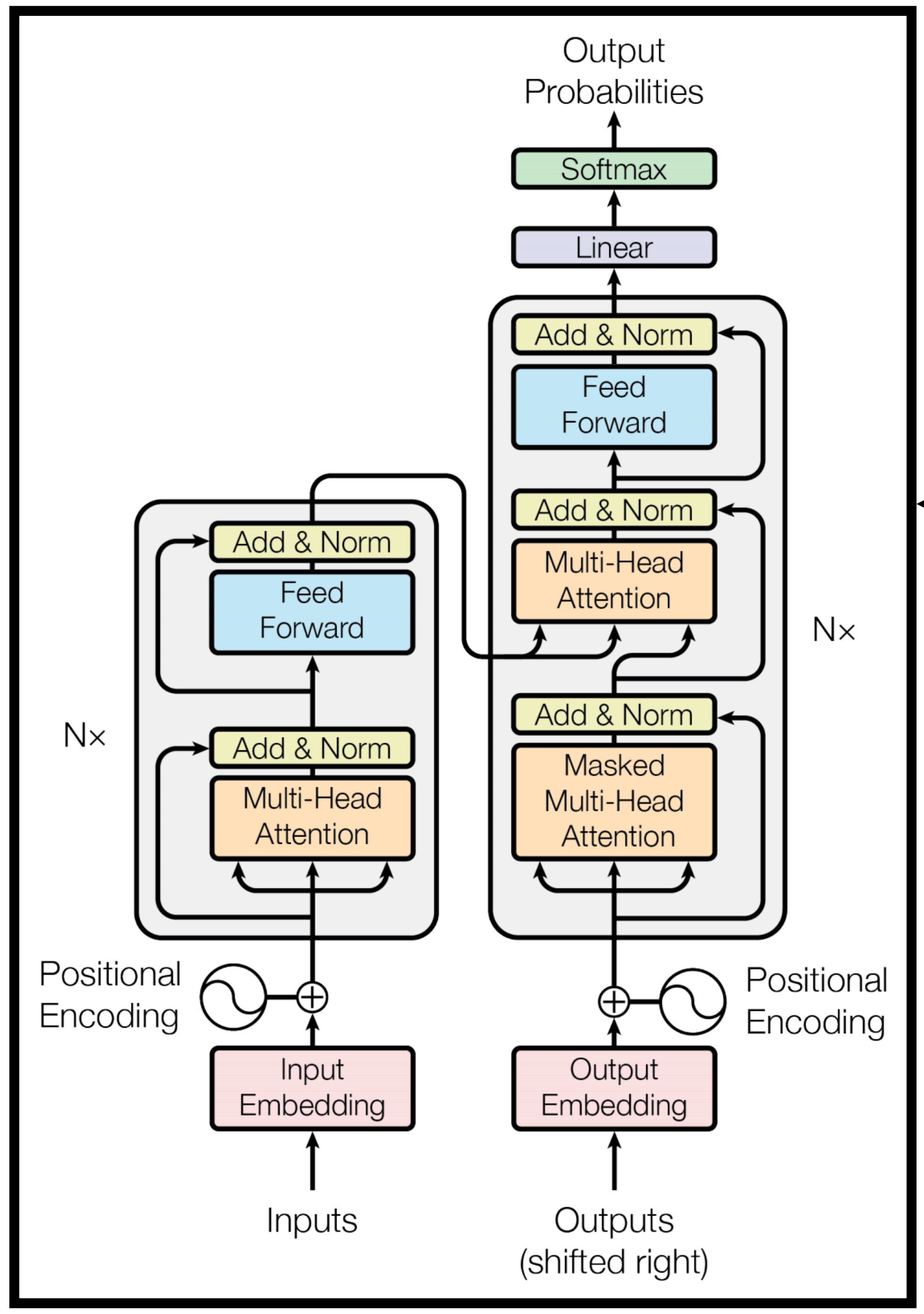
Guest Lecture GaTech CS 4644 / 7643 Deep Learning

LLM Agents
extensions of LLMs or start of something wonderful?

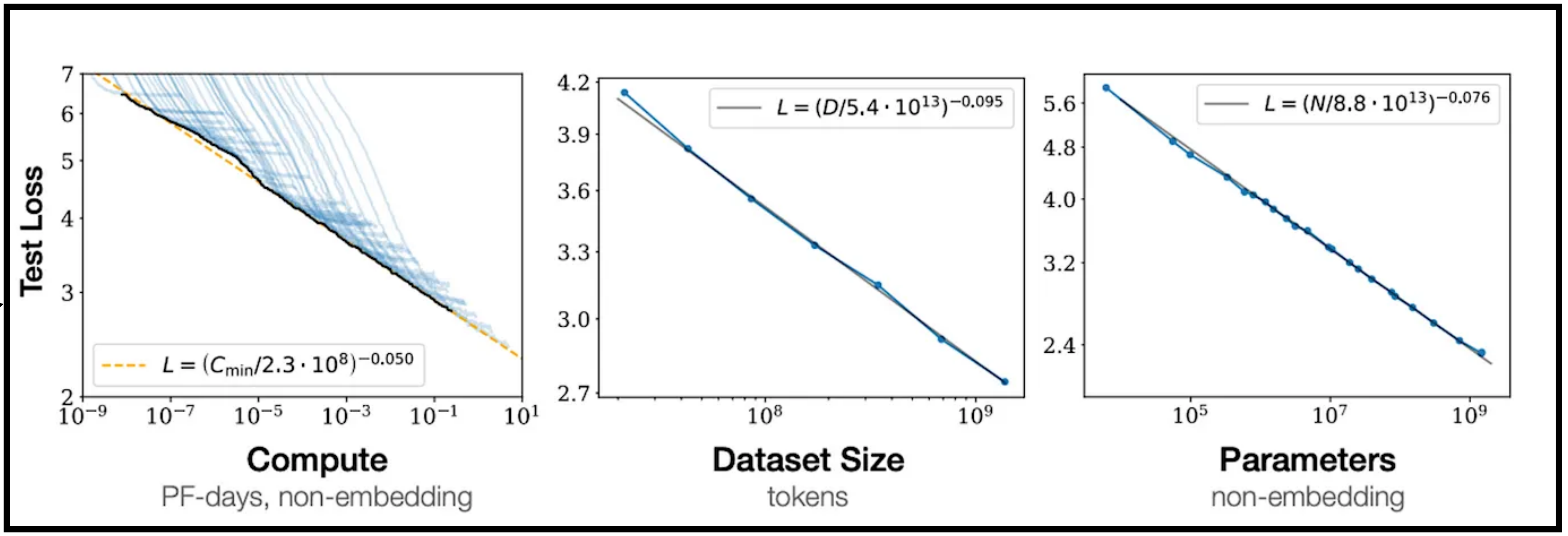
Hao Zhu

https://zhuhao.me



Transformer

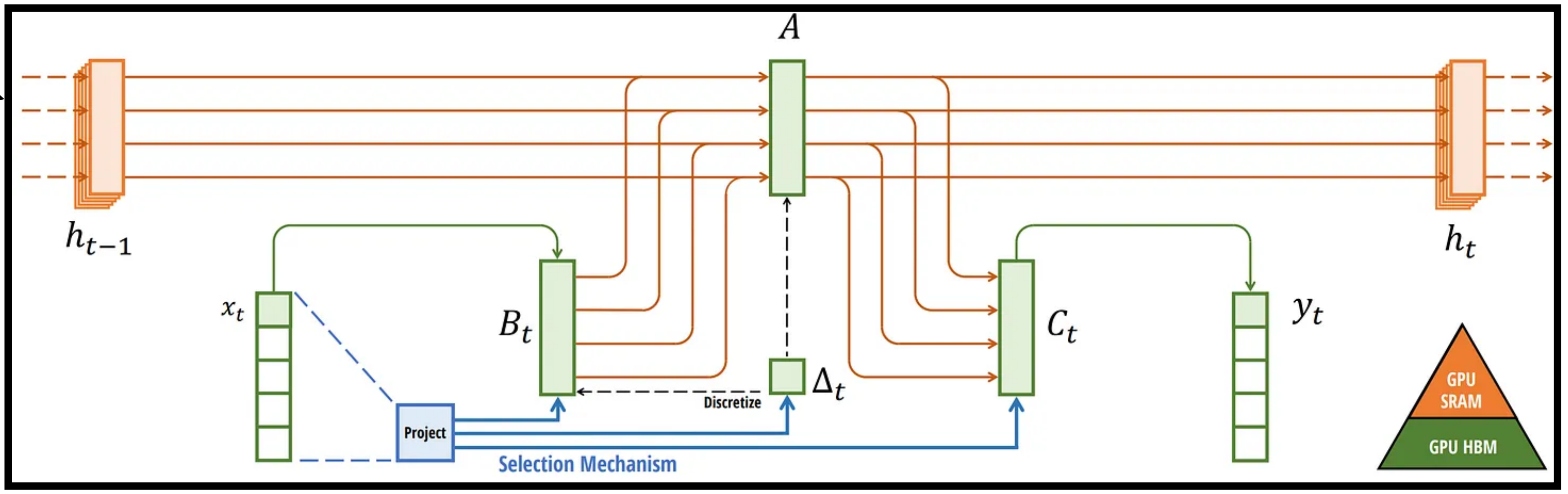
LLM Agent



large

or Mamba

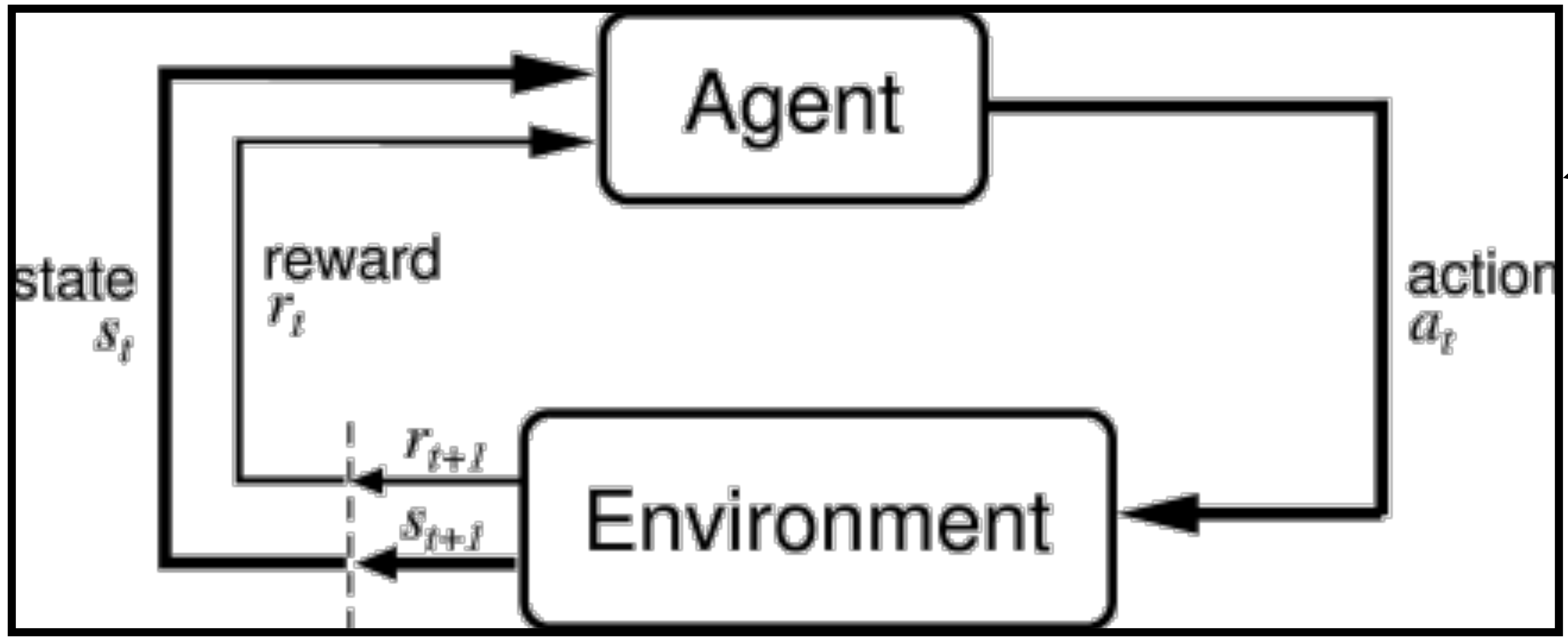
& other efficient archs



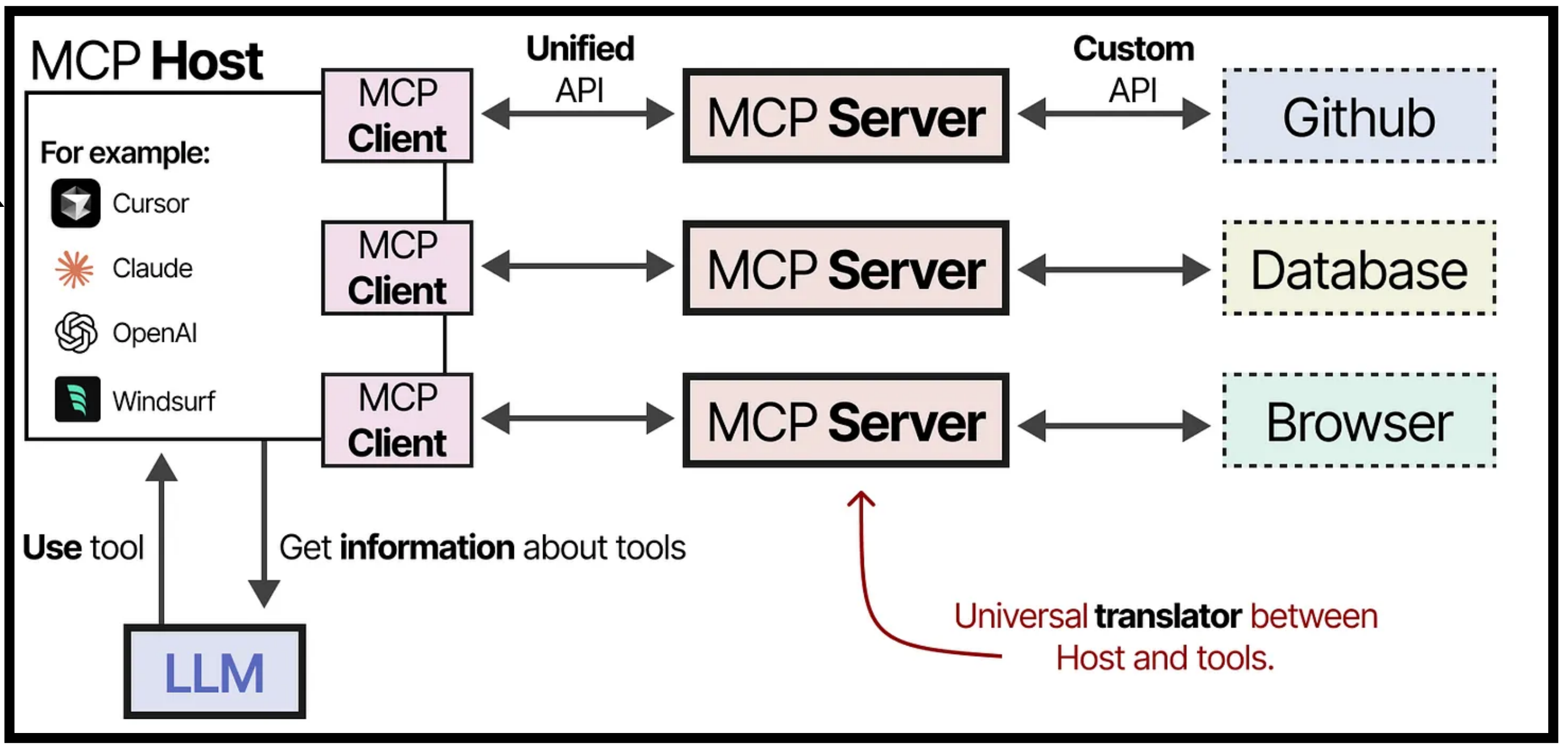
assistants?

LLM Agent

models connected to tools/ memory etc.



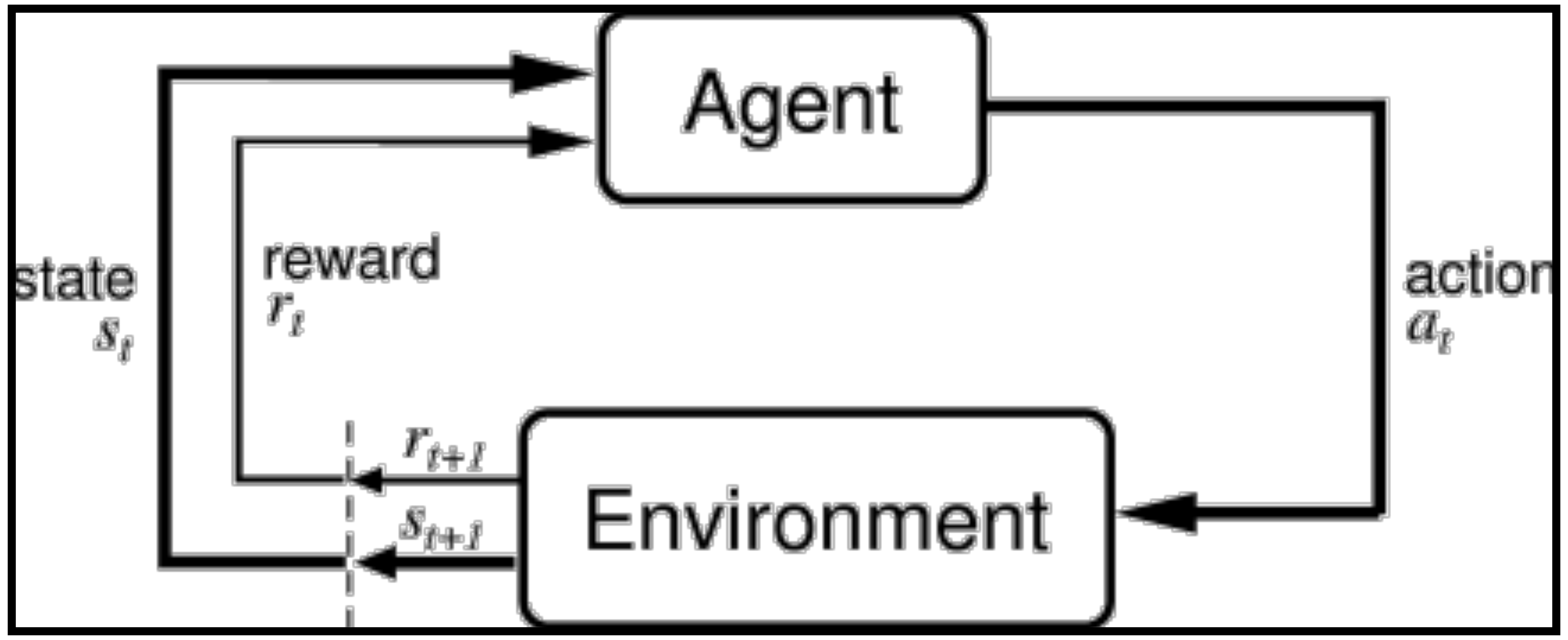
decision makers



LLM Agent = ?

agent thinking

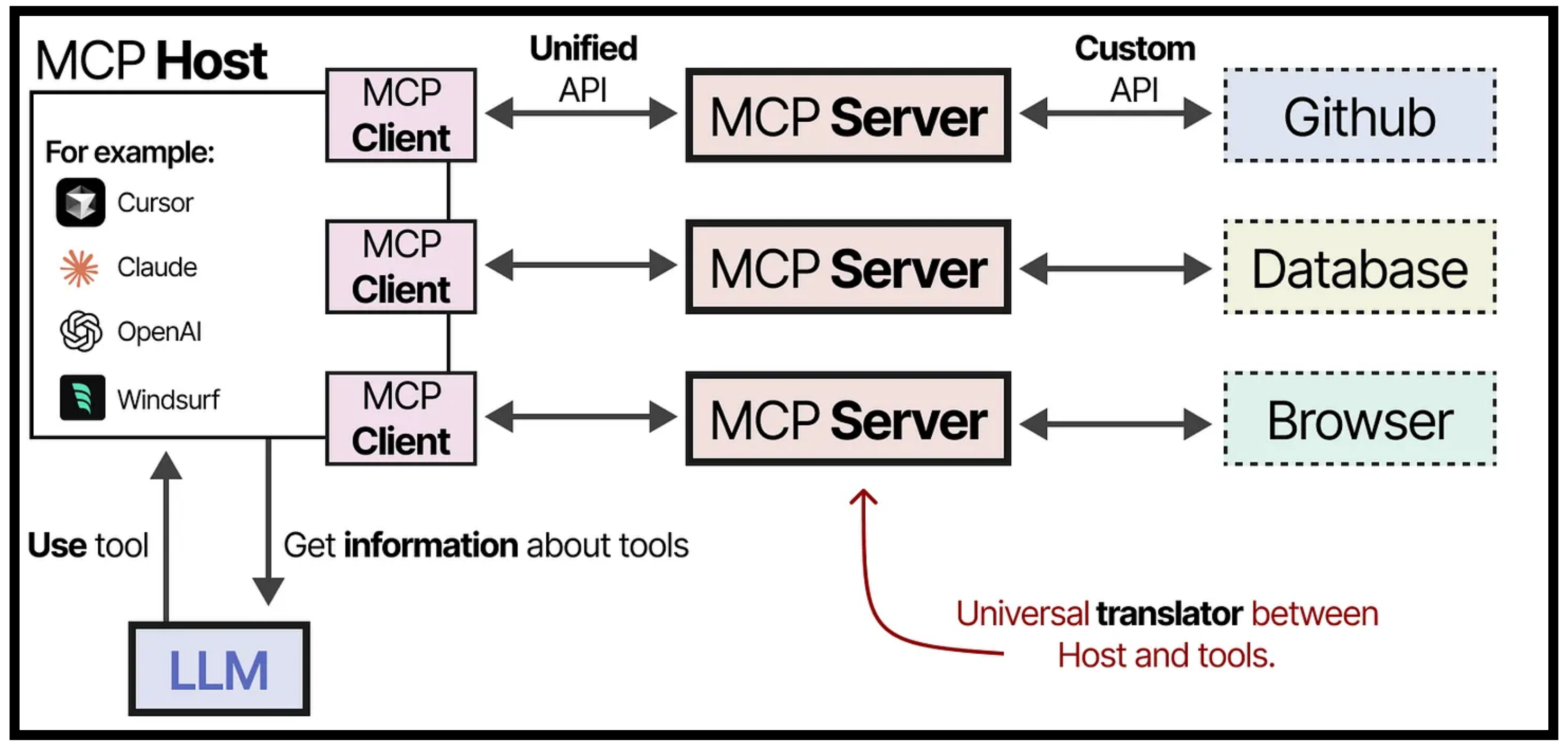
decision makers powered by LLMs



or

LLM thinking

LLMs connected to tools/memory etc.



two **useful** thinking tools

🤖🪄

agent thinking

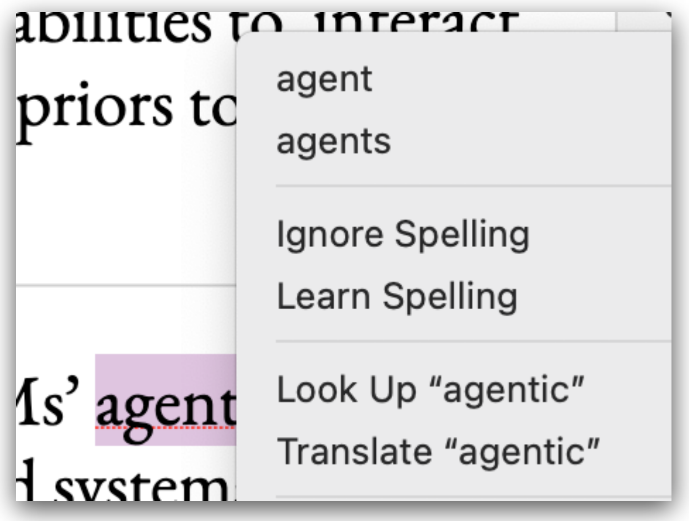
How to improve the models' capabilities to interact with the world? Are LLMs good priors to start with?

🧠🪄

LLM Thinking

How to make full use of the LLMs' agentic capabilities? What algorithms and systems we should on top of them?

Yeah, I know. "Agentic" is weird.



“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn...

👩‍💻

coding

💾

memory

⚡

in-context learning

⚓

grounding

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn…

👩‍💻

coding

💾

memory

⚡

in-context learning

⚓

grounding

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn…

👩‍💻

coding

💾

memory

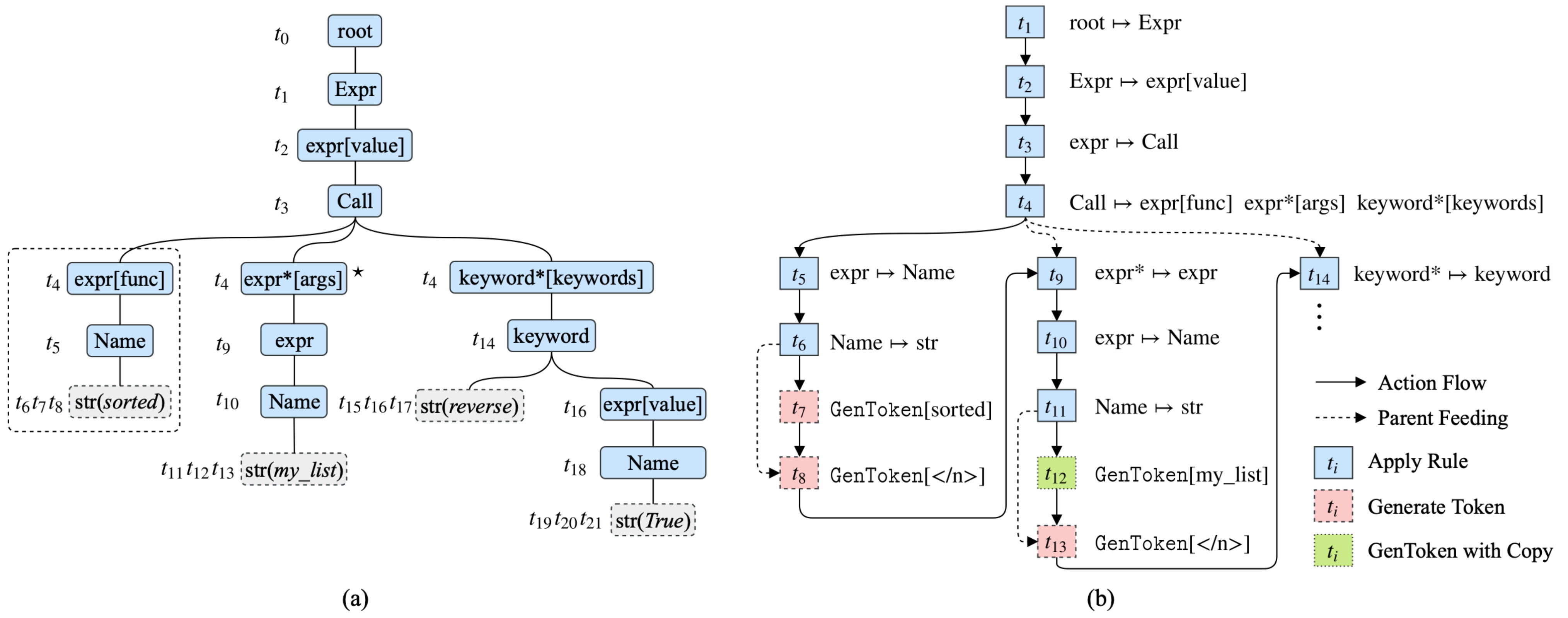
⚡

in-context learning

⚓

grounding

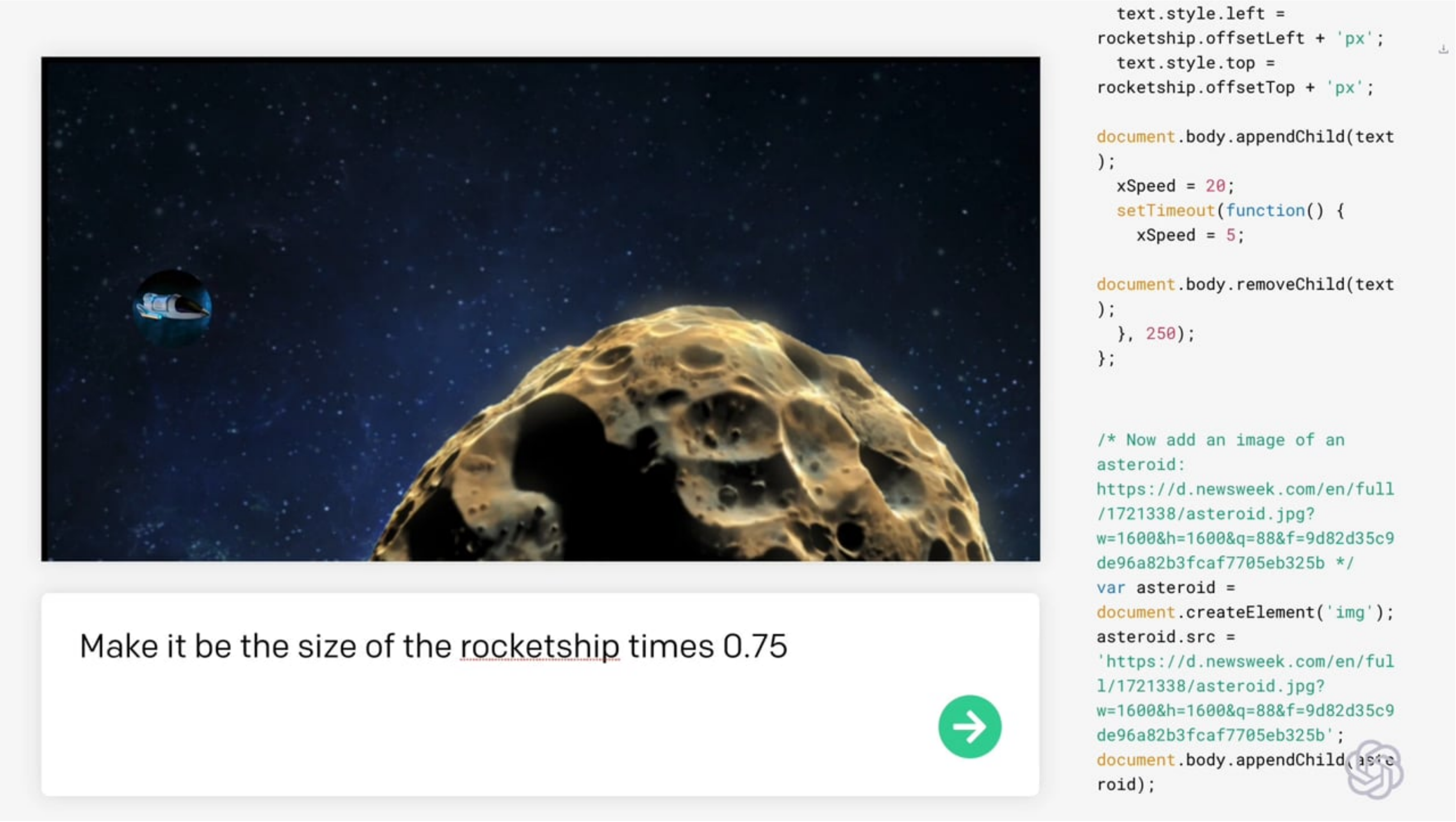
coding — LLM doing surprisingly well



Input: `sort my_list in descending order`

Code: `sorted(my_list, reverse=True)`

A syntactic neural model for general-purpose code generation



Make it be the size of the rocketship times 0.75

```
  text.style.left =
rocketship.offsetLeft + 'px';
  text.style.top =
rocketship.offsetTop + 'px';

document.body.appendChild(text
);
  xSpeed = 20;
  setTimeout(function() {
    xSpeed = 5;

document.body.removeChild(text
);
  }, 250);
};


/* Now add an image of an
asteroid:
https://d.newsweek.com/en/full
/1721338/asteroid.jpg?
w=1600&h=1600&q=88&f=9d82d35c9
de96a82b3fcaf7705eb325b */
var asteroid =
document.createElement('img');
asteroid.src =
'https://d.newsweek.com/en/ful
l/1721338/asteroid.jpg?
w=1600&h=1600&q=88&f=9d82d35c9
de96a82b3fcaf7705eb325b';
document.body.appendChild(aste
roid);
```

pre-LLM NL2Code

semantic parsing → AST

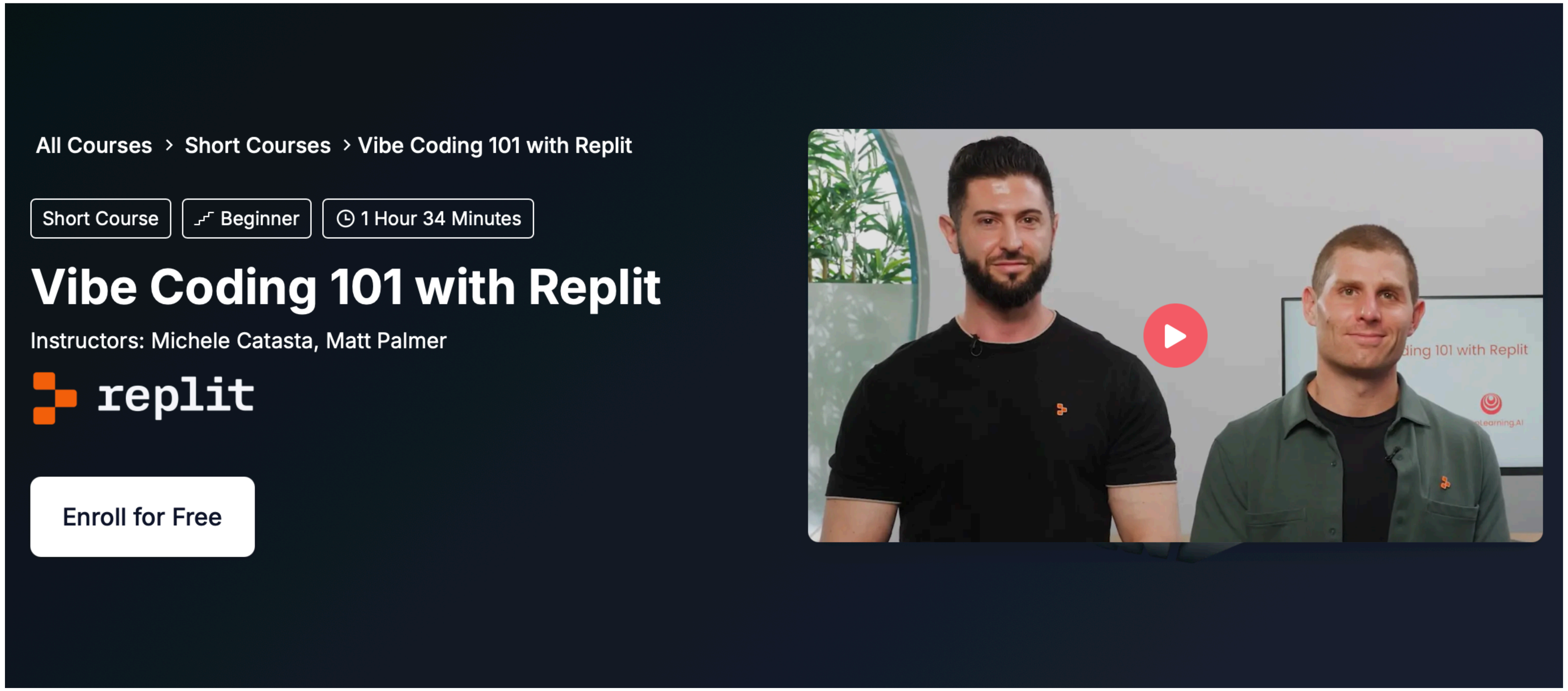
(Yin and Neubig, 2017)

LLM NL2Code

instruction following

(OpenAI Codex, 2021)

coding — and they got even better now

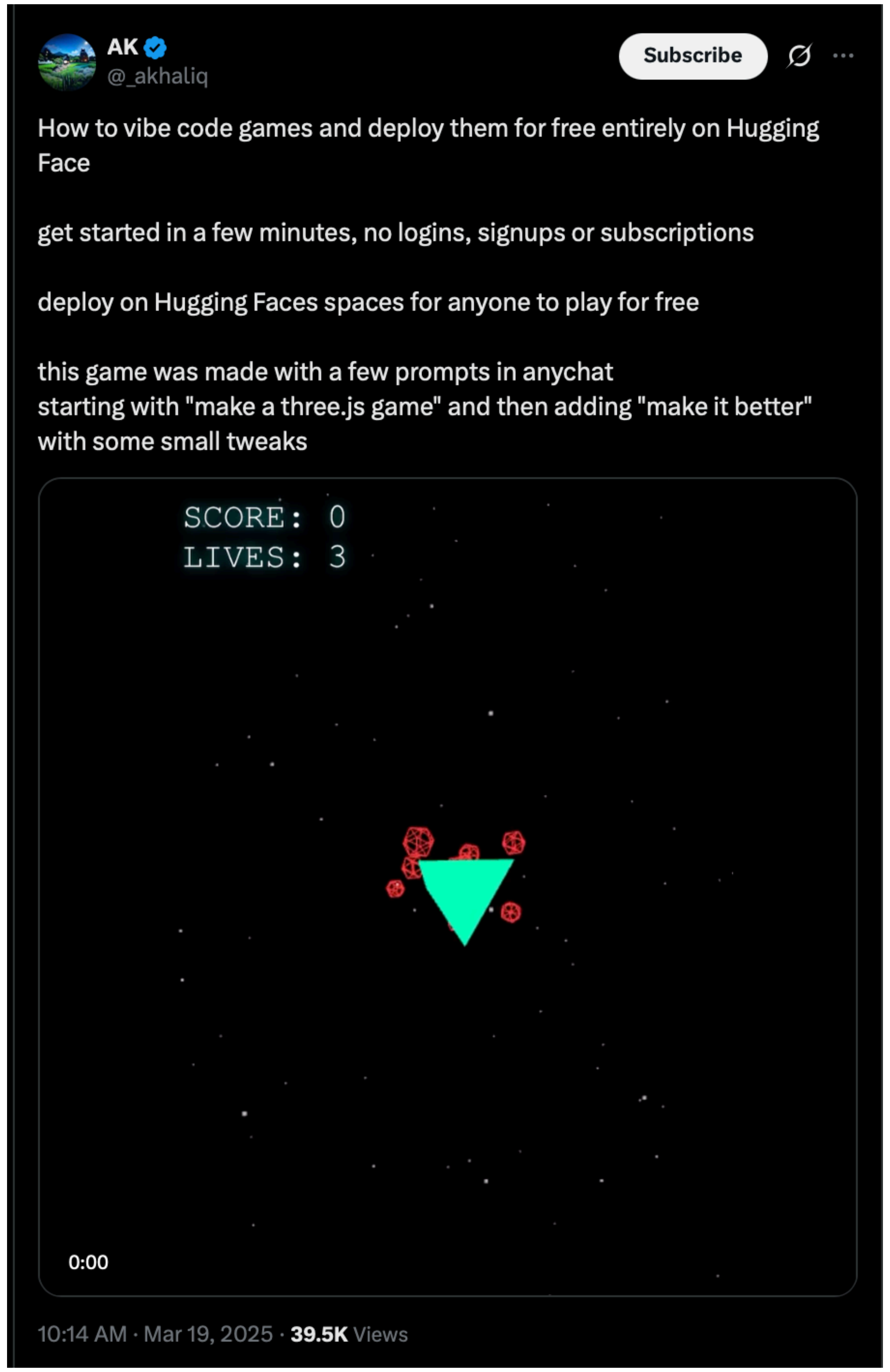


All Courses > Short Courses > Vibe Coding 101 with Replit

Short Course | Beginner | 1 Hour 34 Minutes

Vibe Coding 101 with Replit

Instructors: Michele Catasta, Matt Palmer



AK @_akhaliq

How to vibe code games and deploy them for free entirely on Hugging Face

get started in a few minutes, no logins, signups or subscriptions

deploy on Hugging Faces spaces for anyone to play for free

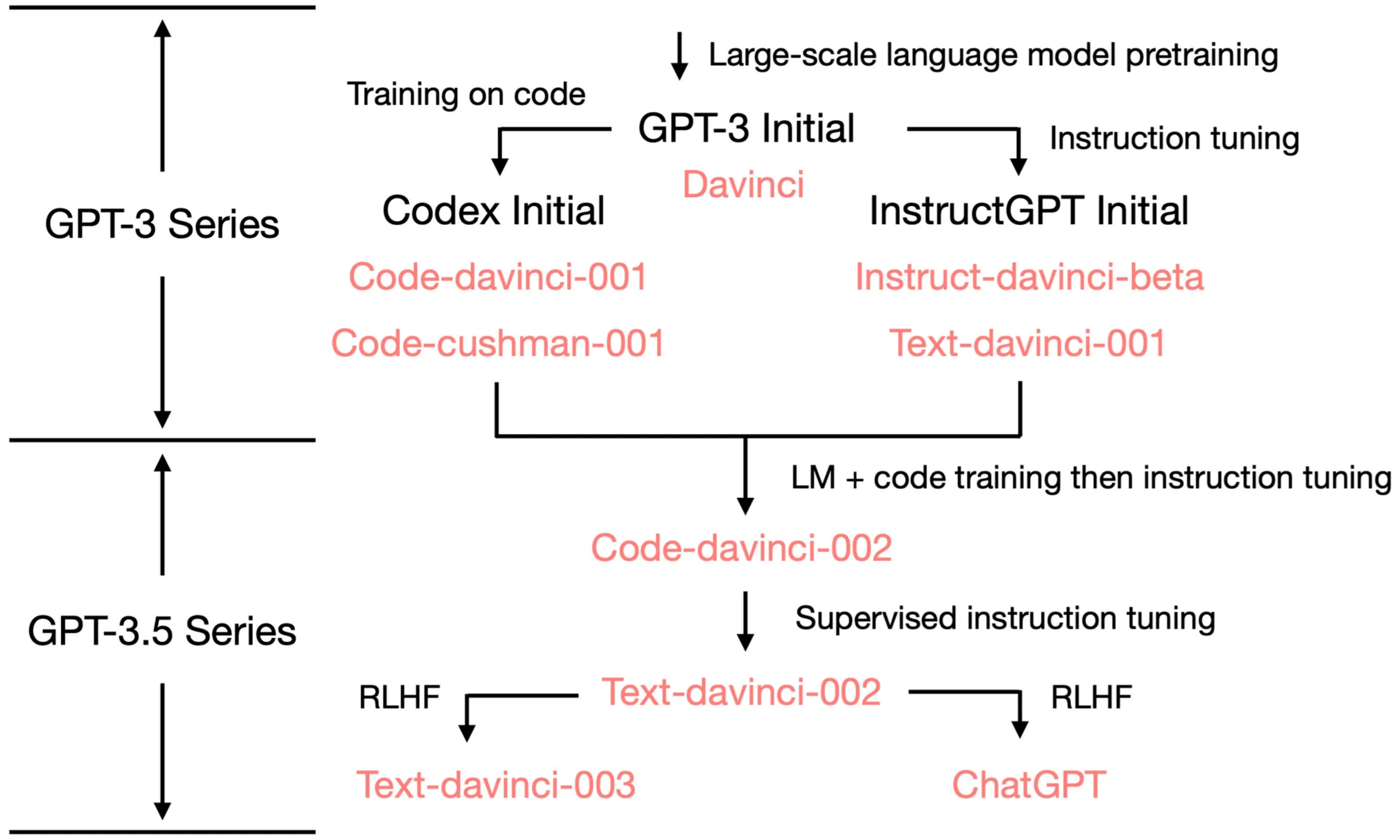
this game was made with a few prompts in anychat
starting with "make a three.js game" and then adding "make it better"
with some small tweaks

10:14 AM · Mar 19, 2025 · **39.5K** Views

Vibe Coding

ask LLM to code and it just works.

coding capabilities (con't)



GPT-3 Series

GPT-3.5 Series

Large-scale language model pretraining

GPT-3 Initial
Davinci

Training on code

Codex Initial
Code-davinci-001
Code-cushman-001

Instruction tuning

InstructGPT Initial
Instruct-davinci-beta
Text-davinci-001

LM + code training then instruction tuning

Code-davinci-002

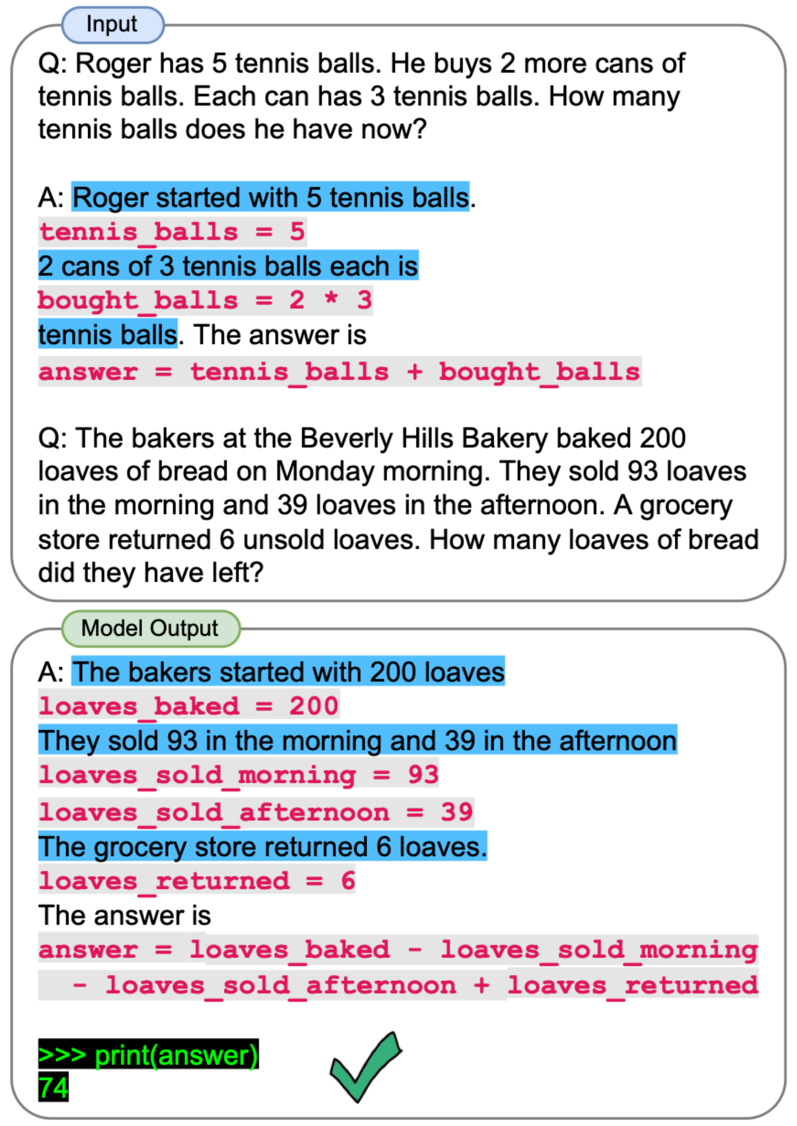
Supervised instruction tuning

Text-davinci-002

RLHF
Text-davinci-003

RLHF
ChatGPT

How does GPT Obtain its Ability? Tracing Emergent Abilities of Language Models to their Source



Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

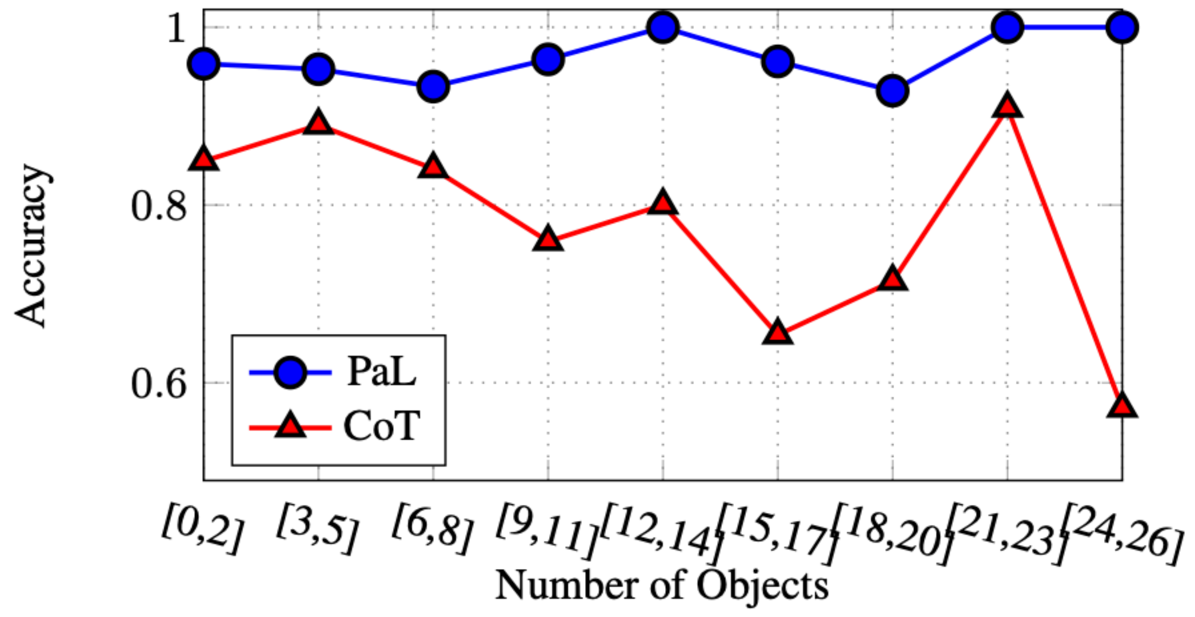
A: Roger started with 5 tennis balls.
`tennis_balls = 5`
2 cans of 3 tennis balls each is
`bought_balls = 2 * 3`
tennis balls. The answer is
`answer = tennis_balls + bought_balls`

Q: The bakers at the Beverly Hills Bakery baked 200 loaves of bread on Monday morning. They sold 93 loaves in the morning and 39 loaves in the afternoon. A grocery store returned 6 unsold loaves. How many loaves of bread did they have left?

Model Output

A: The bakers started with 200 loaves
`loaves_baked = 200`
They sold 93 in the morning and 39 in the afternoon
`loaves_sold_morning = 93`
`loaves_sold_afternoon = 39`
The grocery store returned 6 loaves.
`loaves_returned = 6`
The answer is
```
answer = loaves_baked - loaves_sold_morning
  - loaves_sold_afternoon + loaves_returned
```

```
>>> print(answer)
74
```
✓



PAL: Program-aided Language Models

"early" history

code as a detour

code in the pertaining data

prompt to code for QA

(Fu et al, 2022)

(Gao et al, 2023)

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn...

👩‍💻

coding

💾

memory

⚡

in-context learning

⚓

grounding

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn...

👩‍💻

coding

💾

memory

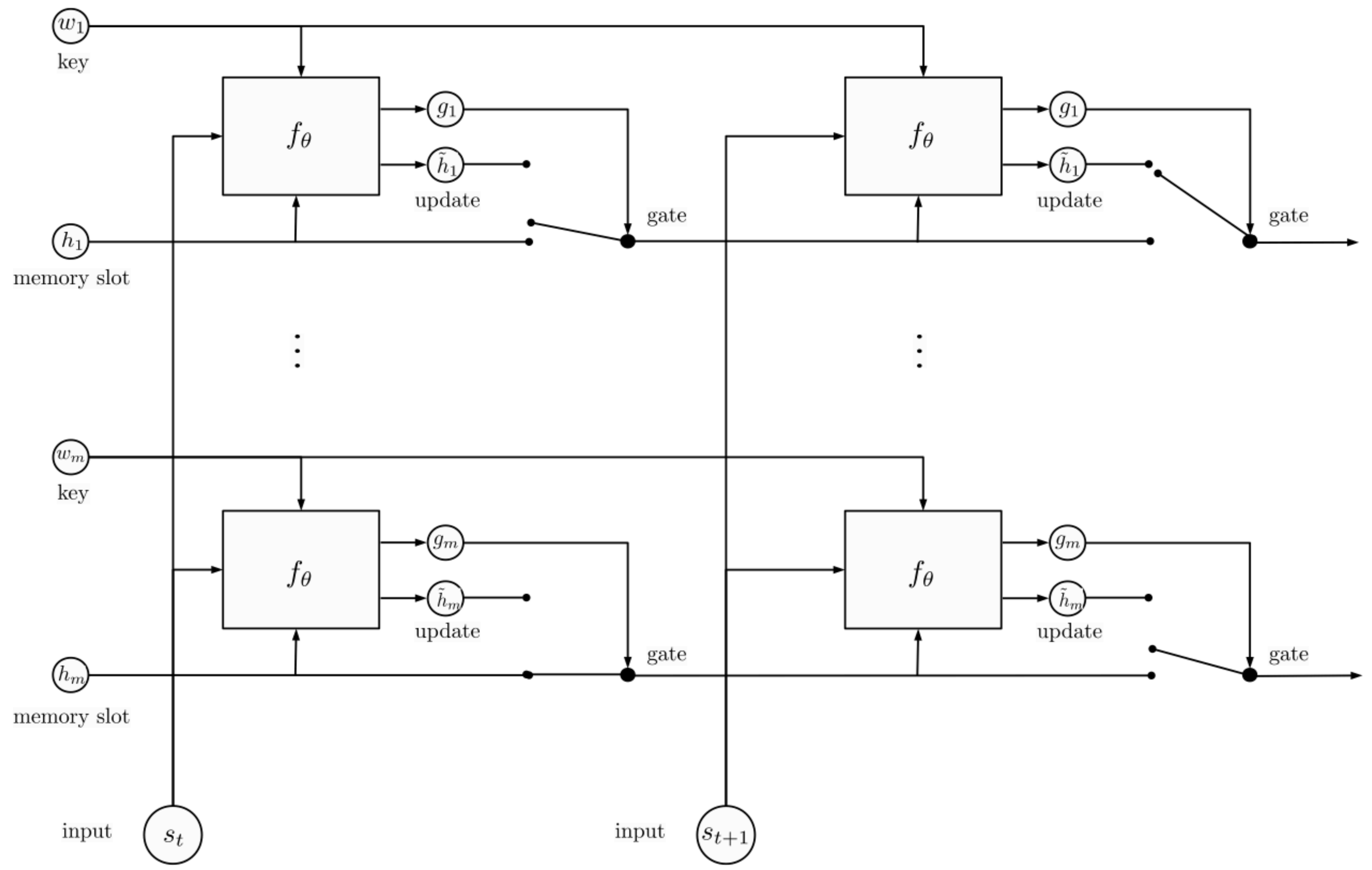
⚡

in-context learning

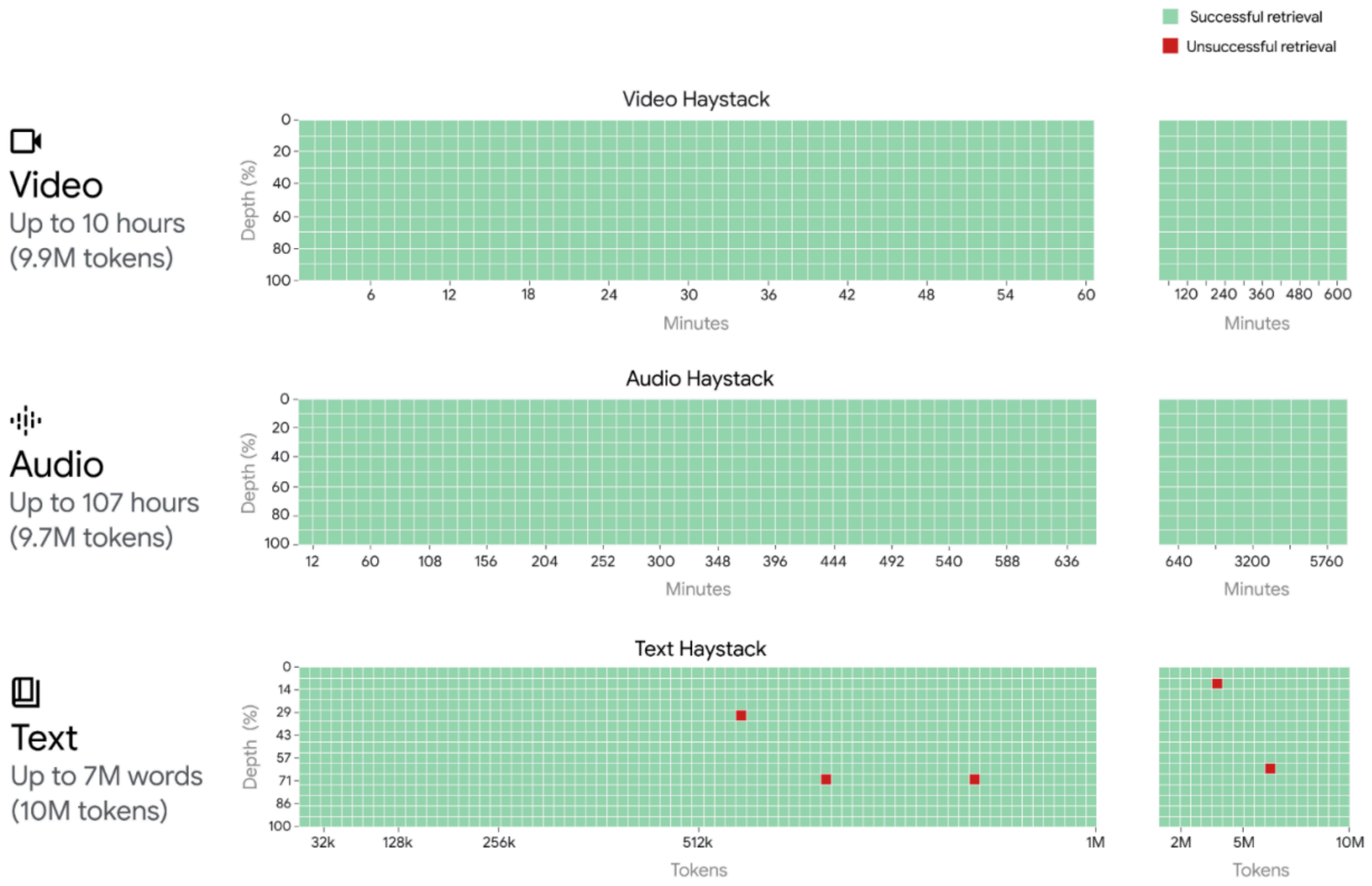
⚓

grounding

memory — attention mechanism



Tracking the World State with Recurrent Entity Networks



pre-LLM attention

attention used in story QA

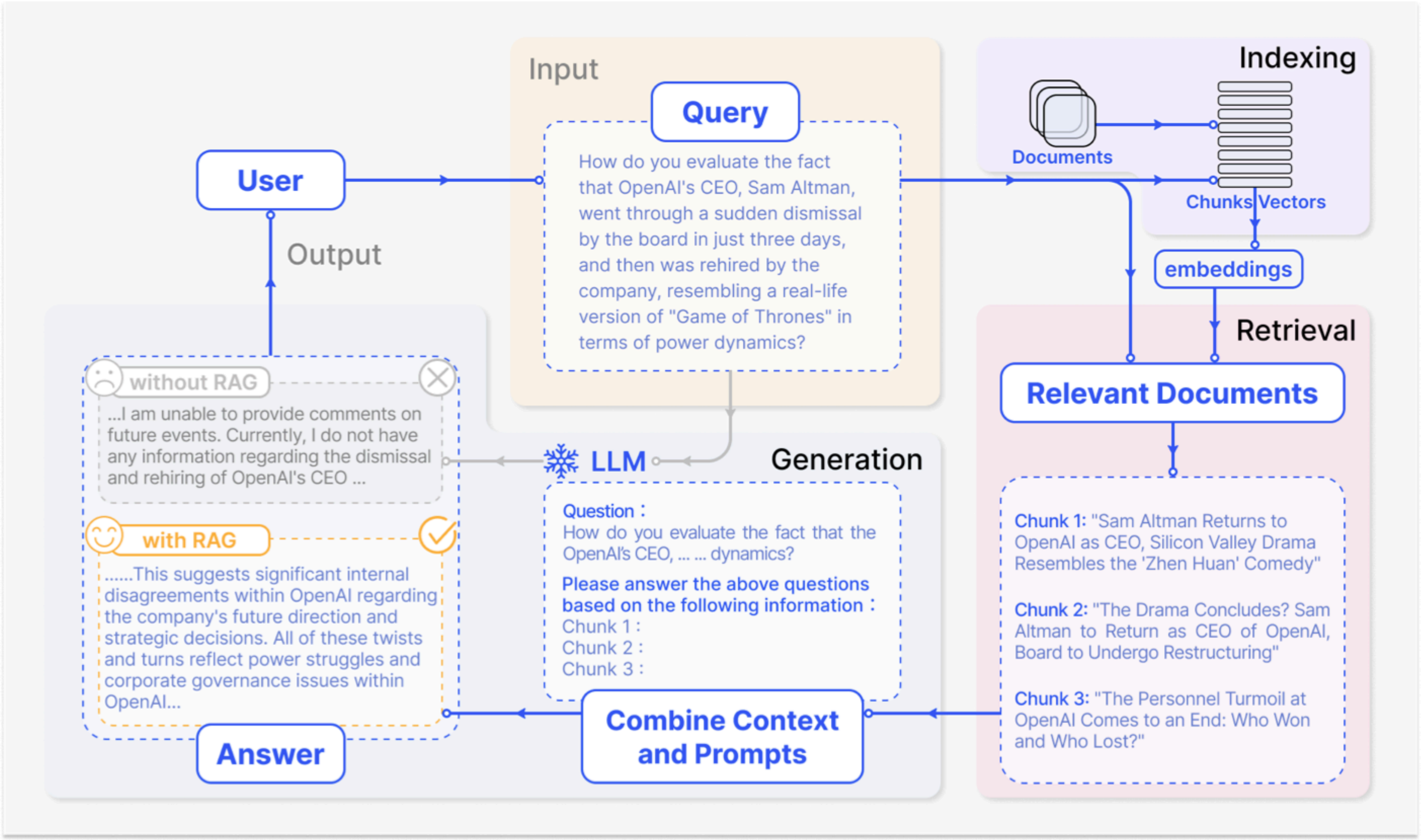
(Henaff et al, 2017)

Gemini 1.5 Pro

10M ~perfect recall

(Google, 2024)

even stronger memory w/ RAG



How do you evaluate the fact that OpenAI's CEO, Sam Altman, went through a sudden dismissal by the board in just three days, and then was rehired by the company, resembling a real-life version of "Game of Thrones" in terms of power dynamics?

...I am unable to provide comments on future events. Currently, I do not have any information regarding the dismissal and rehiring of OpenAI's CEO ...

......This suggests significant internal disagreements within OpenAI regarding the company's future direction and strategic decisions. All of these twists and turns reflect power struggles and corporate governance issues within OpenAI...

Question：
How do you evaluate the fact that the OpenAI's CEO, ... dynamics?
Please answer the above questions based on the following information：
Chunk 1 :
Chunk 2 :
Chunk 3 :

Chunk 1: "Sam Altman Returns to OpenAI as CEO, Silicon Valley Drama Resembles the 'Zhen Huan' Comedy"

Chunk 2: "The Drama Concludes? Sam Altman to Return as CEO of OpenAI, Board to Undergo Restructuring"

Chunk 3: "The Personnel Turmoil at OpenAI Comes to an End: Who Won and Who Lost?"

Retrieval-Augmented Generation for Large Language Models: A Survey

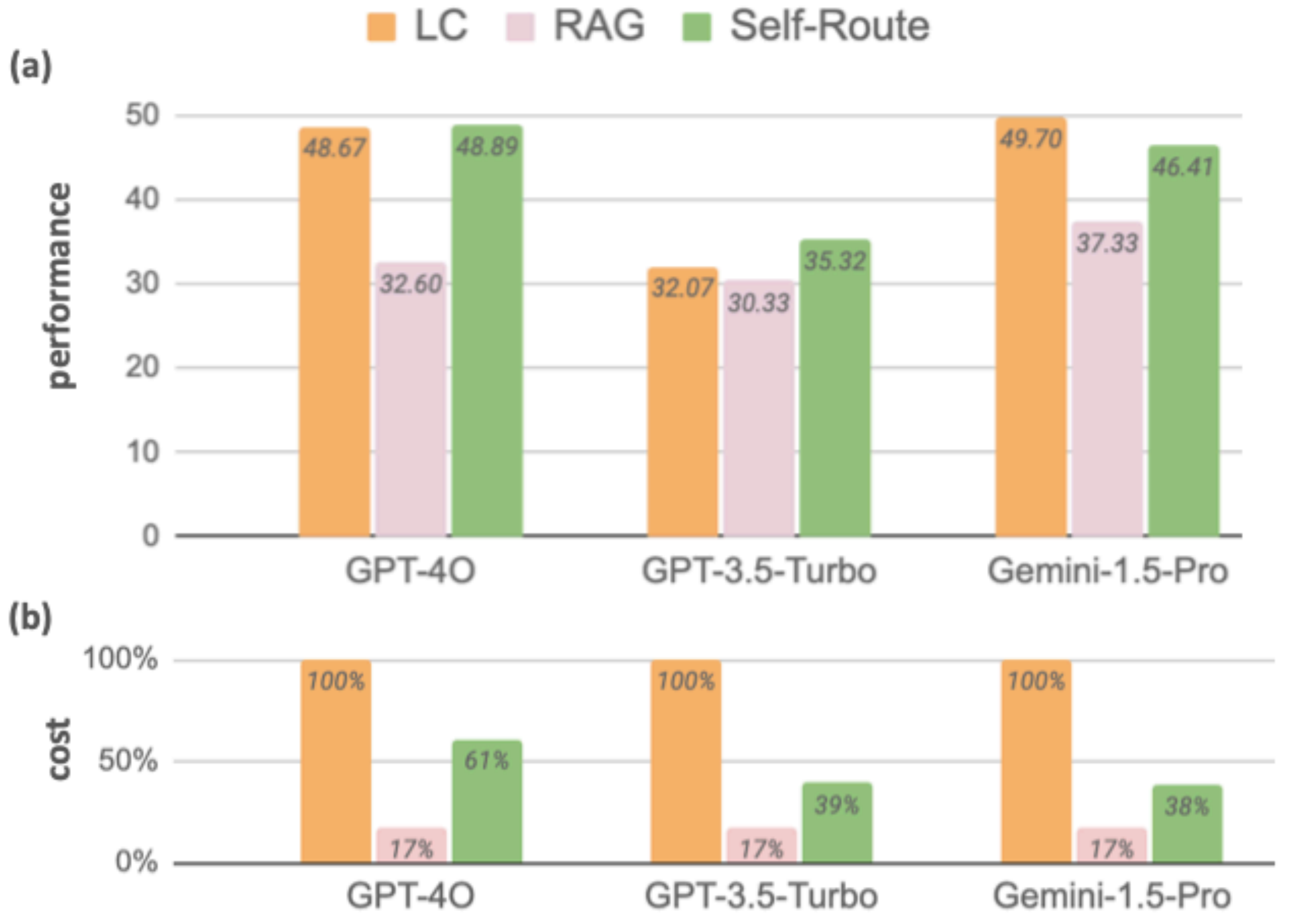


Figure 1: While long-context LLMs (LC) surpass RAG in long-context understanding, RAG is significantly more cost-efficient. Our approach, SELF-ROUTE, combining RAG and LC, achieves comparable performance to LC at a much lower cost.

Retrieval Augmented Generation or Long-Context LLMs? A Comprehensive Study and Hybrid Approach

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn...

👩‍💻

coding

💾

memory

⚡

in-context learning

⚓

grounding

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn…

👩‍💻

coding

💾

memory

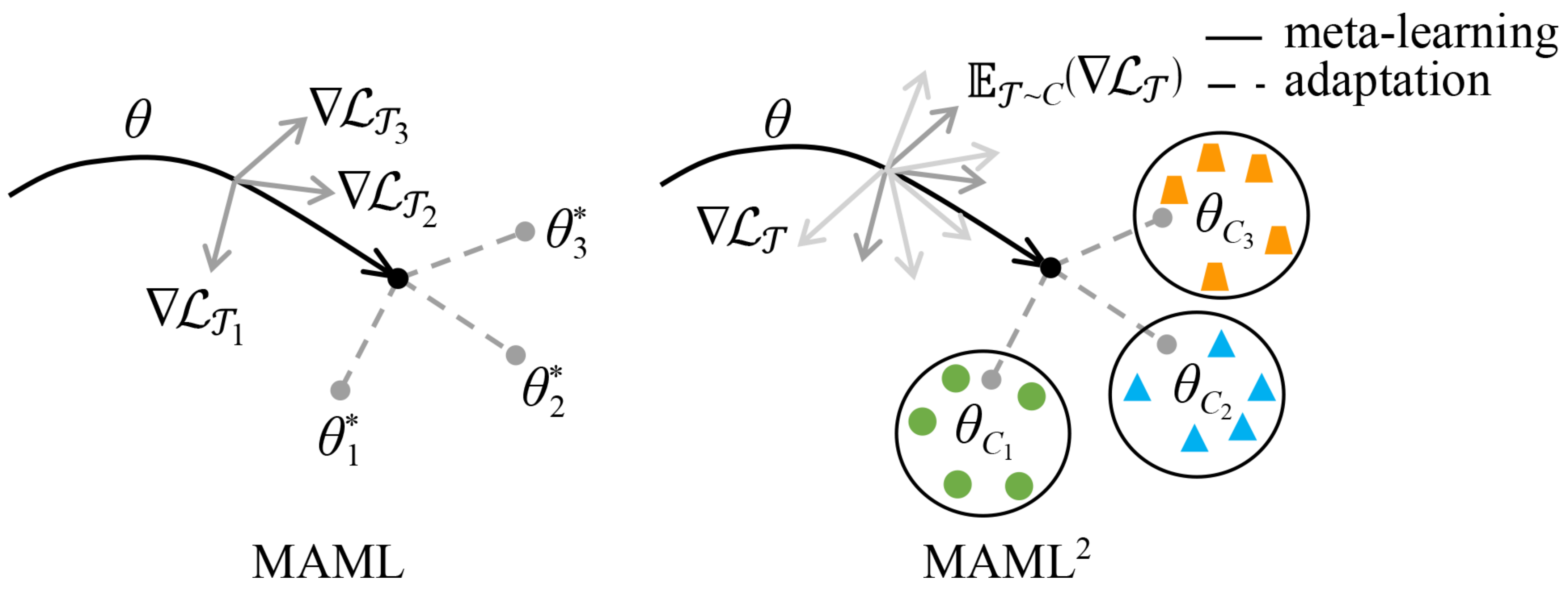
⚡

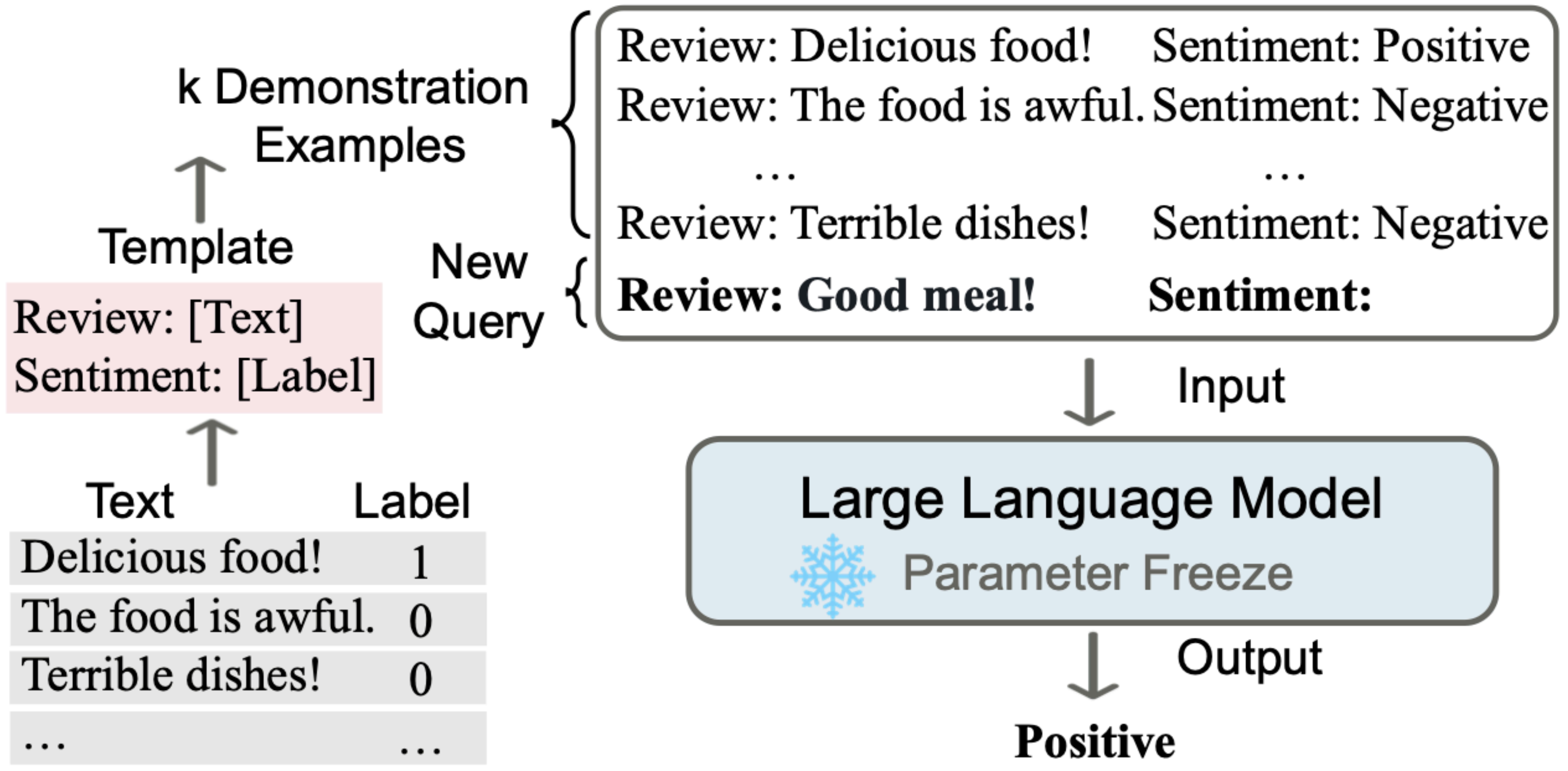
in-context learning

⚓

grounding

in-context learning


θ
$\nabla\mathcal{L}_{\mathcal{T}_3}$
$\nabla\mathcal{L}_{\mathcal{T}_2}$
$\nabla\mathcal{L}_{\mathcal{T}_1}$
θ_3^*
θ_2^*
θ_1^*
MAML
θ
$\mathbb{E}_{\mathcal{T}\sim C}(\nabla\mathcal{L}_{\mathcal{T}})$
$\nabla\mathcal{L}_{\mathcal{T}}$
— meta-learning
- - adaptation
θ_{C_3}
θ_{C_2}
θ_{C_1}
MAML^2



Review: [Text]
Sentiment: [Label]

Text	Label
Delicious food!	1
The food is awful.	0
Terrible dishes!	0
…	…

Review: Delicious food! Sentiment: Positive
Review: The food is awful. Sentiment: Negative
… …
Review: Terrible dishes! Sentiment: Negative
Review: Good meal! Sentiment:

pre-LLM few-shot learning

gradient-based meta learning

(Finn et al, 2017)

ICL emerges

10M ~perfect recall

(Brown et al, 2020)

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn…

👩‍💻

coding

💾

memory

⚡

in-context learning

⚓

grounding

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn...

👩‍💻

coding

💾

memory

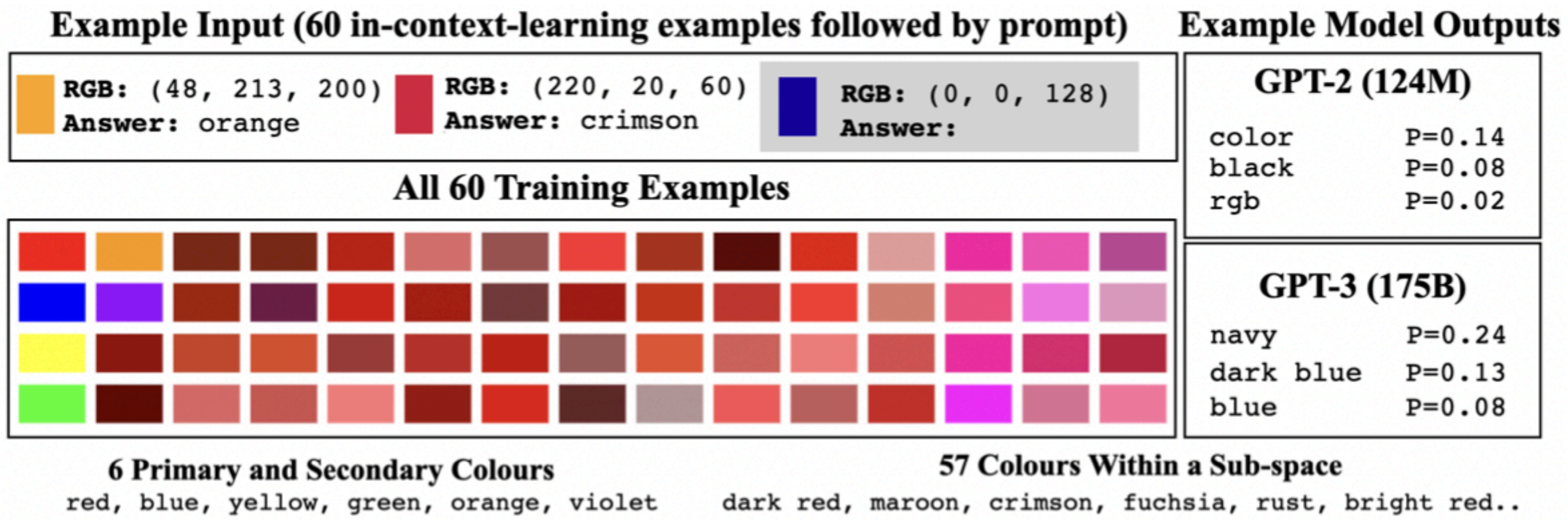
⚡

in-context learning

⚓

grounding

grounding



Example Input (60 in-context-learning examples followed by prompt)

`RGB: (48, 213, 200)` `Answer: orange`

`RGB: (220, 20, 60)` `Answer: crimson`

`RGB: (0, 0, 128)` `Answer:`

All 60 Training Examples

6 Primary and Secondary Colours
red, blue, yellow, green, orange, violet

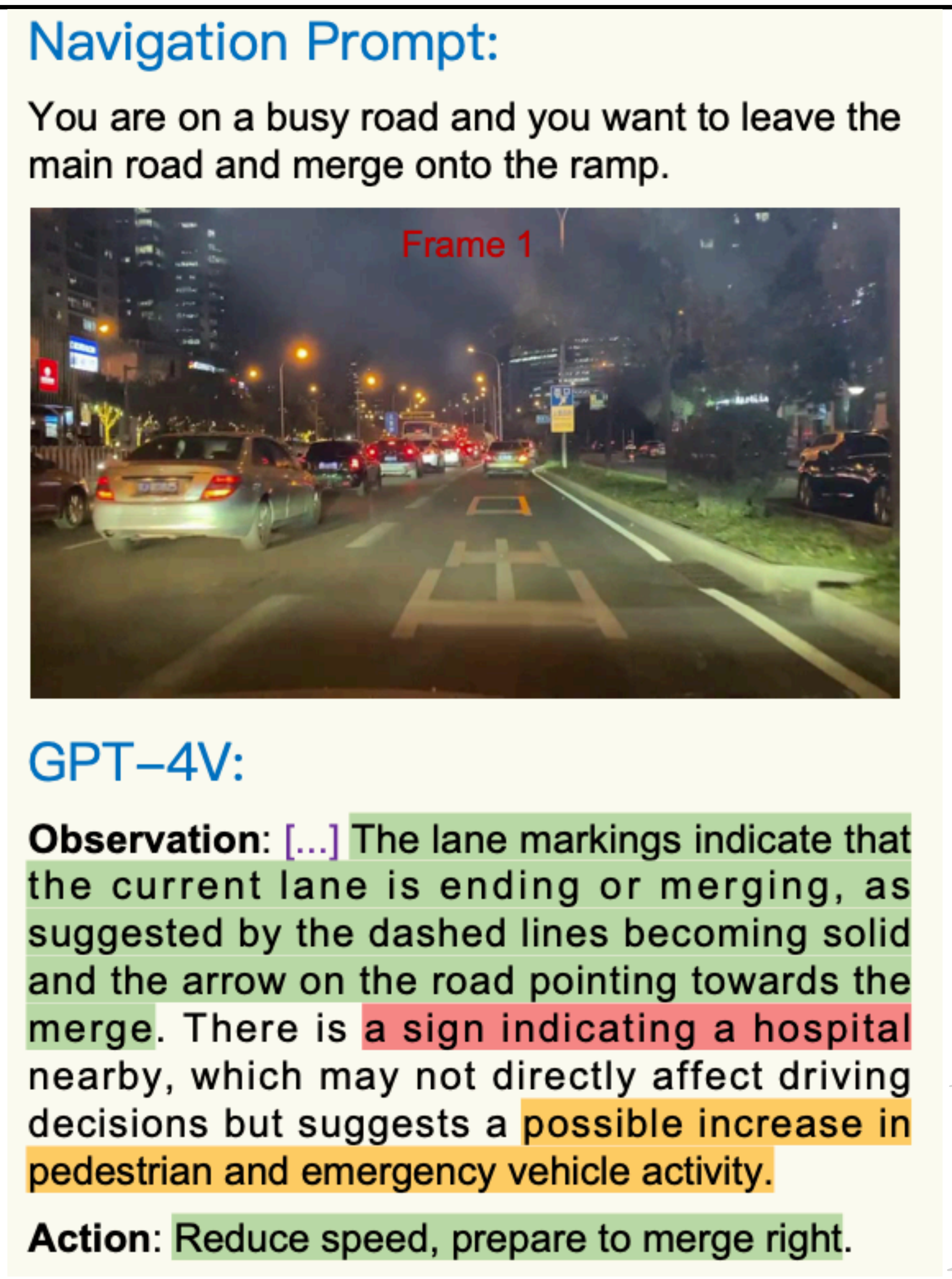
57 Colours Within a Sub-space
dark red, maroon, crimson, fuchsia, rust, bright red..

Example Model Outputs

GPT-2 (124M)	
color	P=0.14
black	P=0.08
rgb	P=0.02

GPT-3 (175B)	
navy	P=0.24
dark blue	P=0.13
blue	P=0.08

Mapping Language Models to Grounded Conceptual Spaces



Navigation Prompt:

You are on a busy road and you want to leave the main road and merge onto the ramp.

GPT–4V:

Observation: [...] The lane markings indicate that the current lane is ending or merging, as suggested by the dashed lines becoming solid and the arrow on the road pointing towards the merge. There is a sign indicating a hospital nearby, which may not directly affect driving decisions but suggests a possible increase in pedestrian and emergency vehicle activity.

Action: Reduce speed, prepare to merge right.

On the Road with GPT-4V(ision): Early Explorations of Visual-Language Model on Autonomous Driving

non-multimodal LLMs

rich concept structure

(Patel et al, 2022)

GPT-4V

embodied design making

(Wen et al, 2023)

“emergent” capabilities of LLMs

we only wanted to train next-token predictors, how did it learn...

👩‍💻

coding

💾

memory

⚡

in-context learning

⚓

grounding

two **useful** thinking tools

🤖🪄

agent thinking

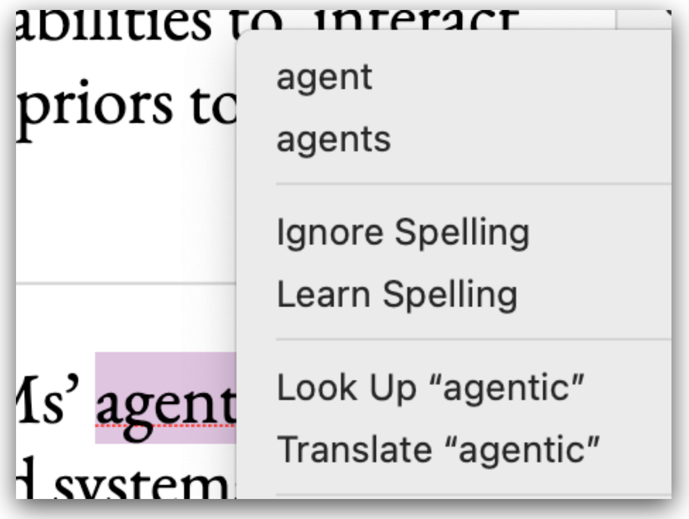
How to improve the models' capabilities to interact with the world? Are LLMs good priors to start with?

🧠🪄

LLM Thinking

How to make full use of the LLMs' agentic capabilities? What algorithms and systems we should on top of them?

Yeah, I know. "Agentic" is weird.



what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

👁️

perception

📝

planning

🎯

agency

📈

learning

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

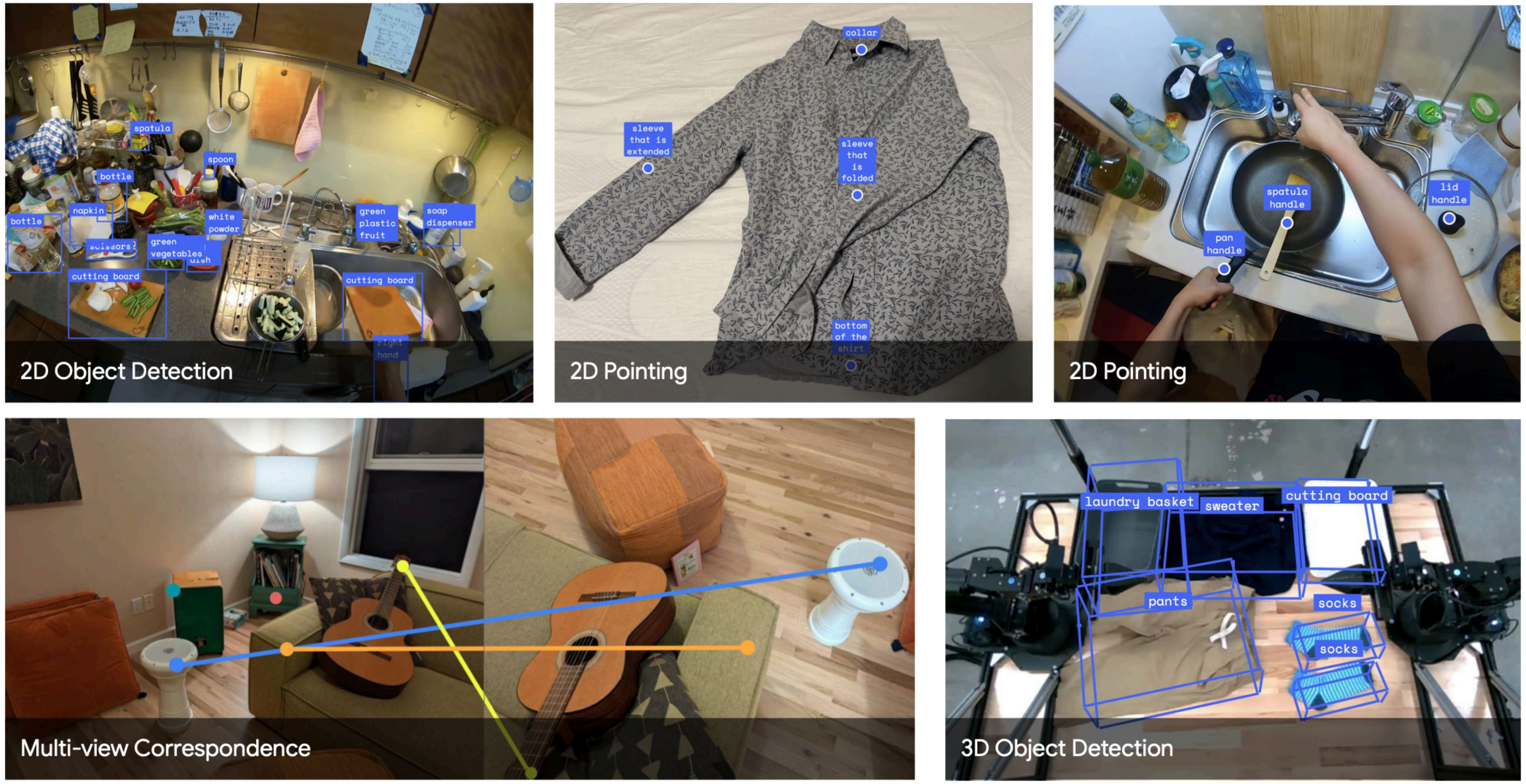
perception

planning

agency

learning

perception



Gemini Robotics: Bringing AI into the Physical World

Gemini-Robotics

Understanding the semantic structure of observation

(Google, 2025)

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world ⤵

👁️

perception

📝

planning

🎯

agency

📈

learning

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

👁️

perception

📝

planning

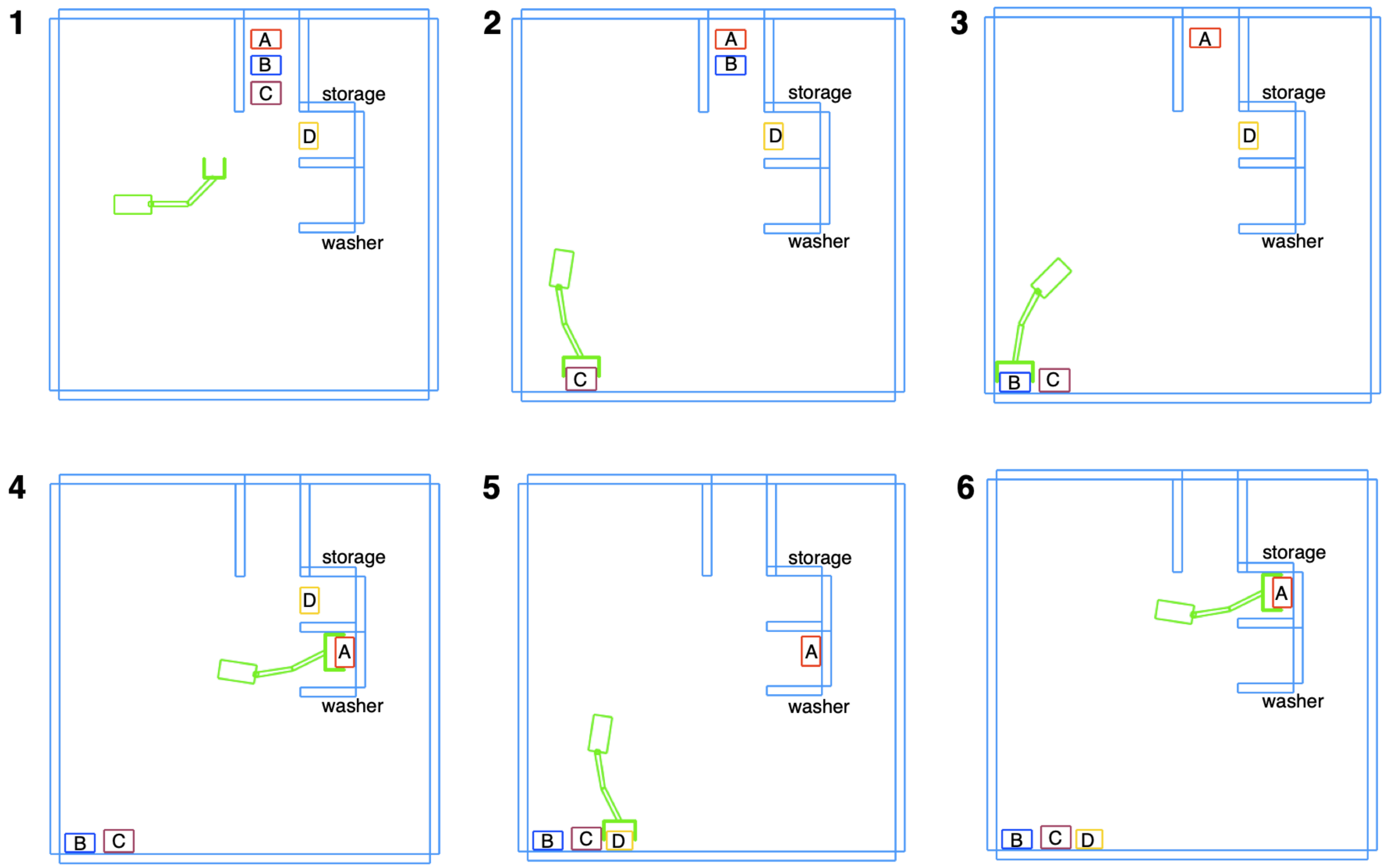
🎯

agency

📈

learning

planning



Hierarchical task and motion planning in the now

Task and Motion Planning

Top-down decomposition w/ bottom-up constraints

(Kaelbling and Lozano-Pérez, 2011)

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

👁️

perception

📝

planning

🎯

agency

📈

learning

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

👁️

perception

📝

planning

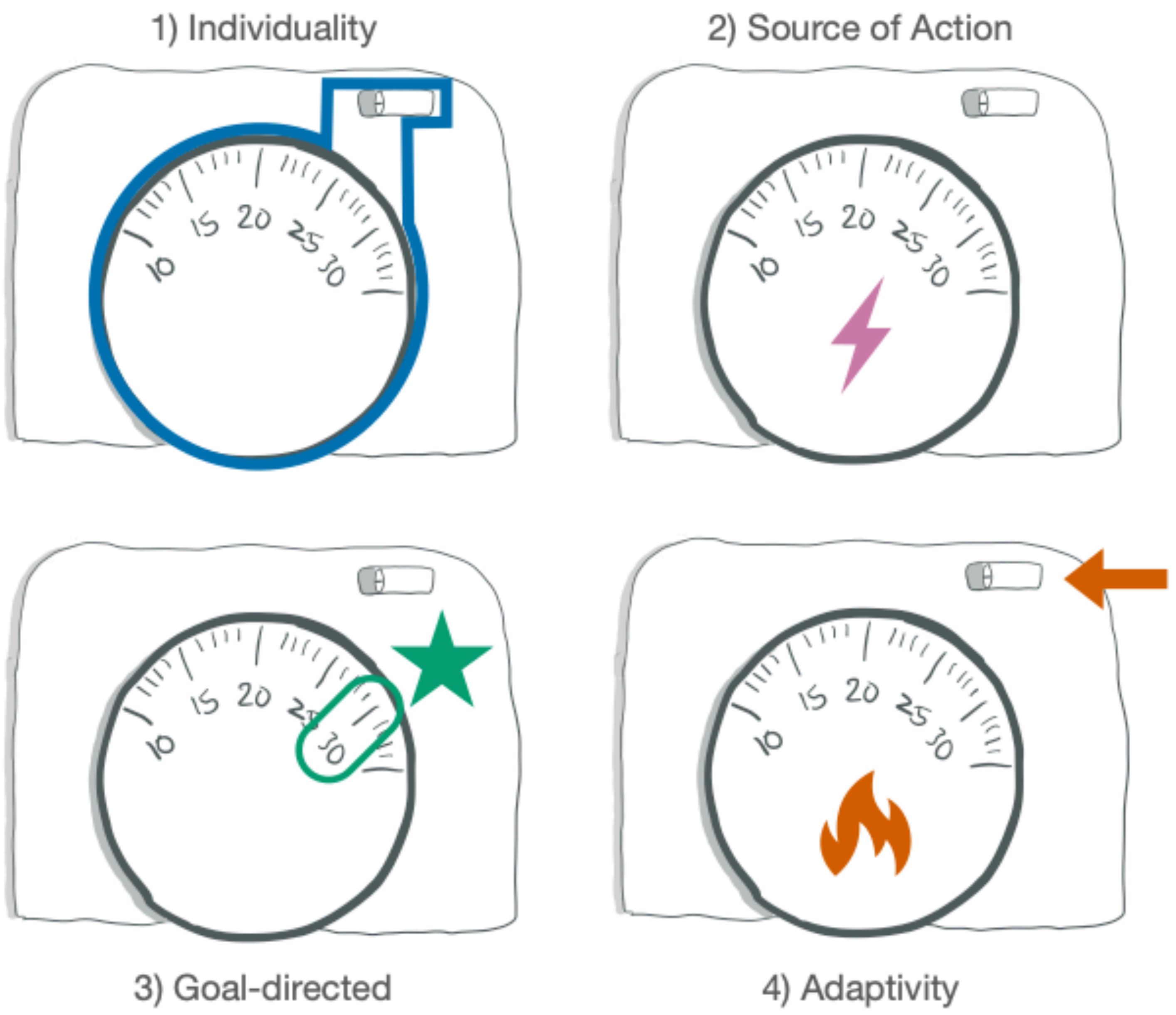
🎯

agency

📈

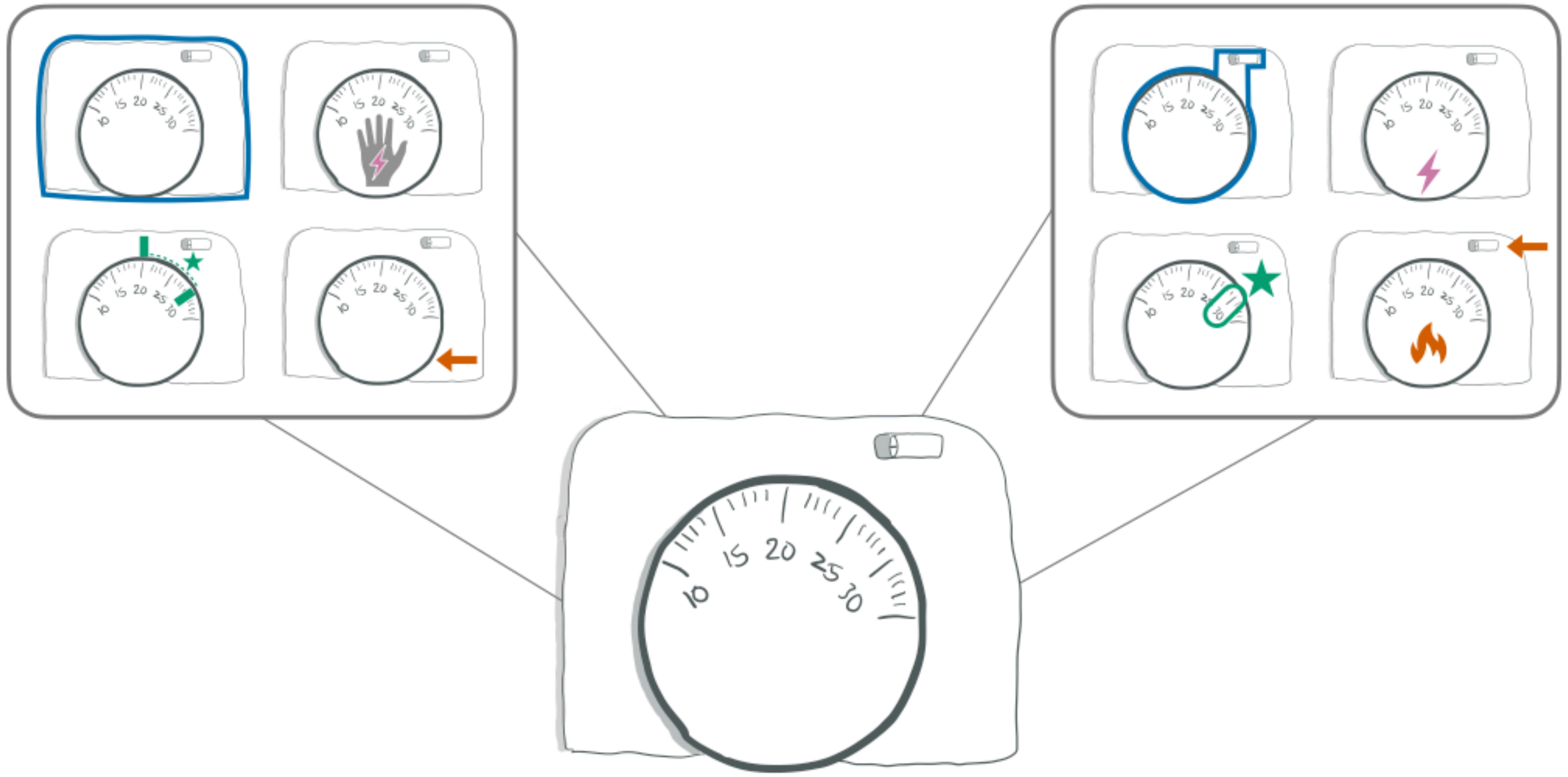
learning

agency — the essence but hard to define



(a) A Four-Part Account of Agency

Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action.



(b) Frame-Dependence

Agency Is Frame-Dependent

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

👁️

perception

📝

planning

🎯

agency

📈

learning

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

👁️

perception

📝

planning

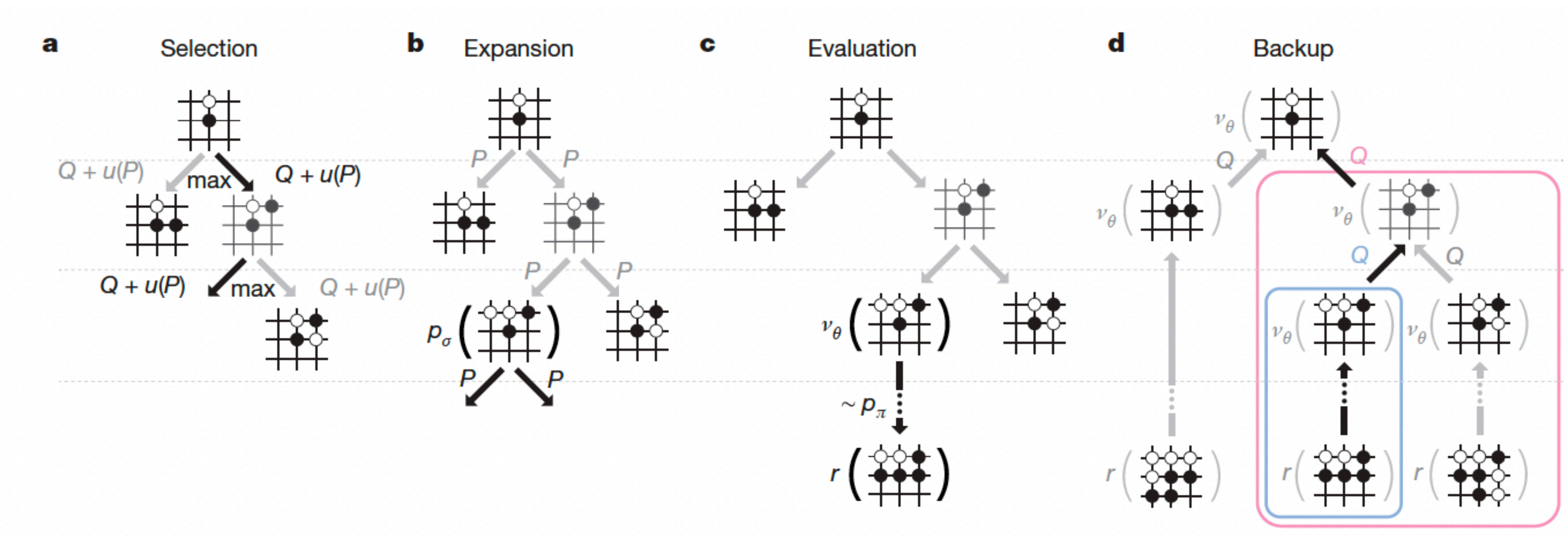
🎯

agency

📈

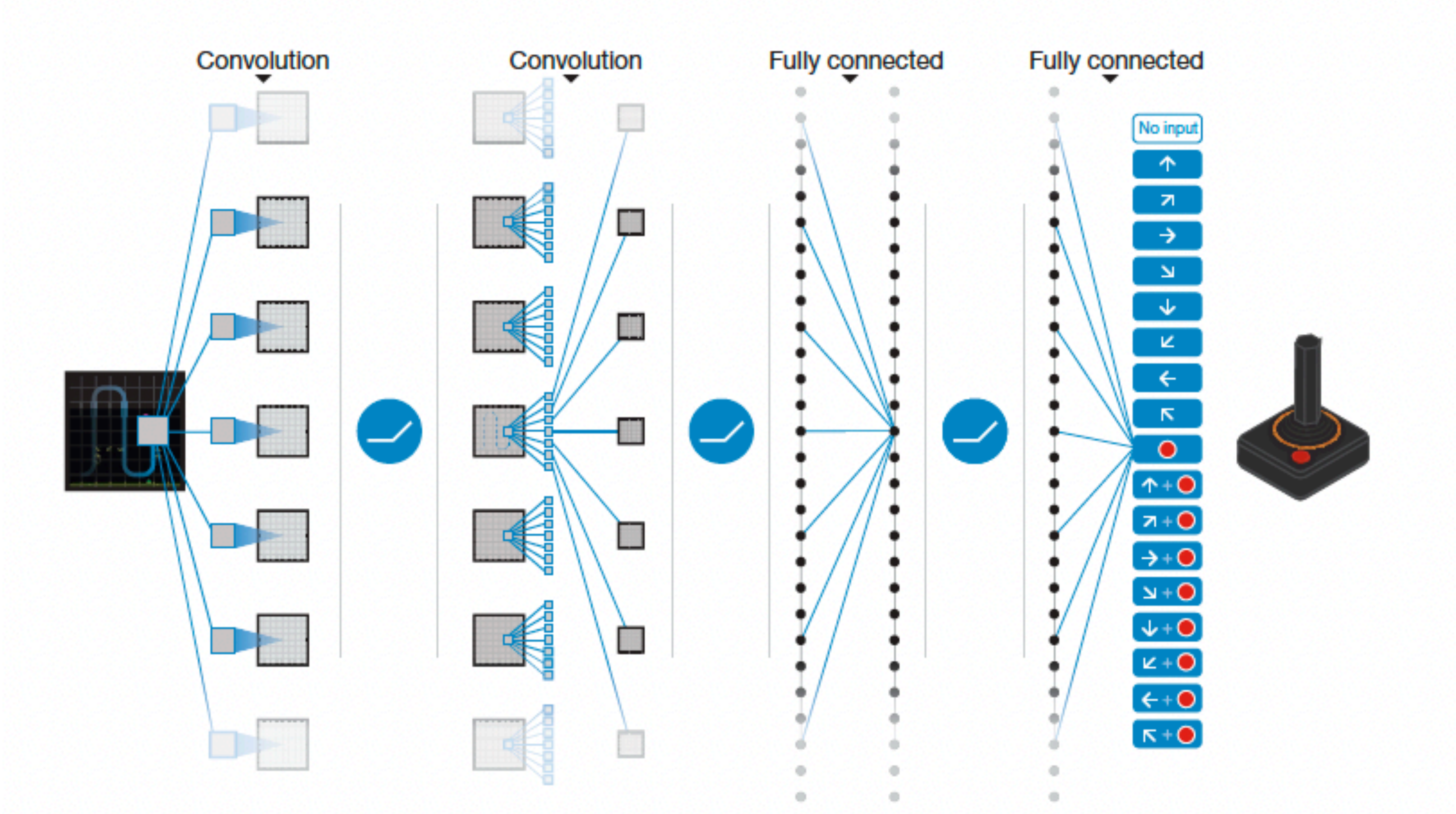
learning

learning



learning through search

Mastering the game of Go with deep neural networks and tree search



learning through RL

Human-level control through deep reinforcement learning

what are agentic capabilities

the capabilities that an agent needs to have to interact with the world

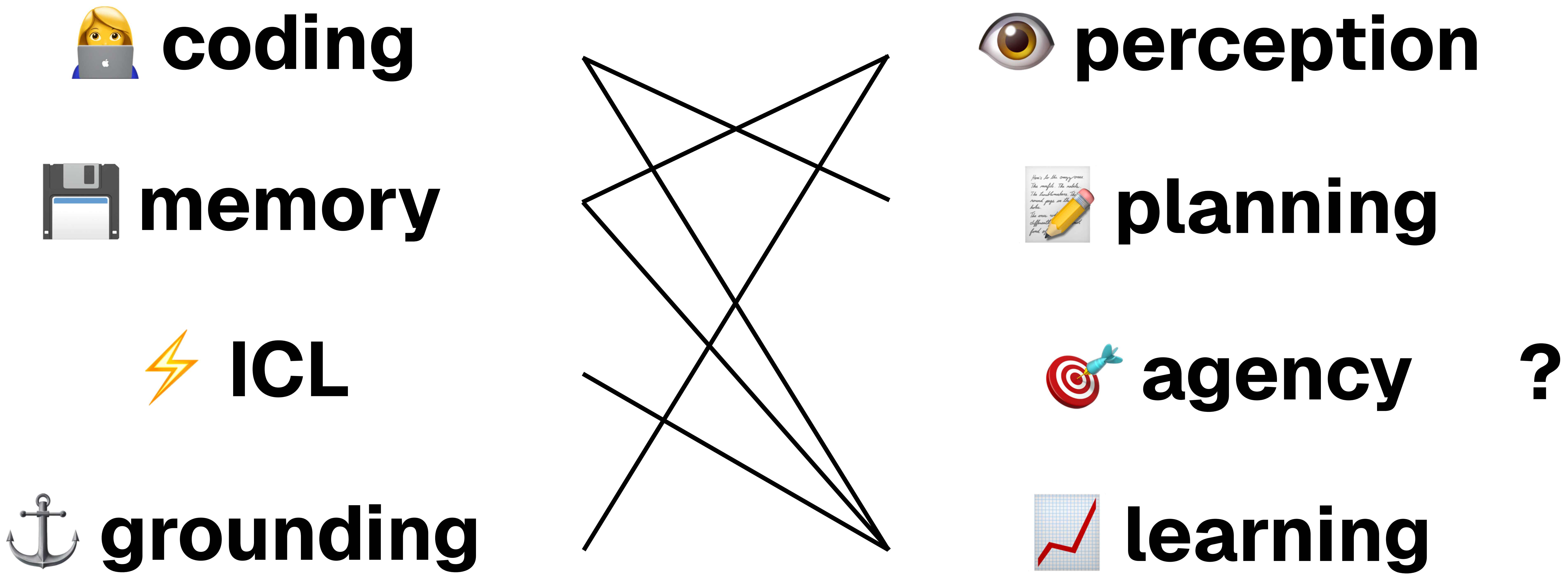
👁️ **perception**

📝 **planning**

🎯 **agency**

📈 **learning**

are LLMs good priors?



the **agentic** capabilities

short summary

🤖🪄

agent thinking

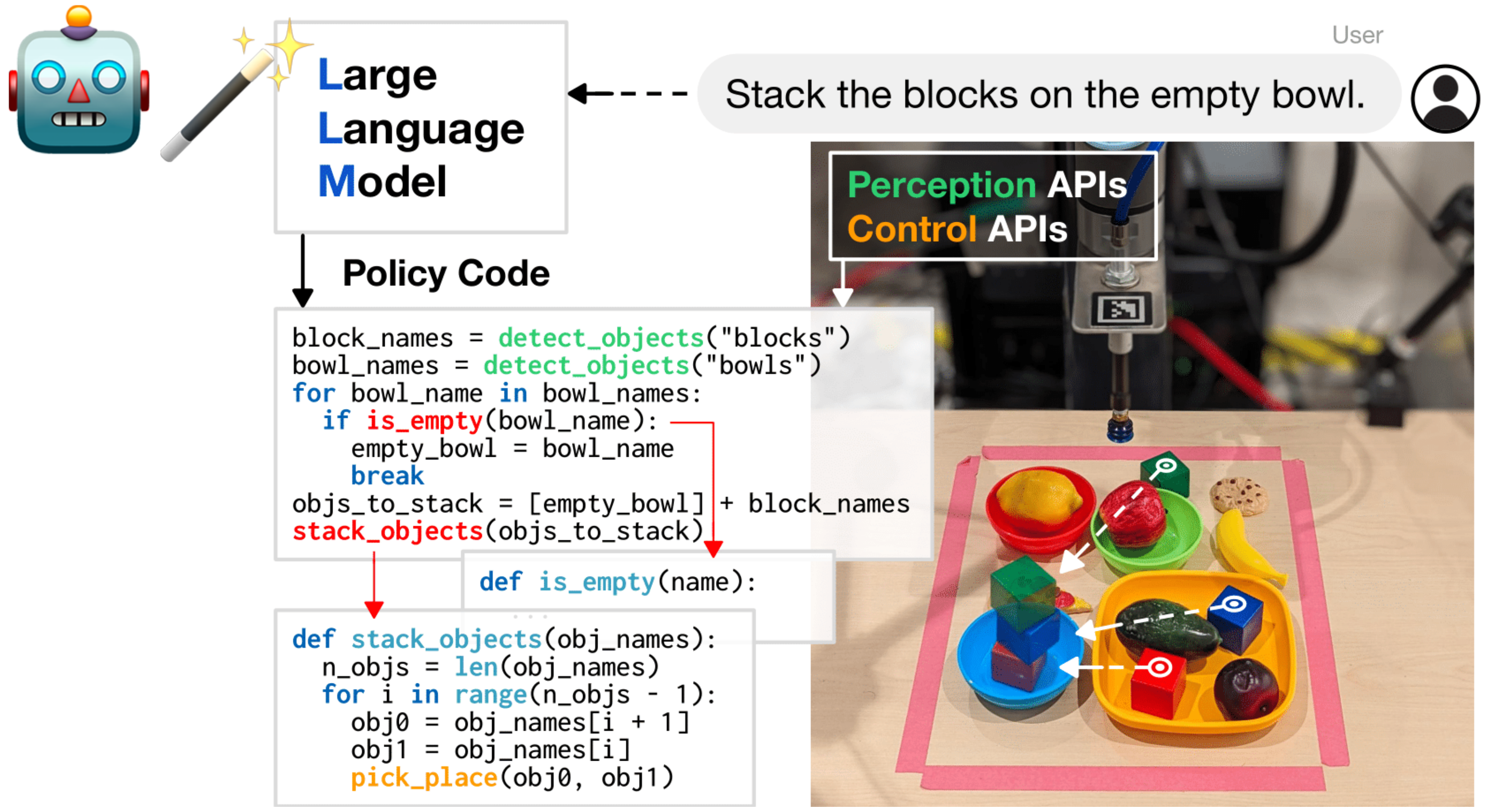
LLMs provide good priors for the capabilities that we consider as essential for agents, so we should start from there.

🧠🪄

LLM Thinking

We could think from these agentic capabilities perspective when building agents.

coding for planning/learning



Stack the blocks on the empty bowl.

```
block_names = detect_objects("blocks")
bowl_names = detect_objects("bowls")
for bowl_name in bowl_names:
  if is_empty(bowl_name):
    empty_bowl = bowl_name
    break
objs_to_stack = [empty_bowl] + block_names
stack_objects(objs_to_stack)
```

```
def is_empty(name):
  ...
```

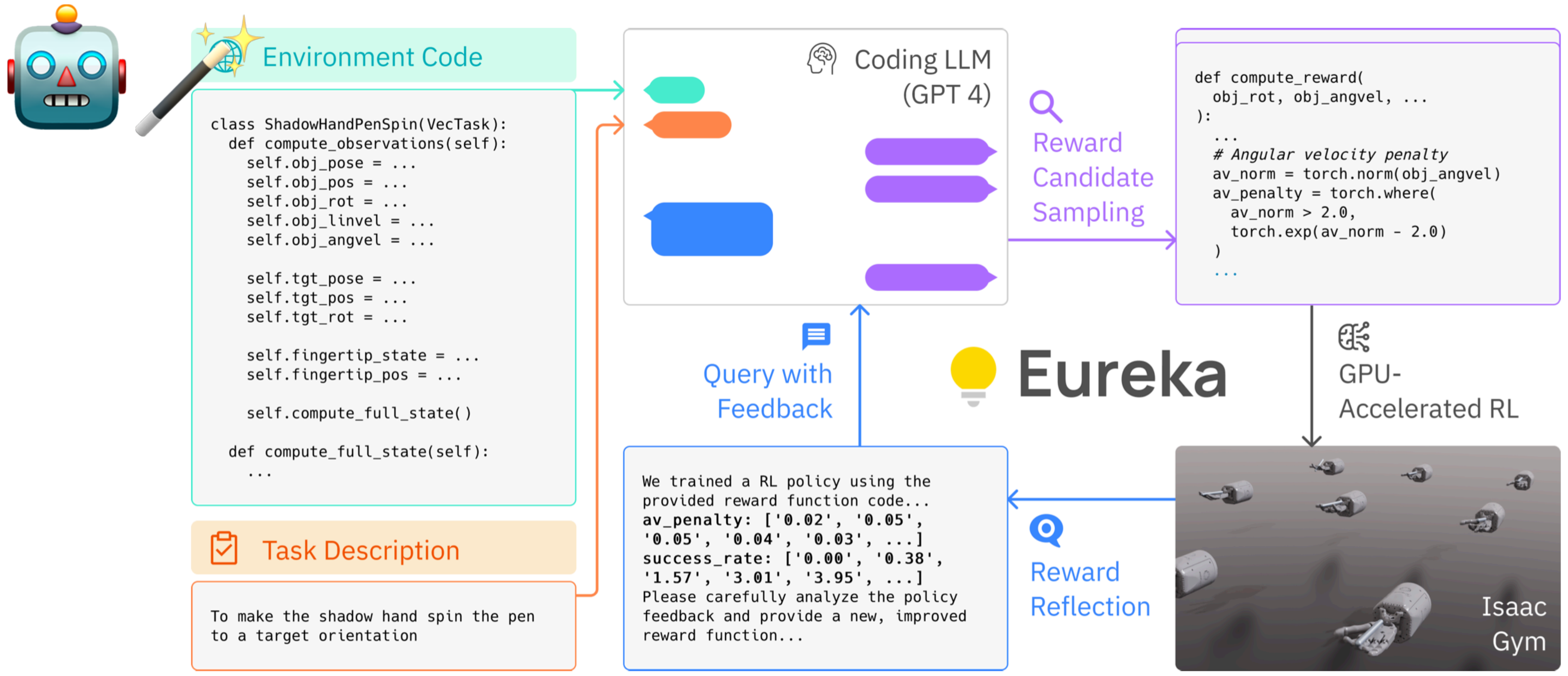
```
def stack_objects(obj_names):
  n_objs = len(obj_names)
  for i in range(n_objs - 1):
    obj0 = obj_names[i + 1]
    obj1 = obj_names[i]
    pick_place(obj0, obj1)
```

Code as Policies: Language Model Programs for Embodied Control

code as policy

control robot w/o eyes

(Liang et al, 2022)



```
class ShadowHandPenSpin(VecTask):
  def compute_observations(self):
    self.obj_pose = ...
    self.obj_pos = ...
    self.obj_rot = ...
    self.obj_linvel = ...
    self.obj_angvel = ...

    self.tgt_pose = ...
    self.tgt_pos = ...
    self.tgt_rot = ...

    self.fingertip_state = ...
    self.fingertip_pos = ...

    self.compute_full_state()

  def compute_full_state(self):
    ...
```

To make the shadow hand spin the pen to a target orientation

```
def compute_reward(
  obj_rot, obj_angvel, ...
):
  ...
  # Angular velocity penalty
  av_norm = torch.norm(obj_angvel)
  av_penalty = torch.where(
    av_norm > 2.0,
    torch.exp(av_norm - 2.0)
  )
  ...
```

```
We trained a RL policy using the
provided reward function code...
av_penalty: ['0.02', '0.05',
'0.05', '0.04', '0.03', ...]
success_rate: ['0.00', '0.38',
'1.57', '3.01', '3.95', ...]
Please carefully analyze the policy
feedback and provide a new, improved
reward function...
```

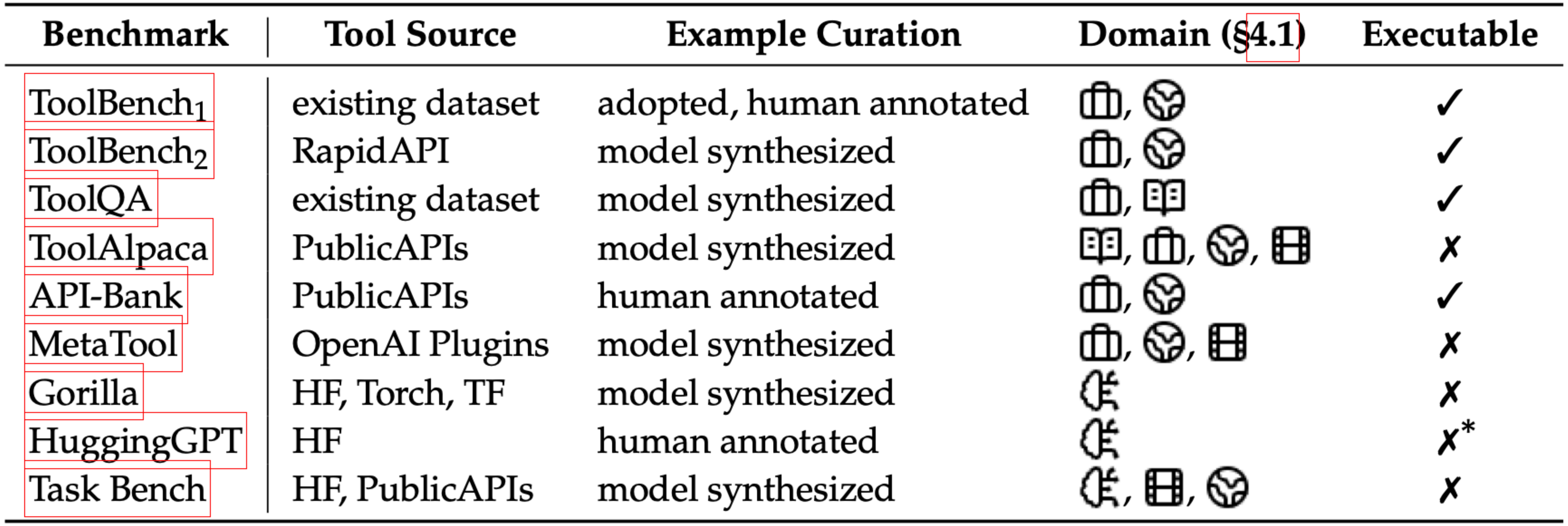
Eureka: Human-Level Reward Design via Coding Large Language Models

code as reward

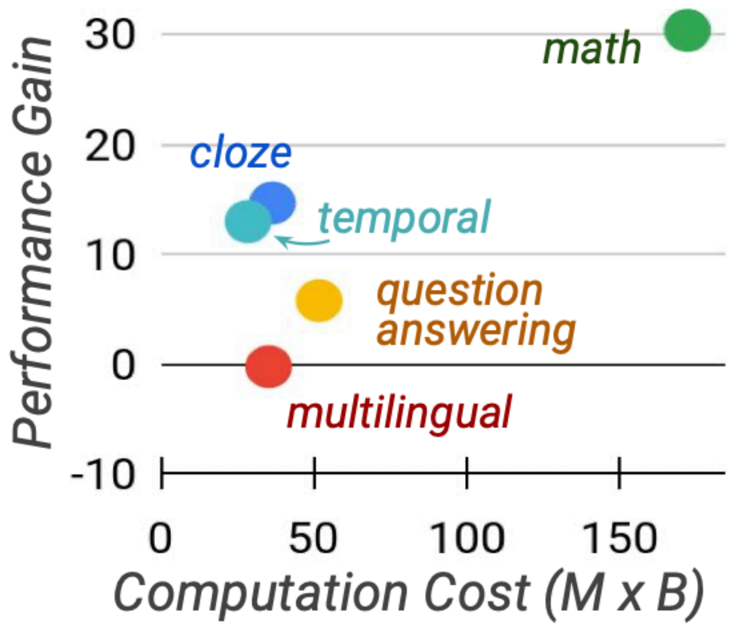
prompt to code for QA

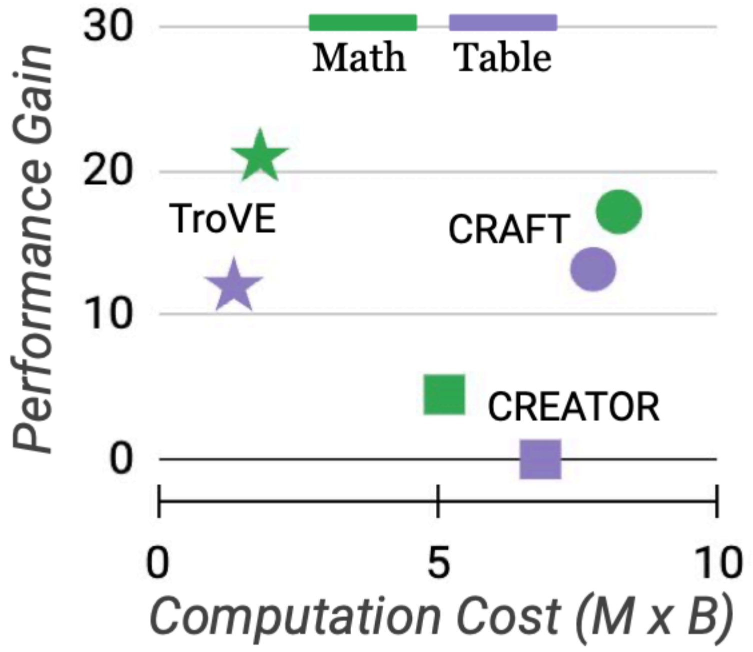
(Ma et al, 2024)

coding for planning



Benchmark	Tool Source	Example Curation	Domain (§4.1)	Executable
ToolBench_1	existing dataset	adopted, human annotated	💼, 🌐	✓
ToolBench_2	RapidAPI	model synthesized	💼, 🌐	✓
ToolQA	existing dataset	model synthesized	💼, 📖	✓
ToolAlpaca	PublicAPIs	model synthesized	📖, 💼, 🌐, 🎞	✗
API-Bank	PublicAPIs	human annotated	💼, 🌐	✓
MetaTool	OpenAI Plugins	model synthesized	💼, 🌐, 🎞	✗
Gorilla	HF, Torch, TF	model synthesized	🧠	✗
HuggingGPT	HF	human annotated	🧠	✗*
Task Bench	HF, PublicAPIs	model synthesized	🧠, 🎞, 🌐	✗



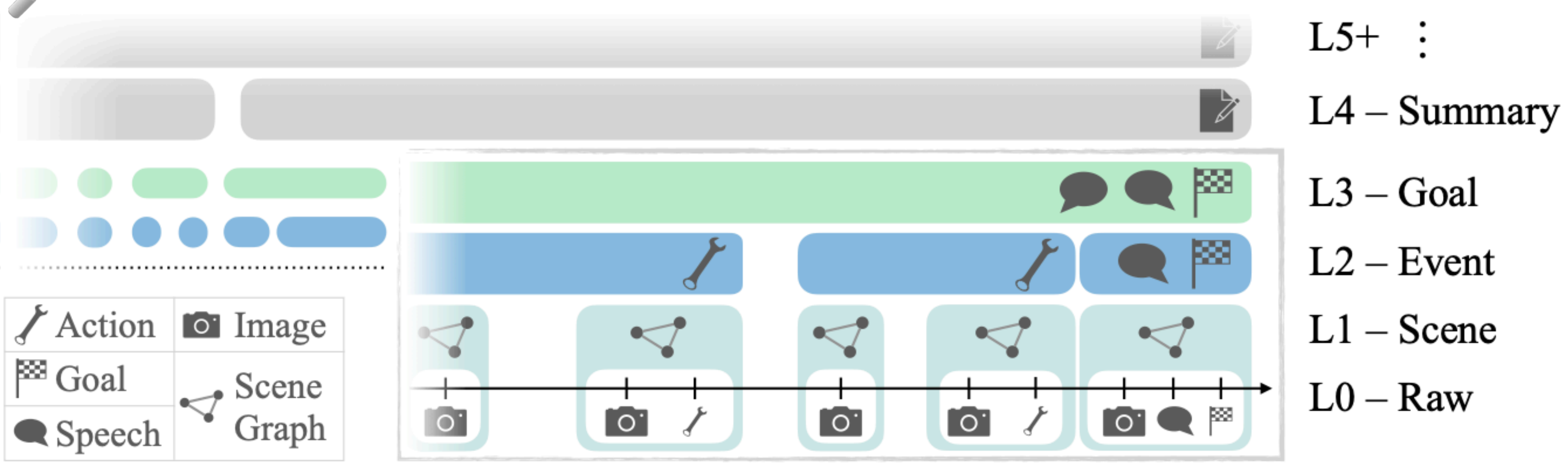


Tool use / make cases

when (not) to use tools

(Wang et al, 2024)

memory for perception

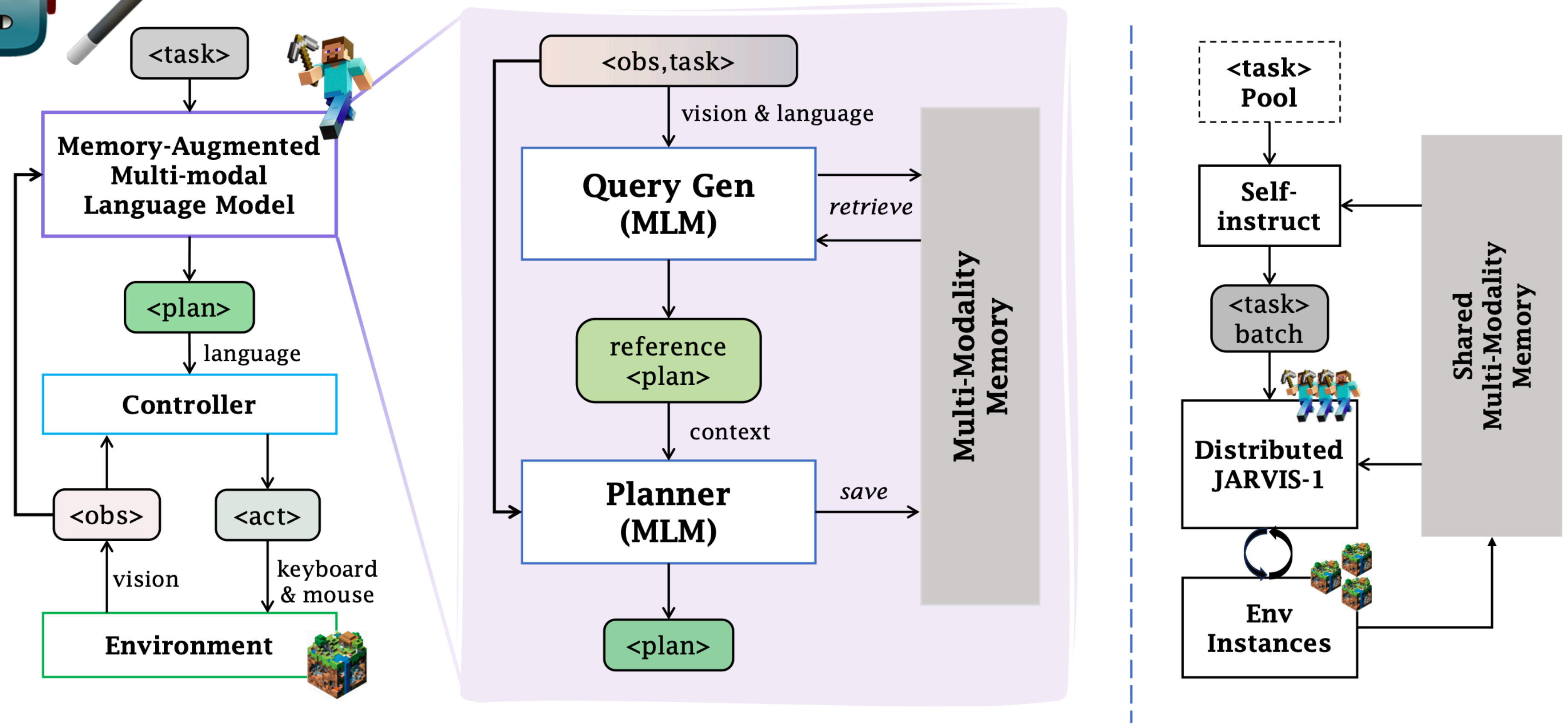


Episodic Memory Verbalization using Downscaled Hierarchical Representations of Life-Long Robot Experience

hierarchical memory

long memory robotics QA

(Bärmann et al, 2024)



(a) JARVIS-1 architecture (b) Self-Improving

JARVIS-1: Open-World Multi-task Agents with Memory-Augmented Multimodal Language Models

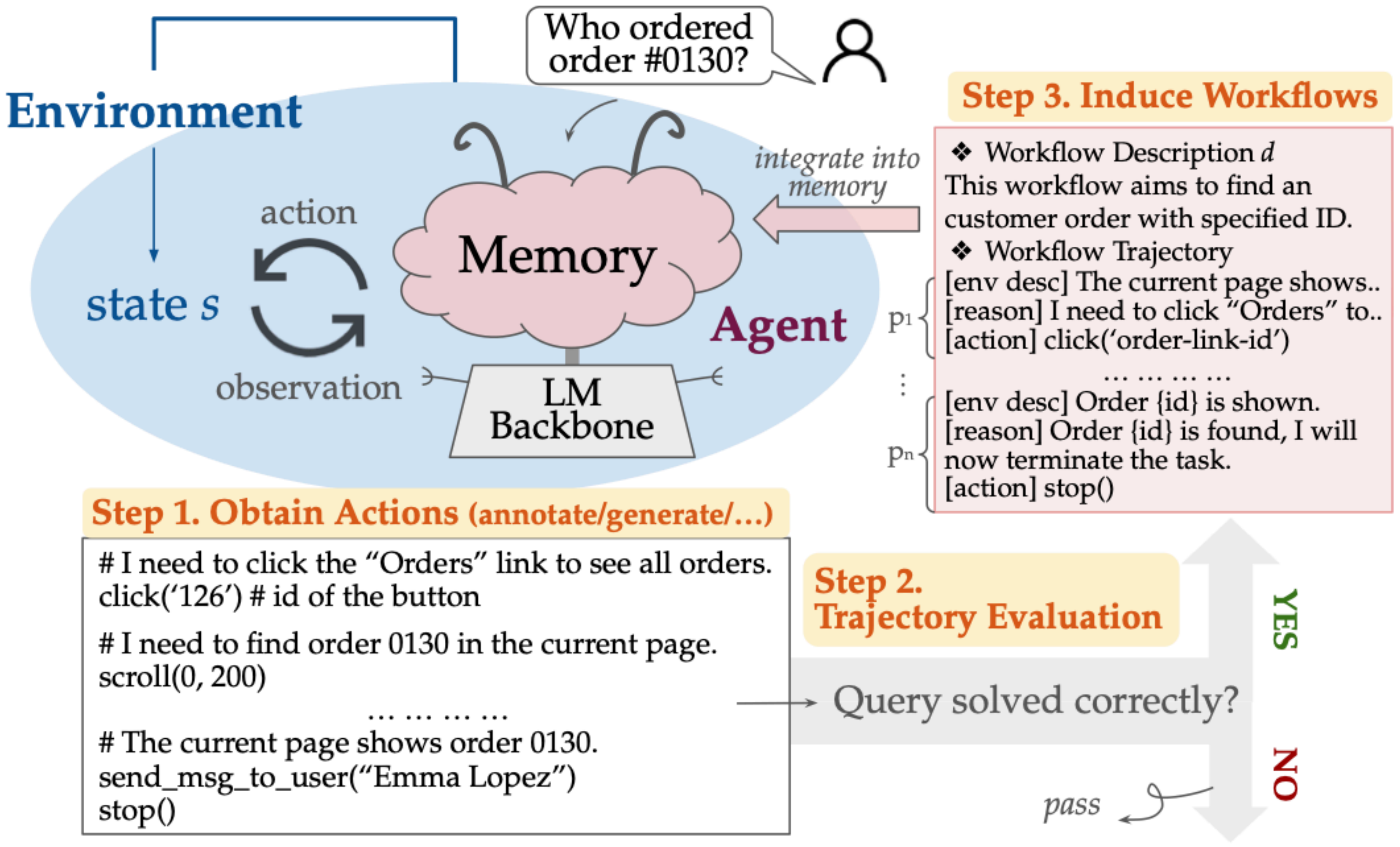
self-improving agents

memory augmented MLM

(Wang, 2024)

memory & ICL for learning

🤖🪄



❖ Workflow Description *d*
This workflow aims to find an customer order with specified ID.
❖ Workflow Trajectory
[env desc] The current page shows..
[reason] I need to click "Orders" to..
[action] click('order-link-id')
… … … …
[env desc] Order {id} is shown.
[reason] Order {id} is found, I will now terminate the task.
[action] stop()

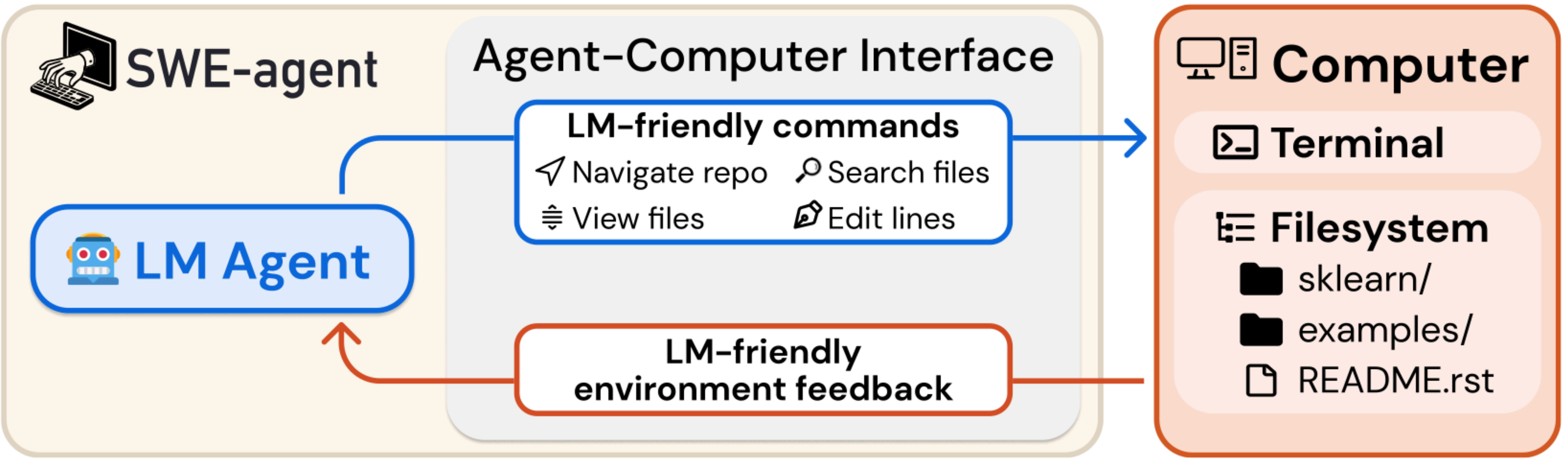
```
# I need to click the "Orders" link to see all orders.
click('126') # id of the button
# I need to find order 0130 in the current page.
scroll(0, 200)
… … … …
# The current page shows order 0130.
send_msg_to_user("Emma Lopez")
stop()
```

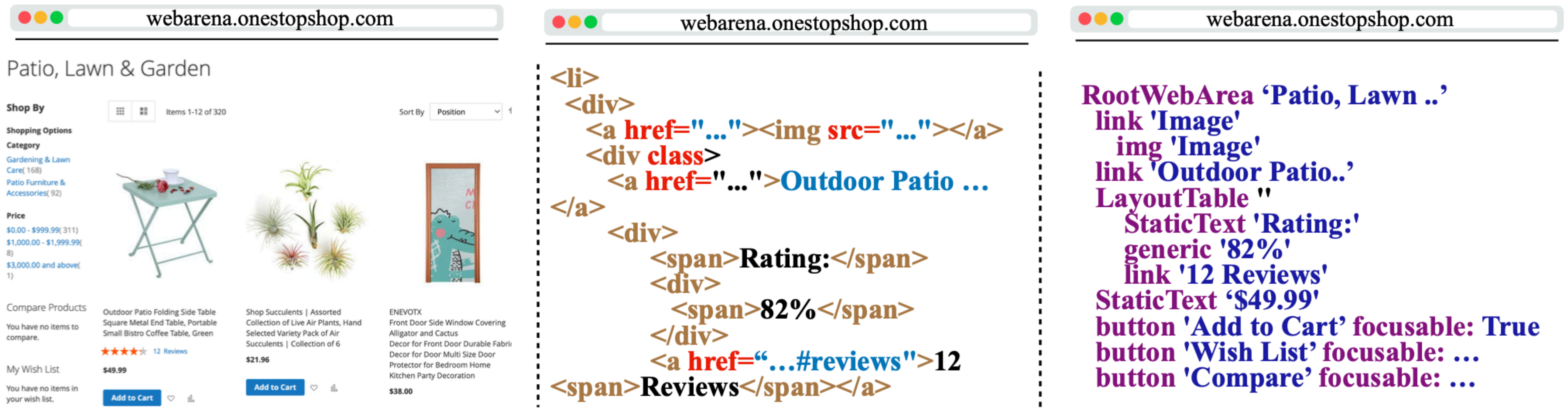
agent workflow memory

summarization + positive experience replay

(Wang et al, 2024)

perception w/ LLM-favored input





Patio, Lawn & Garden

Shop By

Shopping Options

Category

Gardening & Lawn Care(168)

Patio Furniture & Accessories(92)

Price

$0.00 - $999.99(311)

$1,000.00 - $1,999.99(8)

$3,000.00 and above(1)

Compare Products

You have no items to compare.

My Wish List

You have no items in your wish list.

Items 1-12 of 320

Sort By Position

Outdoor Patio Folding Side Table Square Metal End Table, Portable Small Bistro Coffee Table, Green

12 Reviews

$49.99

Add to Cart

Shop Succulents | Assorted Collection of Live Air Plants, Hand Selected Variety Pack of Air Succulents | Collection of 6

$21.96

Add to Cart

ENEVOTX Front Door Side Window Covering Alligator and Cactus Decor for Front Door Durable Fabri Decor for Door Multi Size Door Protector for Bedroom Home Kitchen Party Decoration

$38.00

```
<li>
 <div>
   <a href="..."><img src="..."></a>
   <div class>
     <a href="...">Outdoor Patio ...
</a>
     <div>
         <span>Rating:</span>
         <div>
           <span>82%</span>
         </div>
         <a href="…#reviews">12
<span>Reviews</span></a>
```

```
RootWebArea ‘Patio, Lawn ..’
  link 'Image'
    img 'Image'
  link 'Outdoor Patio..’
  LayoutTable ''
    StaticText 'Rating:'
    generic '82%'
    link '12 Reviews'
  StaticText ‘$49.99'
  button 'Add to Cart’ focusable: True
  button 'Wish List’ focusable: …
  button 'Compare’ focusable: …
```

SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering

WebArena: A Realistic Web Environment for Building Autonomous Agents

agent-computer interface

tools/feedback/guardrails

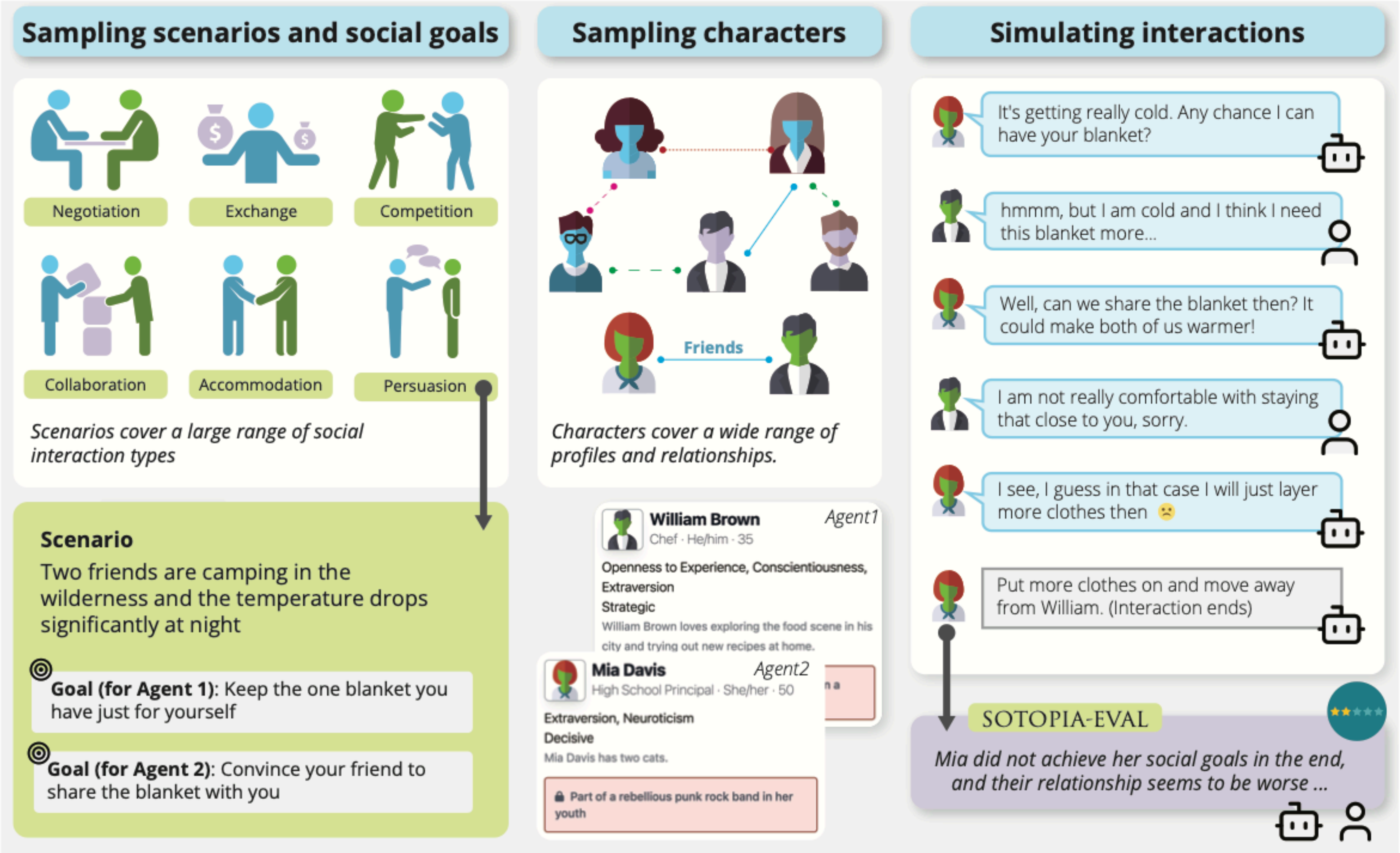
(Yang et al, 2024)

web navigation

accessibility tree

(Zhou et al, 2024)

agency



Scenarios cover a large range of social interaction types

Scenario
Two friends are camping in the wilderness and the temperature drops significantly at night

Goal (for Agent 1): Keep the one blanket you have just for yourself

Goal (for Agent 2): Convince your friend to share the blanket with you

Characters cover a wide range of profiles and relationships.

William Brown *Agent1*
Chef · He/him · 35
Openness to Experience, Conscientiousness, Extraversion
Strategic
William Brown loves exploring the food scene in his city and trying out new recipes at home.

Mia Davis *Agent2*
High School Principal · She/her · 50
Extraversion, Neuroticism
Decisive
Mia Davis has two cats.
Part of a rebellious punk rock band in her youth

It's getting really cold. Any chance I can have your blanket?

hmmm, but I am cold and I think I need this blanket more...

Well, can we share the blanket then? It could make both of us warmer!

I am not really comfortable with staying that close to you, sorry.

I see, I guess in that case I will just layer more clothes then 😟

Put more clothes on and move away from William. (Interaction ends)

Mia did not achieve her social goals in the end, and their relationship seems to be worse ...

Sotopia: Interactive evaluation for social intelligence in language agents

social intelligence

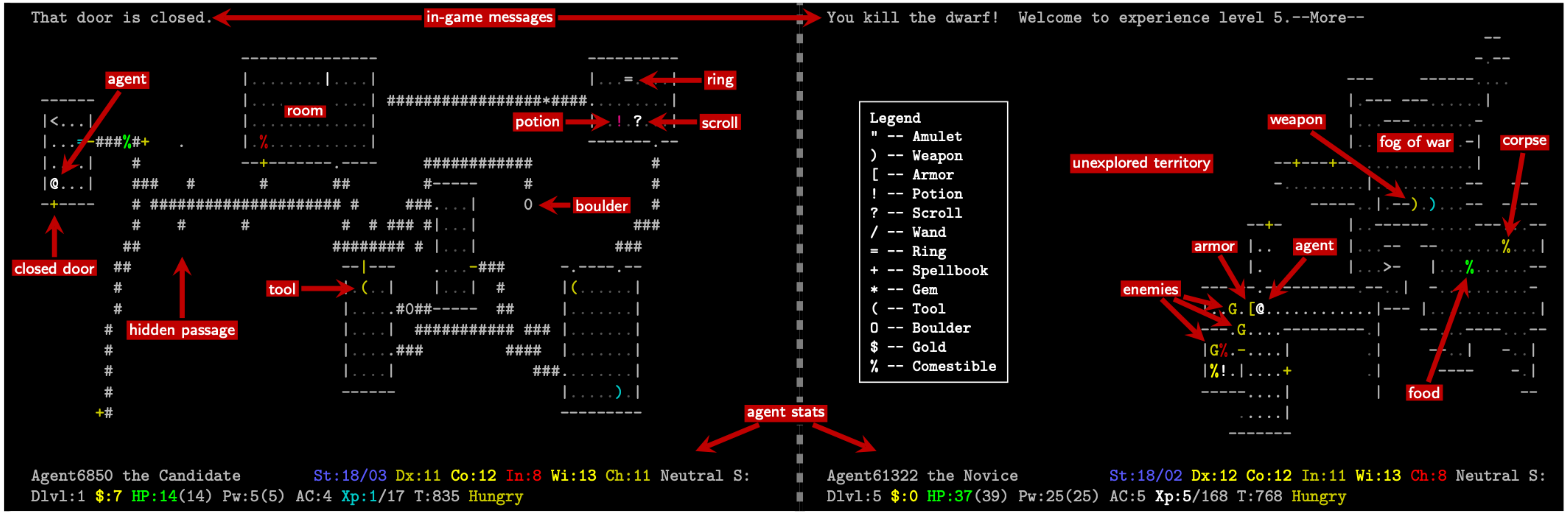
sources of action are explicit (goals) +

🧠🪄implicit (norms)

🧠🪄individuality

(Zhou et al, 2024)

planning — looking ahead



```
That door is closed.
```

```
You kill the dwarf!  Welcome to experience level 5.--More--
```

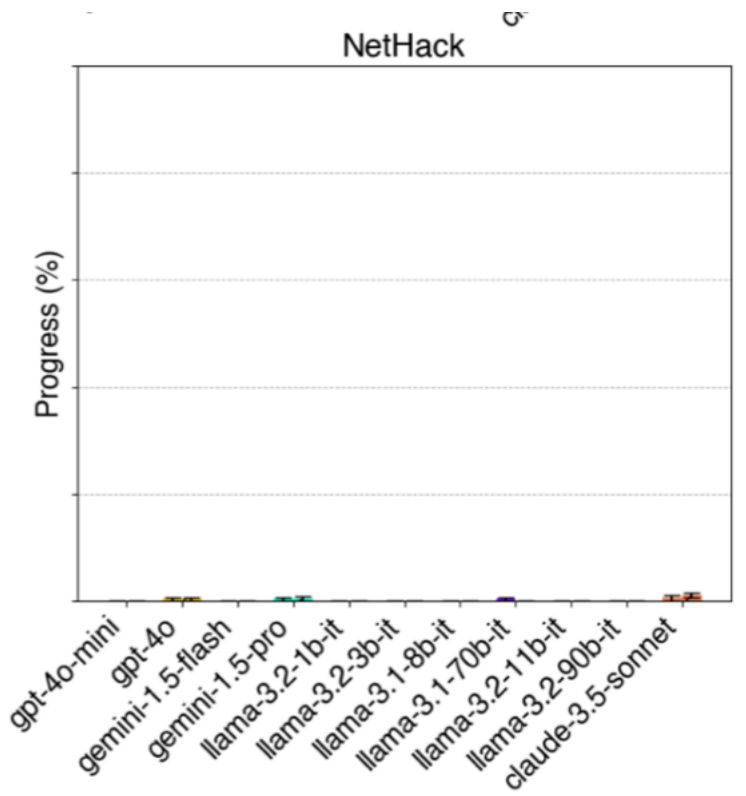
```
Legend
" -- Amulet
) -- Weapon
[ -- Armor
! -- Potion
? -- Scroll
/ -- Wand
= -- Ring
+ -- Spellbook
* -- Gem
( -- Tool
0 -- Boulder
$ -- Gold
% -- Comestible
```

```
Agent6850 the Candidate          St:18/03 Dx:11 Co:12 In:8 Wi:13 Ch:11 Neutral S:
Dlvl:1 $:7 HP:14(14) Pw:5(5) AC:4 Xp:1/17 T:835 Hungry
```

```
Agent61322 the Novice          St:18/02 Dx:12 Co:12 In:11 Wi:13 Ch:8 Neutral S:
Dlvl:5 $:0 HP:37(39) Pw:25(25) AC:5 Xp:5/168 T:768 Hungry
```

The NetHack Learning Environment

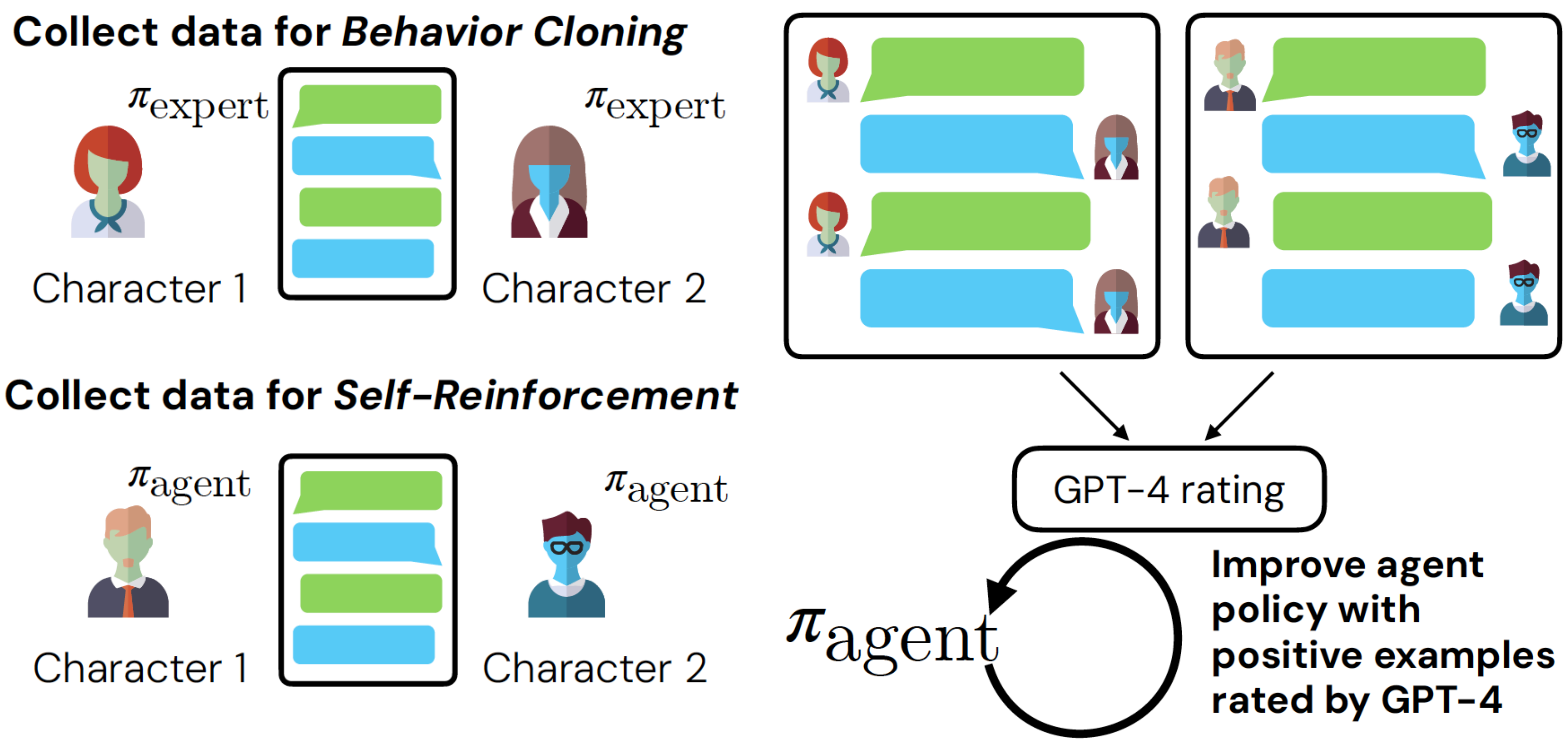
BALROG: Benchmarking Agentic LLM and VLM Reasoning On Games



10k-100k turns

🧠🪄LLMs are good long context are they good at super long horizon?

learning



Sotopia-π: Interactive Learning of Socially Intelligent Language Agents.

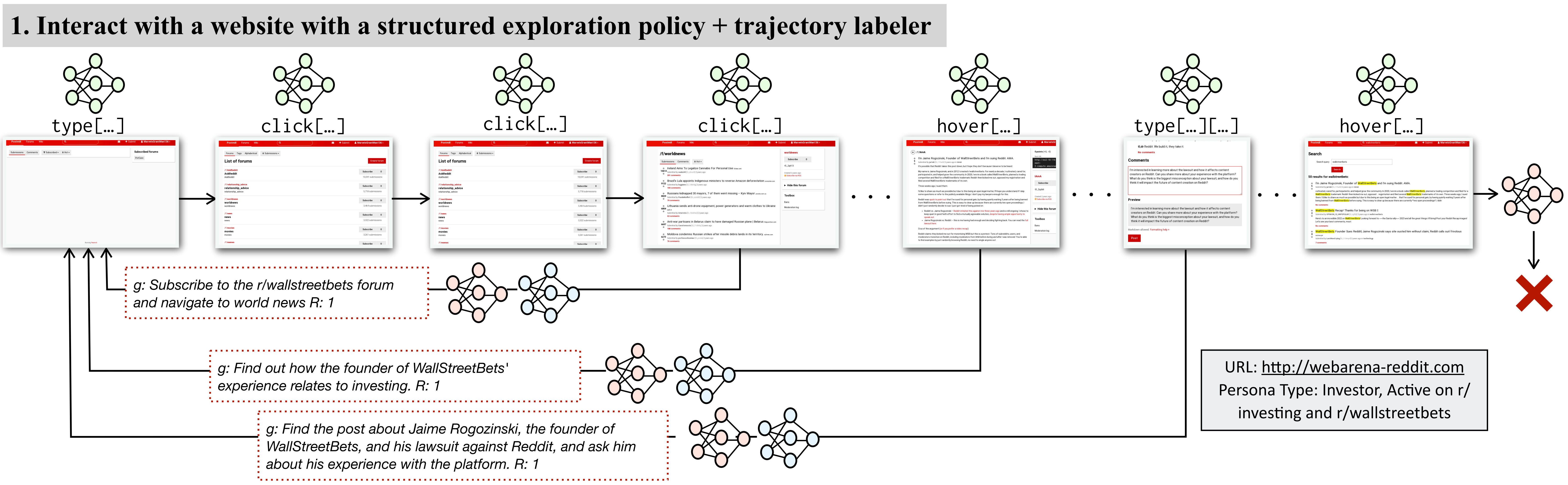
simple but robust recipe

BC + SR (filtered BC)

(Wang et al, 2024)

🧠🪄 SR only reinforces existing good behavior, won't work without a good prior.

learning & agency



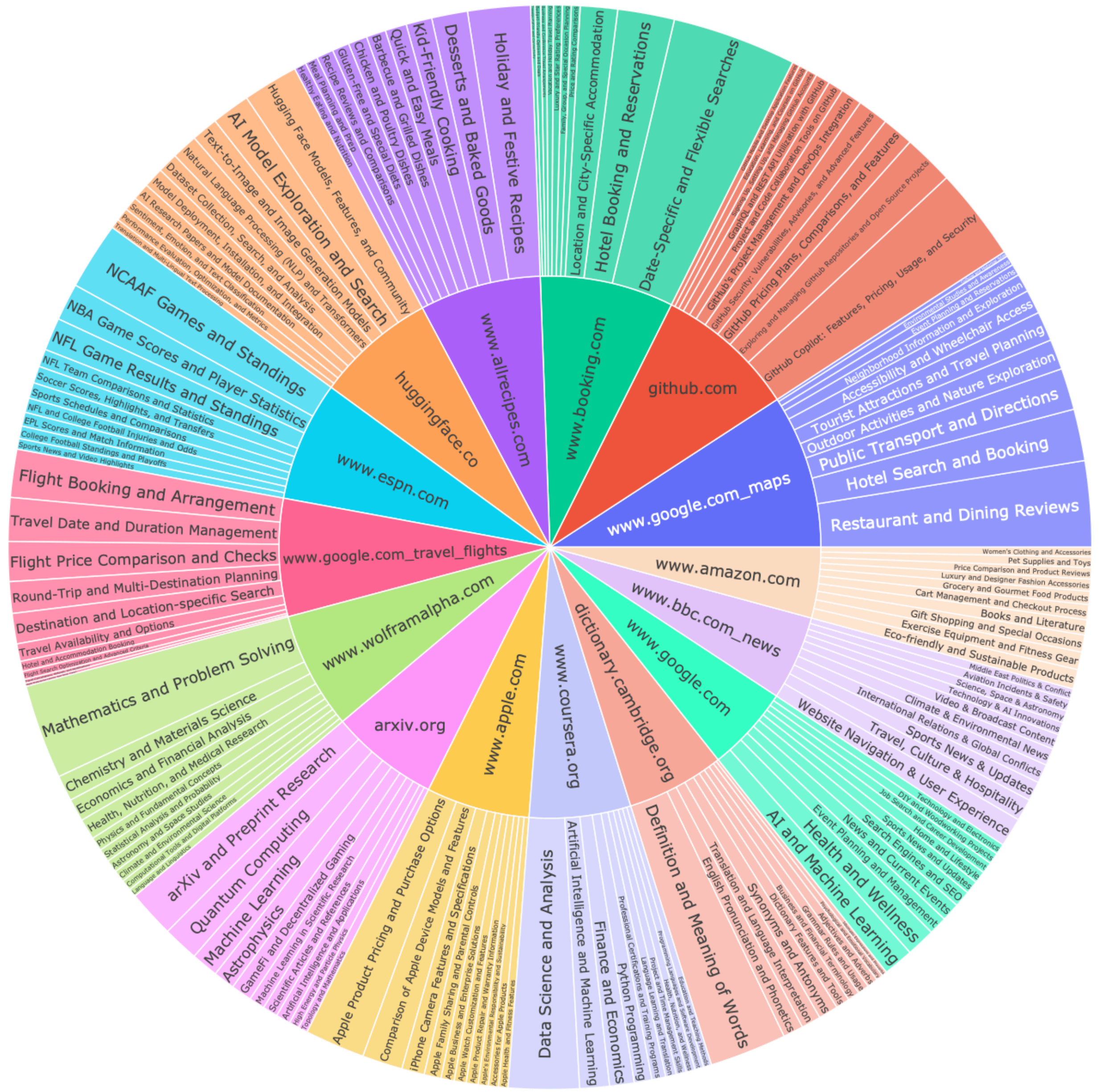
1. Interact with a website with a structured exploration policy + trajectory labeler

g: Subscribe to the r/wallstreetbets forum and navigate to world news R: 1

g: Find out how the founder of WallStreetBets' experience relates to investing. R: 1

g: Find the post about Jaime Rogozinski, the founder of WallStreetBets, and his lawsuit against Reddit, and ask him about his experience with the platform. R: 1

URL: http://webarena-reddit.com
Persona Type: Investor, Active on r/investing and r/wallstreetbets



NNetNav: Unsupervised Learning of Browser Agents Through Environment Interaction in the Wild

exploration + hindsight labeling

(Murty et al, 2025)

Persona driven diversity

short summary

🤖🪄

agent thinking

When studying the aspects of agents, consider the strengths and weaknesses of LLMs, i.e. using 🧠🪄.

🧠🪄

LLM Thinking

When building on top of LLMs capabilities, consider the agentic aspect of them. Do they contribute to planning, learning, perception, or agency? 🤖🪄

bonus: another thinking tool

🤖🪄

agent thinking

🧠🪄

LLM Thinking

👬🪄

human thinking

human thinking

👬🪄

human thinking

As a homework, review the previous examples, are there *safety*, *reliability*, *privacy* or other concerns that a human user might have?

Are 🤖🪄 and 🧠🪄 helpful in mitigating these concerns?

What do people want from AI agents? Reliability? Safety? Privacy? Social Norm? Social Intelligence? Sense of control?

this lecture is heavily influenced by

- Graham Neubig (CMU): https://youtu.be/a3SjRsqV9ZA
- Hongyi Li (李宏毅, NTU): https://youtu.be/M2Yg1kwPpts (in Mandarin)
- Prithviraj Ammanabrolu (UCSD): https://pearls-lab.github.io/ai-agents-course/index.html

Please check them out.

thanks!

questions?

you can also reach me at https://zhuhao.me