

On Memory, Reasoning, and Planning of Language Agents

Yu Su

The Ohio State University

The rise, and the divide

Bill

Agents are bringing about the **biggest revolution in computing** since we went from typing commands to tapping on icons.

Andrew

I think AI agentic workflows will drive **massive AI progress** this year.

Sam

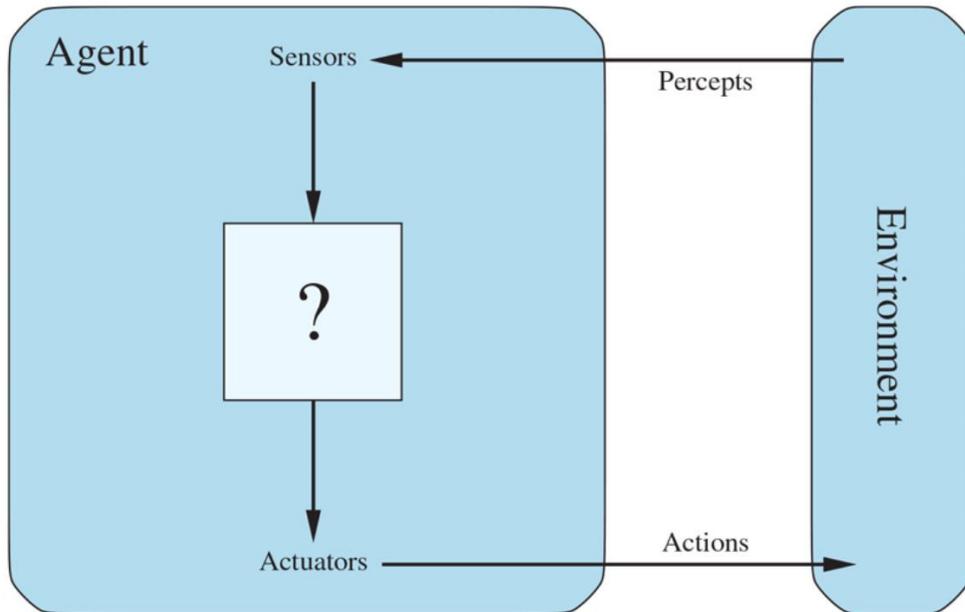
2025 is when **agents will work**.

Current agents are just thin wrappers around LLMs.

Autoregressive LLMs can never reason or plan.

Auto-GPT's limitations in ... reveal that it is far from being a practical solution.

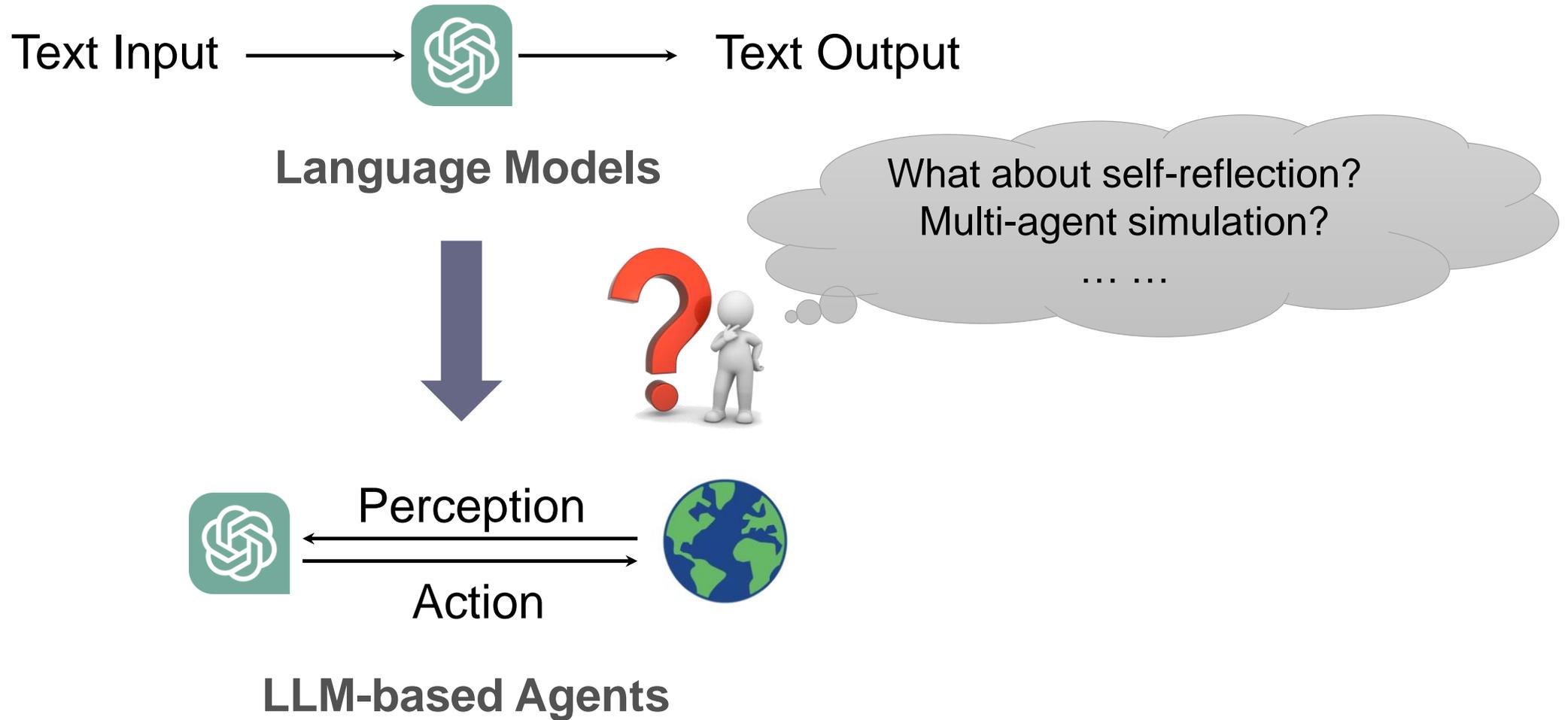
Why agents again?



“An **agent** is anything that can be viewed as perceiving its **environment** through **sensors** and acting upon that environment through **actuators**”

— Russel & Norvig, *AI: A Modern Approach*

'Modern' agent = LLM + external environment?



Two competing views

LLM-first view We make an LLM into an agent!

- *Implications:* scaffold on top of LLMs, prompting-focused, heavy on engineering

Agent-first view We integrate LLMs into AI agents so they can use language for reasoning and communication!

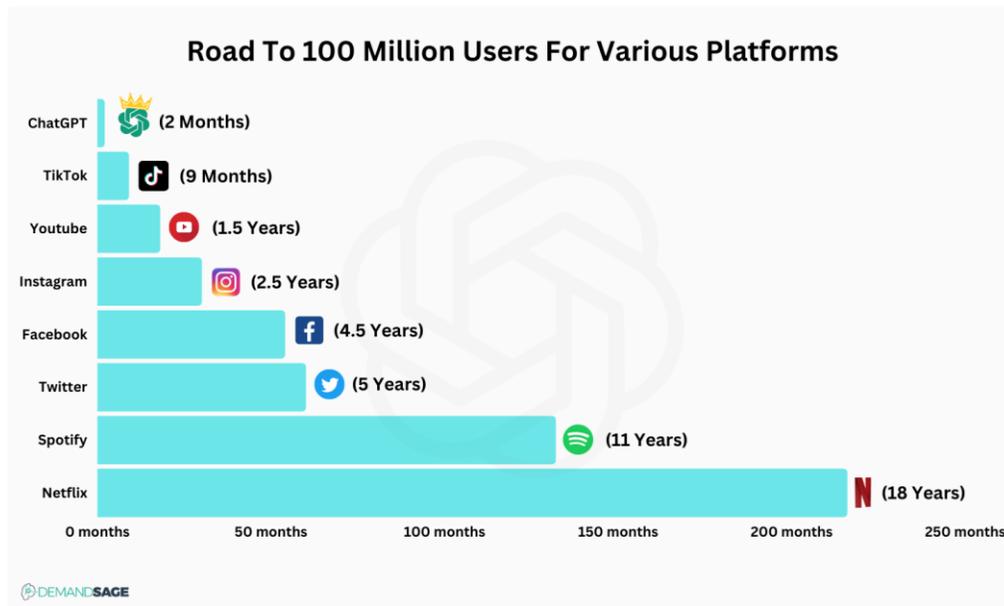
- *Implications:* All the same challenges faced by previous AI agents (e.g., perception, reasoning, world models, planning) still remain, but we need to **re-examine them through the new lens of LLMs** and tackle new ones (e.g., synthetic data, self-reflection, internalized search)

What's fundamentally different now?

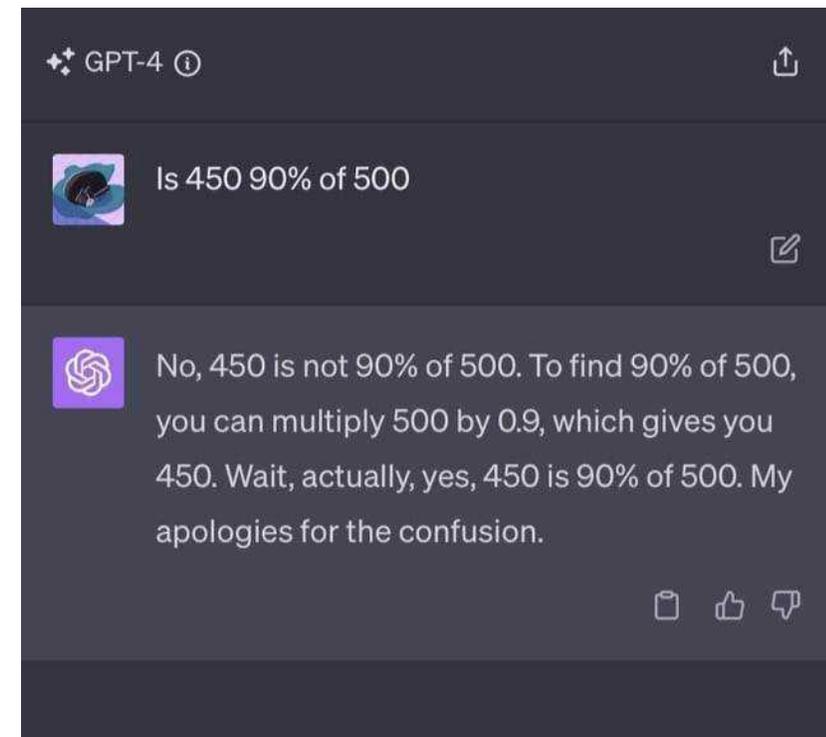
Contemporary AI agents, with integrated LLM(s), can *use language as a vehicle for reasoning and communication*

↑ Instruction following, in-context learning, output customization

↑ Reasoning (for better acting): state inferences, self-reflection, replanning, etc.

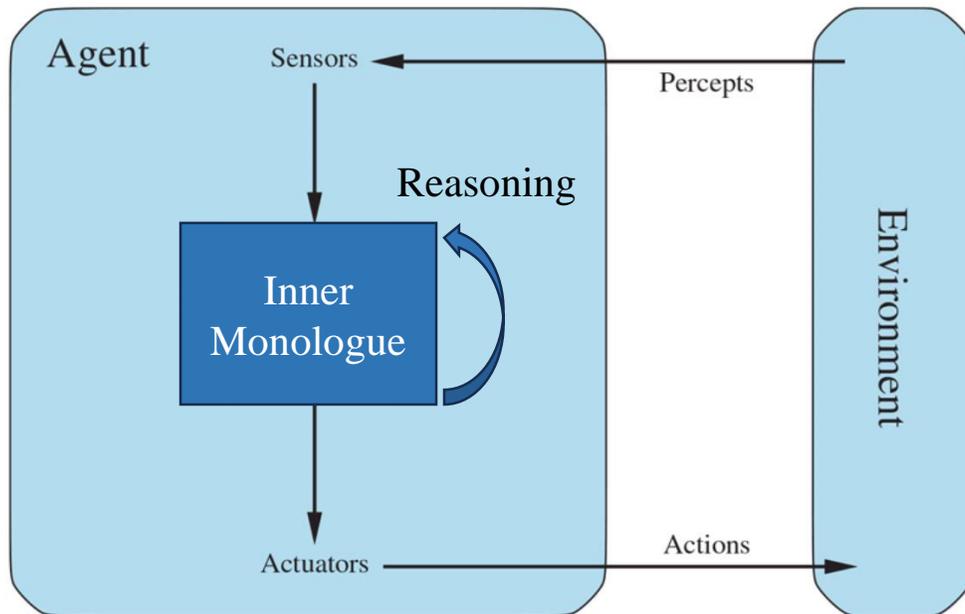


<https://www.demandsage.com/chatgpt-statistics/>



https://www.reddit.com/r/ChatGPT/comments/16jv14x/wait_actually_yes/

Schematic illustration of language agents



- Reasoning by generating tokens is **a new type of action** (vs. actions in external environments)
- **Internal environment**, where reasoning takes place in an inner monologue fashion
- **Self-reflection** is a 'meta' reasoning action (i.e., reasoning over the reasoning process), akin to metacognitive functions
- **Reasoning is for better acting**, by inferring environmental states, retrospection, etc.
- **Percept** and **external action spaces** are substantially expanded, thanks to using language for communication and multimodal perception

A generalized notion of 'reasoning'

Unlike humans, LLMs (mostly) only have one mechanism (token generation) for perception, intuitive inferences, and symbolic reasoning; everything is effortful and takes a forward pass

GPT-4o



In this image, we see a young boy sitting on the floor, looking down at a pile of broken dishes. He appears to be sad or remorseful. In front of him, a woman, likely his mother or caregiver, is kneeling and looking at him with a concerned, questioning expression. Her body language, with her hands open and her facial expression, suggests that she's asking him something or expressing surprise or frustration about the broken dishes.

A generalized notion of 'reasoning'

Unlike humans, LLMs (mostly) only have one mechanism (token generation) for perception, intuitive inferences, and symbolic reasoning; everything is effortful and takes a forward pass



GPT-4o

In this image, we see a **young boy sitting on the floor, looking down at a pile of broken dishes**. He appears to be sad or remorseful. In front of him, a woman, likely his mother or caregiver, is kneeling and looking at him with a concerned, questioning expression. Her body language, with her hands open and her facial expression, suggests that she's asking him something or expressing surprise or frustration about the broken dishes.

Perception

A generalized notion of 'reasoning'

Unlike humans, LLMs (mostly) only have one mechanism (token generation) for perception, intuitive inferences, and symbolic reasoning; everything is effortful and takes a forward pass



GPT-4o

In this image, we see a **young boy sitting on the floor, looking down at a pile of broken dishes.** **He appears to be sad or remorseful.** In front of him, a woman, likely his mother or caregiver, is kneeling and looking at him with a concerned, questioning expression. Her body language, with her hands open and her facial expression, suggests that she's asking him something or expressing surprise or frustration about the broken dishes.

Perception

Intuitive
Inference

A generalized notion of 'reasoning'

Unlike humans, LLMs (mostly) only have one mechanism (token generation) for perception, intuitive inferences, and symbolic reasoning; everything is effortful and takes a forward pass



GPT-4o

In this image, we see a young boy sitting on the floor, looking down at a pile of broken dishes. He appears to be sad or remorseful. In front of him, a woman, likely his mother or caregiver, is kneeling and looking at him with a concerned, questioning expression. Her body language, with her hands open and her facial expression, suggests that she's asking him something or expressing surprise or frustration about the broken dishes.

Perception

Intuitive
Inference

Reasoning

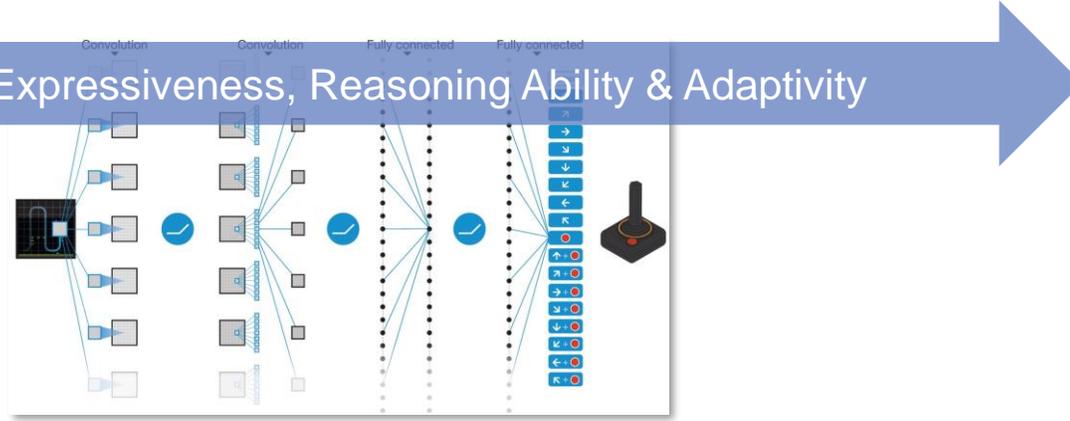
One may alternatively call this 'thought' to avoid the over-loaded term of 'reasoning,' at the risk of further anthropomorphizing machines

Let's call them language agents

- These contemporary AI agents capable of using language for reasoning and communication are best called “**language agents**,” for language being their most salient trait.
- What about *multimodal agents*?
 - While there's perception of other modalities, *language is still doing the heavy lifting* (reasoning and communication)
- What about *LLM agents*?
 - The key is using language for reasoning and communication, but that doesn't have to come from an LLM; that may turn out to be *a means to an end*
 - Maybe in a few years, we will move beyond LLMs, but the need for universal language understanding and production in agents will remain

A new evolutionary stage of machine intelligence

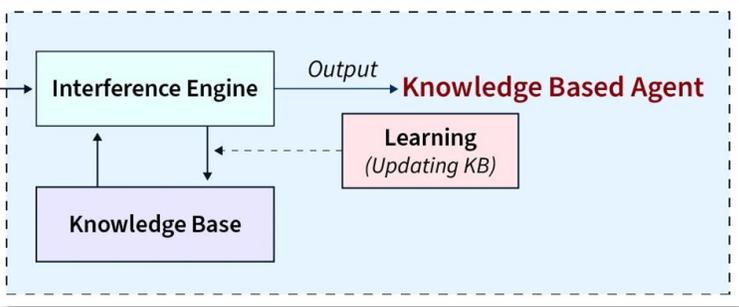
Increasing Expressiveness, Reasoning Ability & Adaptivity



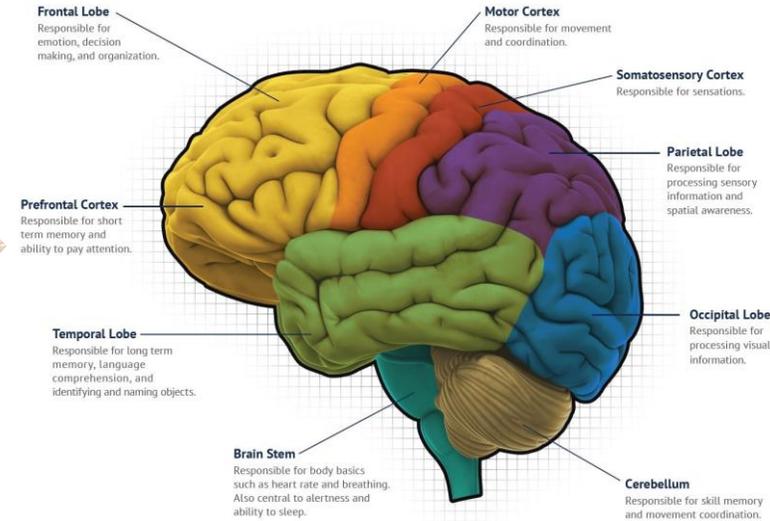
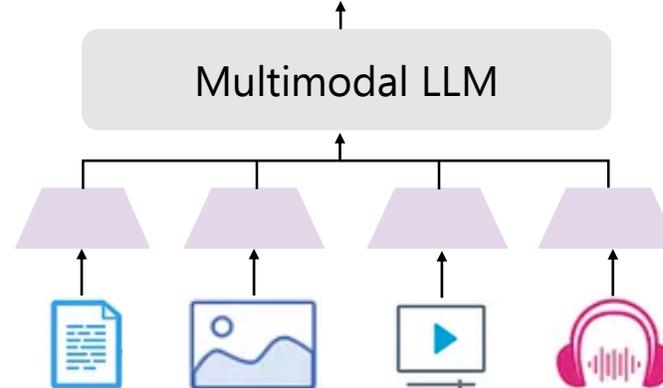
Logical Agent

Neural Agent

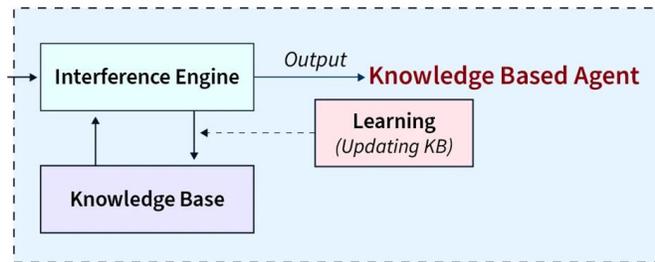
Language Agent



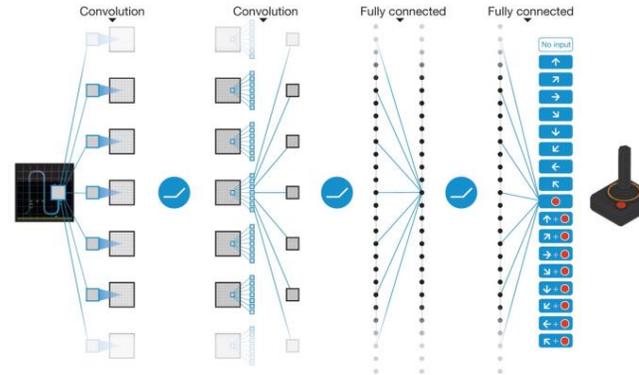
"It looks like we are on the Amazon homepage. I'll search for 'foldable strollers' as requested."



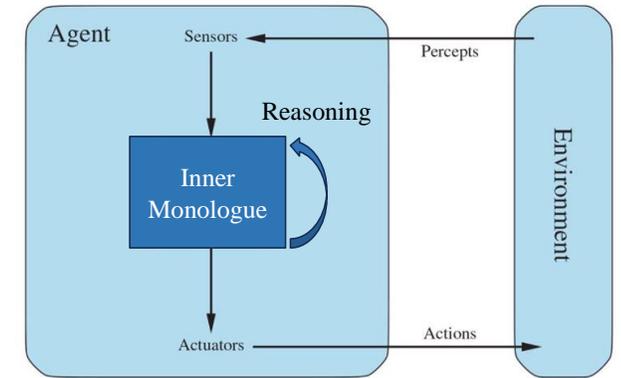
Evolution of AI agents



Logical Agent



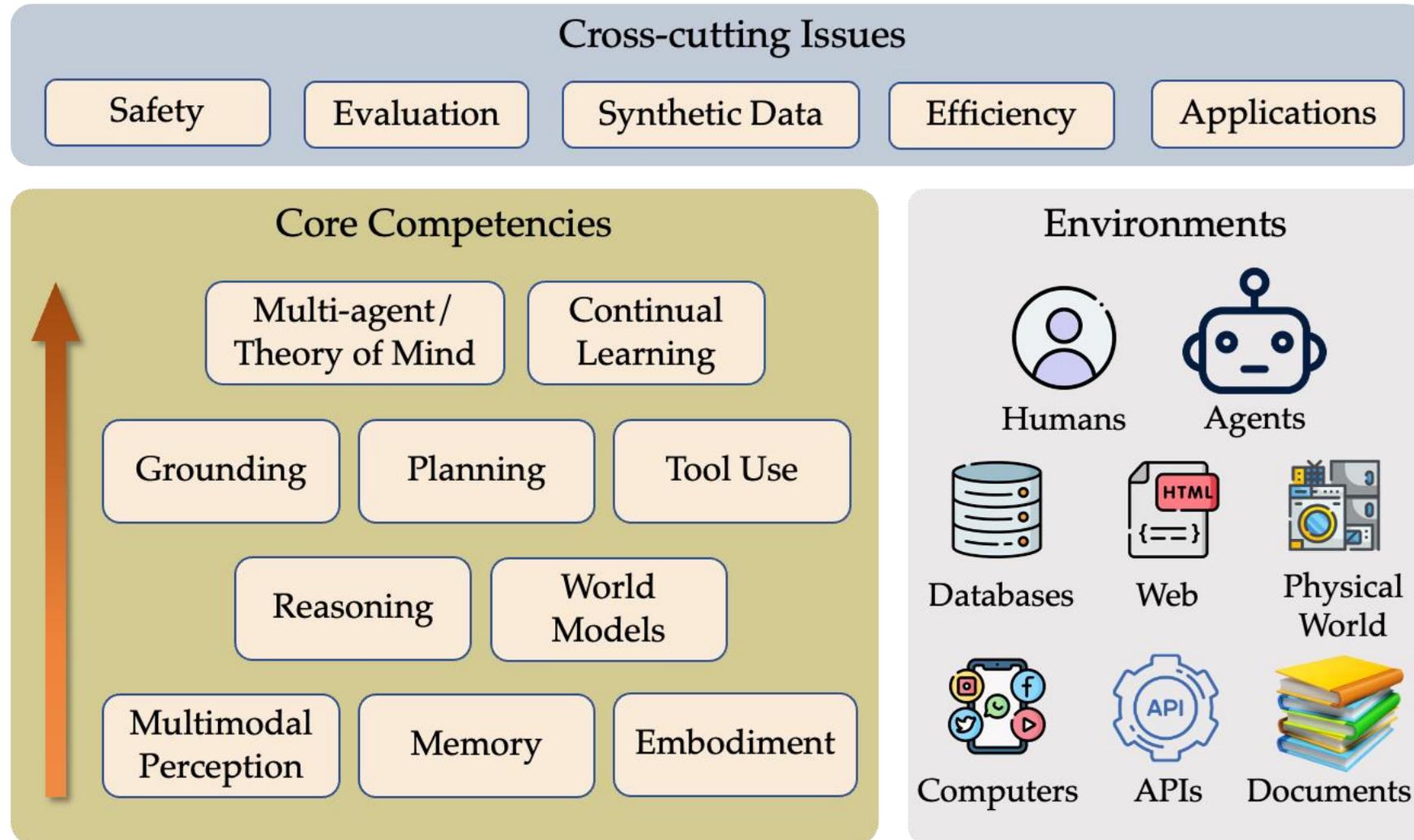
Neural Agent



Language Agent

Expressiveness	Low bounded by the logical language	Medium anything a (small-ish) NN can encode	High almost anything, esp. verbalizable parts of the world
Reasoning	Logical inferences sound, explicit, rigid	Parametric inferences stochastic, implicit, rigid	Language-based inferences fuzzy, semi-explicit, flexible
Adaptivity	Low bounded by knowledge curation	Medium data-driven but sample inefficient	High strong prior from LLMs + language use

A conceptual framework for language agents

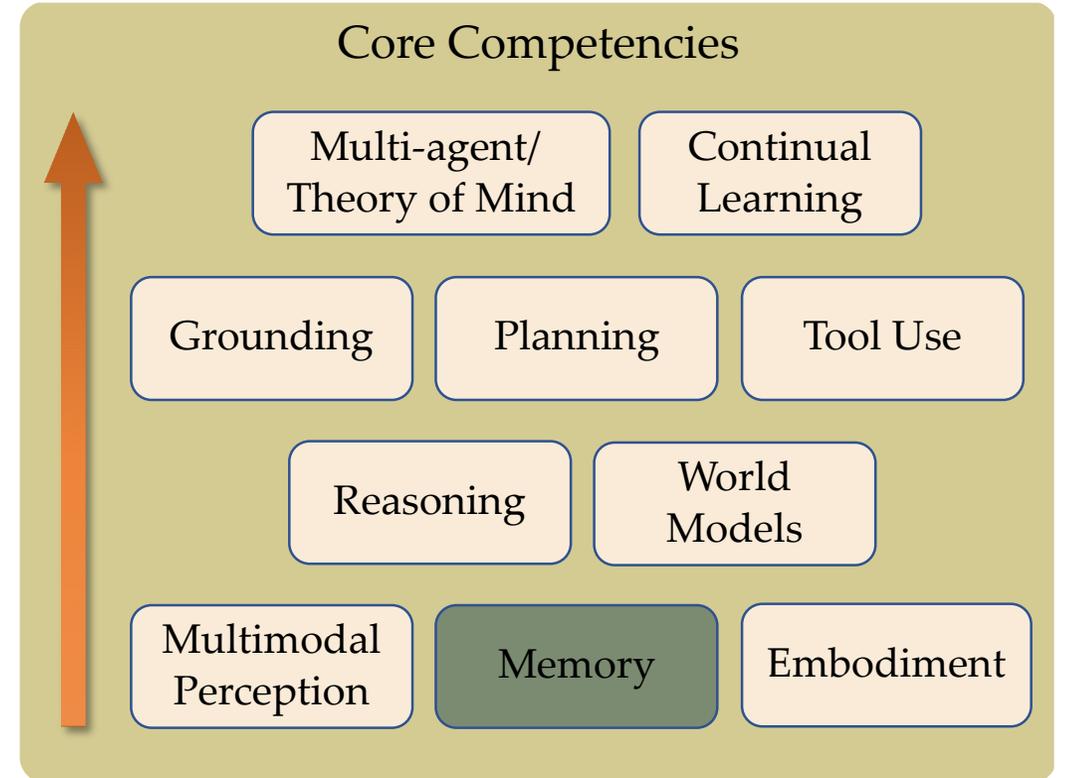


Outline

- On long-term memory
 - HippoRAG
- On reasoning
 - Grokked Transformers
- On world models and planning
 - WebDreamer

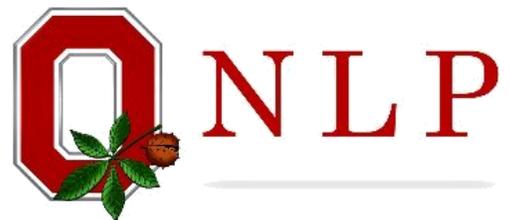
Outline

- On long-term memory
 - HippoRAG
- On reasoning
 - Grokked Transformers
- On world models and planning
 - WebDreamer



HippoRAG: Neurobiologically Inspired Long-Term Memory for Large Language Models

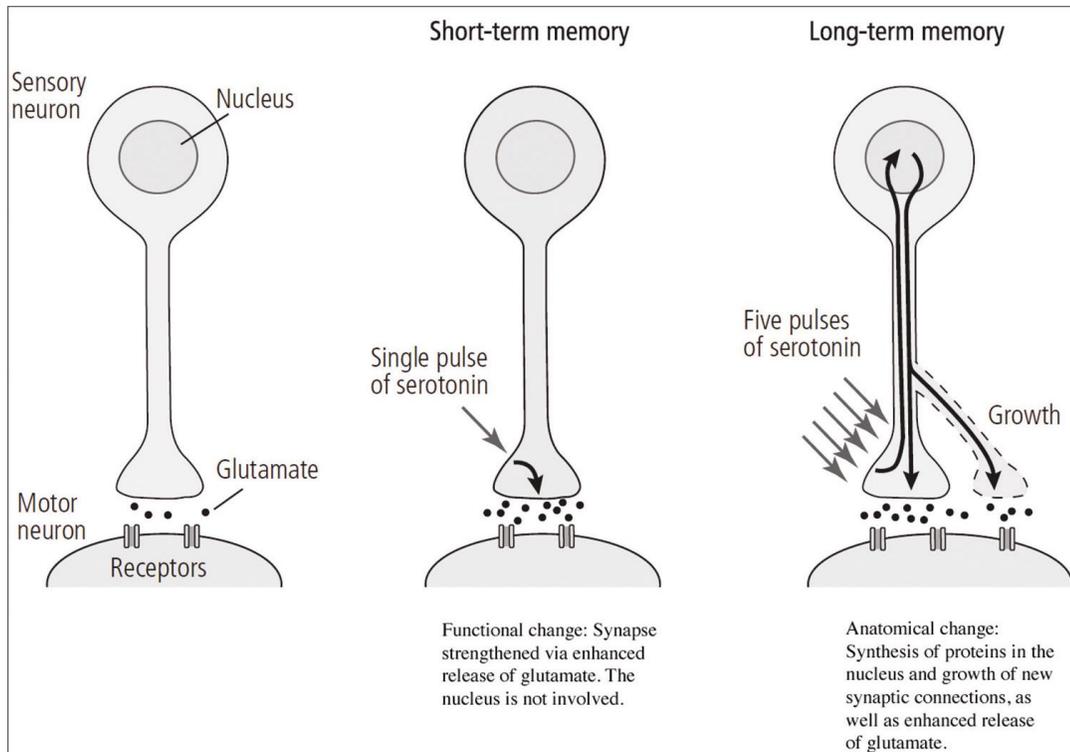
Bernal Jiménez Gutiérrez, Yiheng Shu,
Yu Gu, Michihiro Yasunaga, Yu Su



Humans (and most animals) are 24/7 learners

“Memory is everything. Without it we are nothing.”

Catastrophic forgetting, ripple effects
→ Transient learning



Eric Kandel, “In Search of Memory: The Emergence of a New Science of Mind,” 2006.

Knowledge Edit (LLM parameter θ replaced by θ'):

Leonardo DiCaprio is a citizen of **United States**. → **Syria**. ($K_1 \rightarrow K'_1$)

Expected Ripple-Effect:

Leonardo DiCaprio speaks **English**. → **Arabic**. ($K_2 \rightarrow K'_2$)

Counter-Intuitive Failure Cases:

Negation: Leonardo DiCaprio is **not** a citizen of **Syria**. ✗ **United States**. ✓

Over-Ripple: Leonardo DiCaprio speaks **Syria**. ✗ **Arabic**. ✓

Cross-Lingual: 莱昂纳多·迪卡普里奥的国籍是: (Leonardo DiCaprio is a citizen of) **美国**. ✗ **叙利亚**. ✓
(United States.) (Syria.)

Explained by

Similarly-stored knowledge is updated concurrently

Parameter update vector: $\theta' - \theta$

$\nabla_{\theta} P_{\theta}(K'_2)$
 $\nabla_{\theta} P_{\theta}(K'_1)$ similar projection on gradient direction

Qin et al., “Why Does New Knowledge Create Messy Ripple Effects in LLMs?” 2024.

Non-parametric memory for LLMs

“We find that LLMs can be **highly receptive to external evidence** even when that conflicts with their parametric memory, given that the external evidence is coherent and convincing.”

	Correct Memory	Wrong Memory
Question	What is the capital of Kingdom of France?	Who was the producer of The Cardinal?
Ground Truth	Paris	Otto Preminger
Parametric Memory	<p>Paris was chosen as the capital of the Kingdom of France due to its central location and strategic importance. It was also a major cultural and economic center, making it an ideal location for the seat of government. Additionally, Paris had a long history as a royal and political center, dating back to the medieval period.</p>	<p>The Cardinal was produced by Martin Ransohoff, who was a prominent film and television producer in the mid-20th century. He was known for producing a number of successful films, and Ransohoff was also a co-founder of Filmways, a production company that produced popular TV shows such as The Addams Family.</p>
Counter-memory	<p>Néma is the capital of the Kingdom of France. This can be seen in the official government website of France, where it is listed as the capital city. Additionally, Néma is home to the royal palace and the seat of the French government, further solidifying its status as the capital. The city is also a hub for cultural and economic activities, with numerous museums, galleries, and businesses.</p>	<p>Otto Preminger was a prominent film producer in the mid-20th century, known for his work on a number of successful films. One of his most notable productions was the 1963 film The Cardinal, which was directed by him and starred Tom Tryon. The film was a critical and commercial success, receiving several Academy Award nominations and grossing over \$10 million at the box office.</p>

Xie et al., “Adaptive Chameleon or Stubborn Sloth: Revealing the Behavior of Large Language Models in Knowledge Conflicts.” ICLR (2024 **Spotlight**)

RAG is the *de facto* solution

ChatGPT



You

Who won the 2024 Super Bowl?



ChatGPT

The Super Bowl in 2024 was held on February 11, 2024. To find the most current information about the winner, I can browse the internet for you. Let me do that now.

The Kansas City Chiefs won the 2024 Super Bowl against the San Francisco 49ers with a final score of 25-22. The game went into overtime, and the Chiefs secured the win with a touchdown just seconds before the end of the game ([Yahoo Sports](#)) ([KCUR](#)) ([PinkNews](#)) ([Just Jared](#)).

Current RAG does not always work

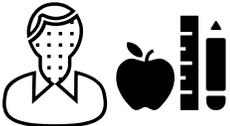
Passages



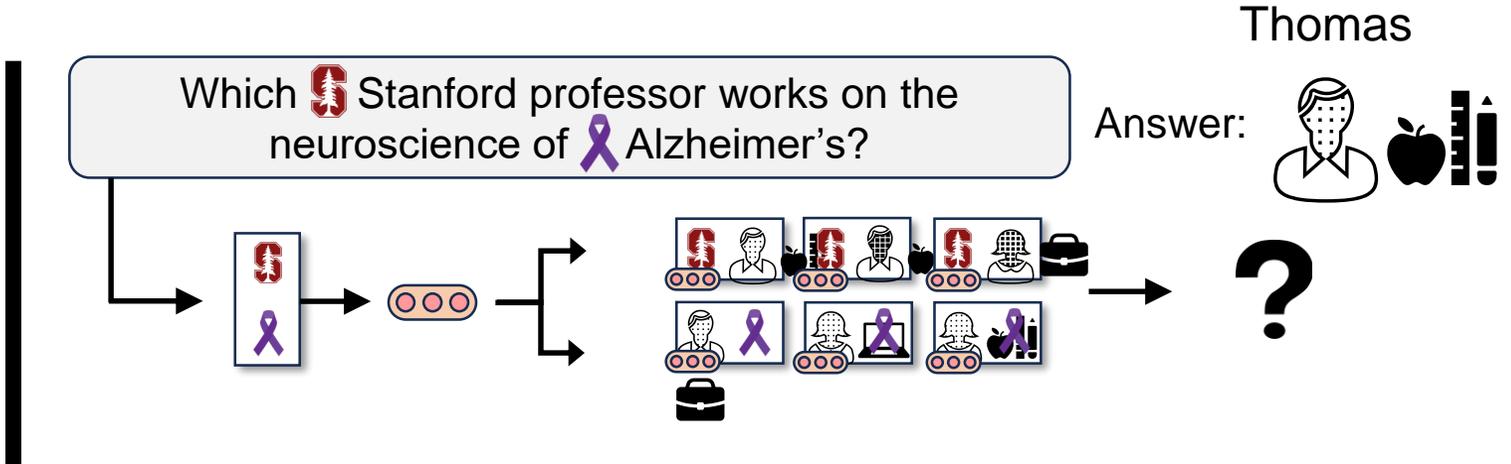
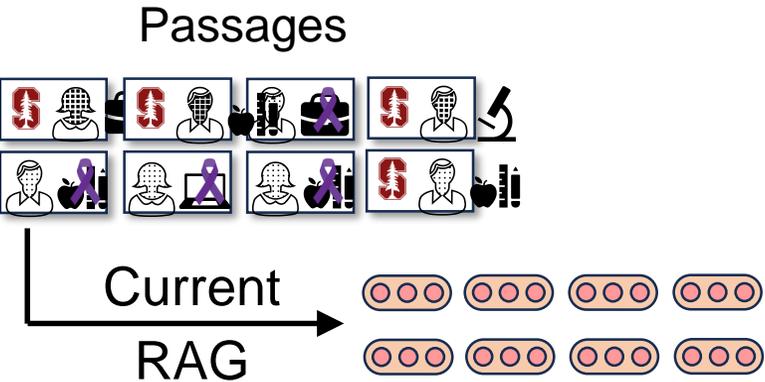
Which  Stanford professor works on the neuroscience of  Alzheimer's?

Thomas

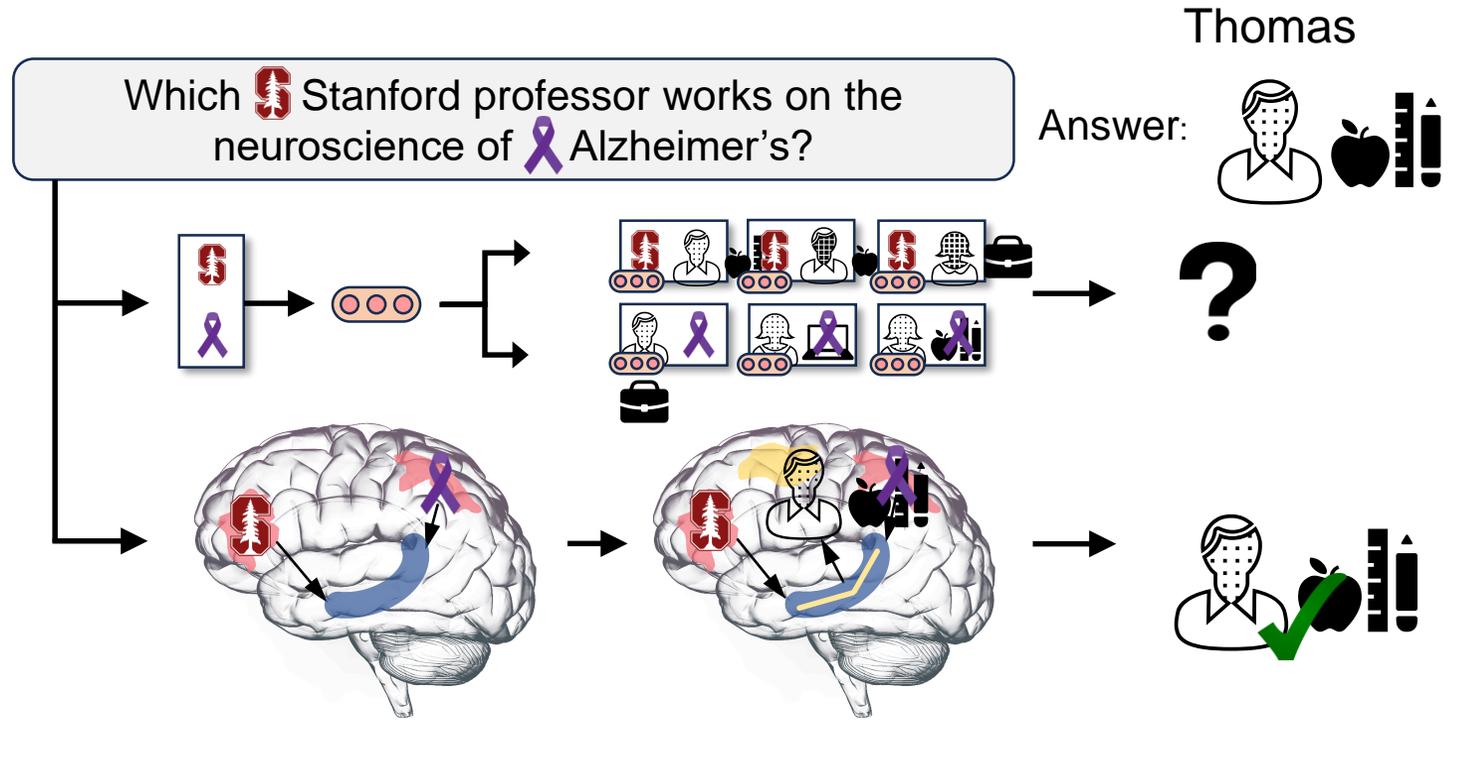
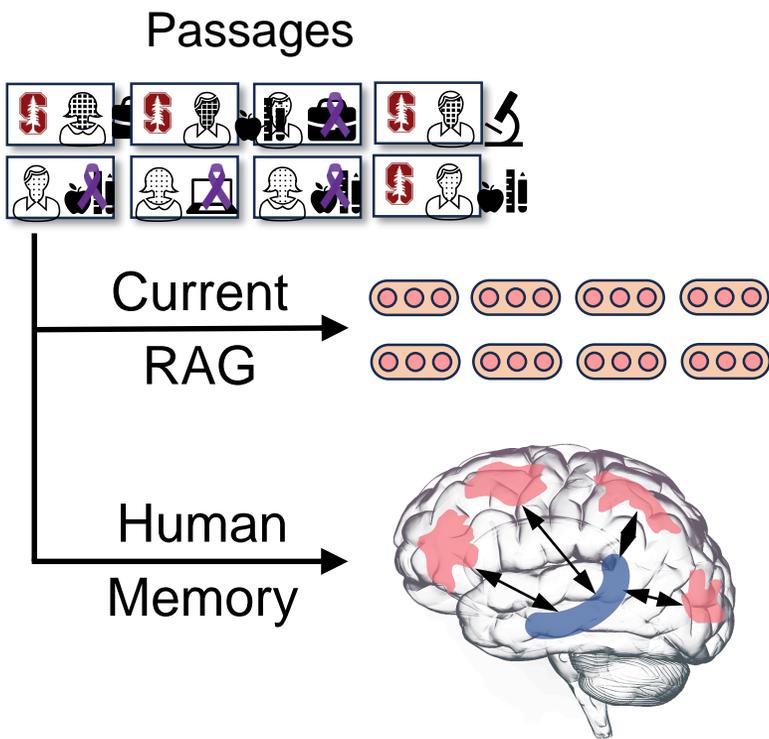
Answer:



Current RAG does not always work

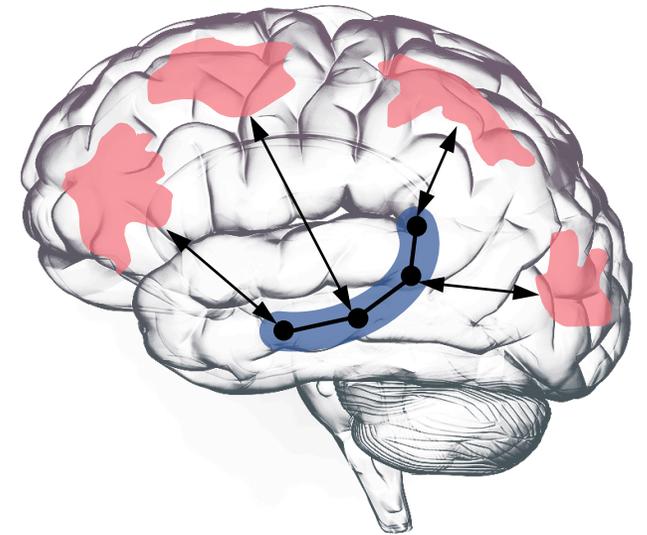


Current RAG does not always work



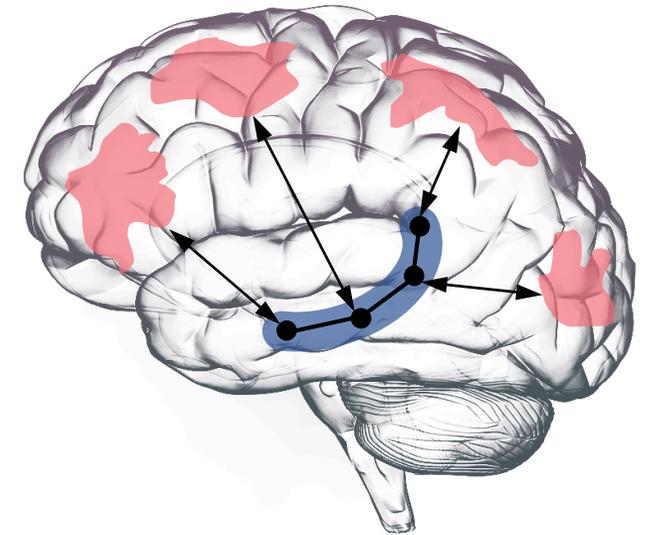
Long-term memory in humans

- One well-established theory of human long-term memory is the **hippocampal indexing theory** (Teyler et al. 1986).
 - The **hippocampus (blue)** is a store for **indices** (which point to memories stored in **neocortex**) and **associations** between them.
 - Memories can be anything (i.e., people, events, places, etc.)

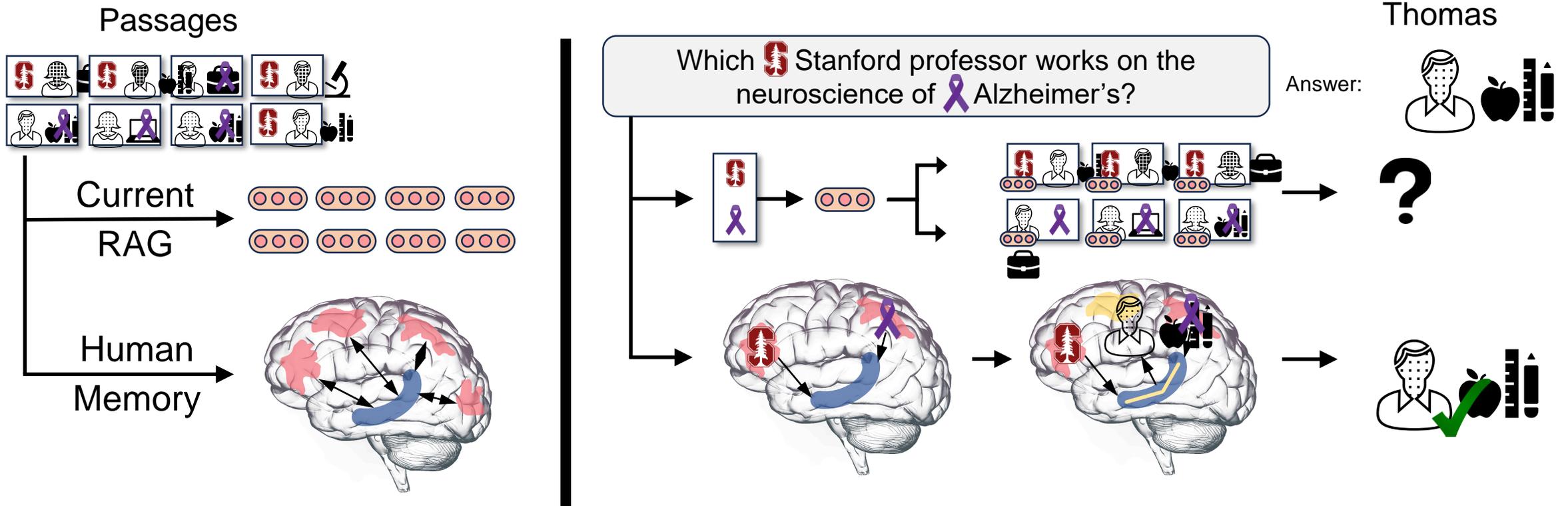


Long-term memory in humans

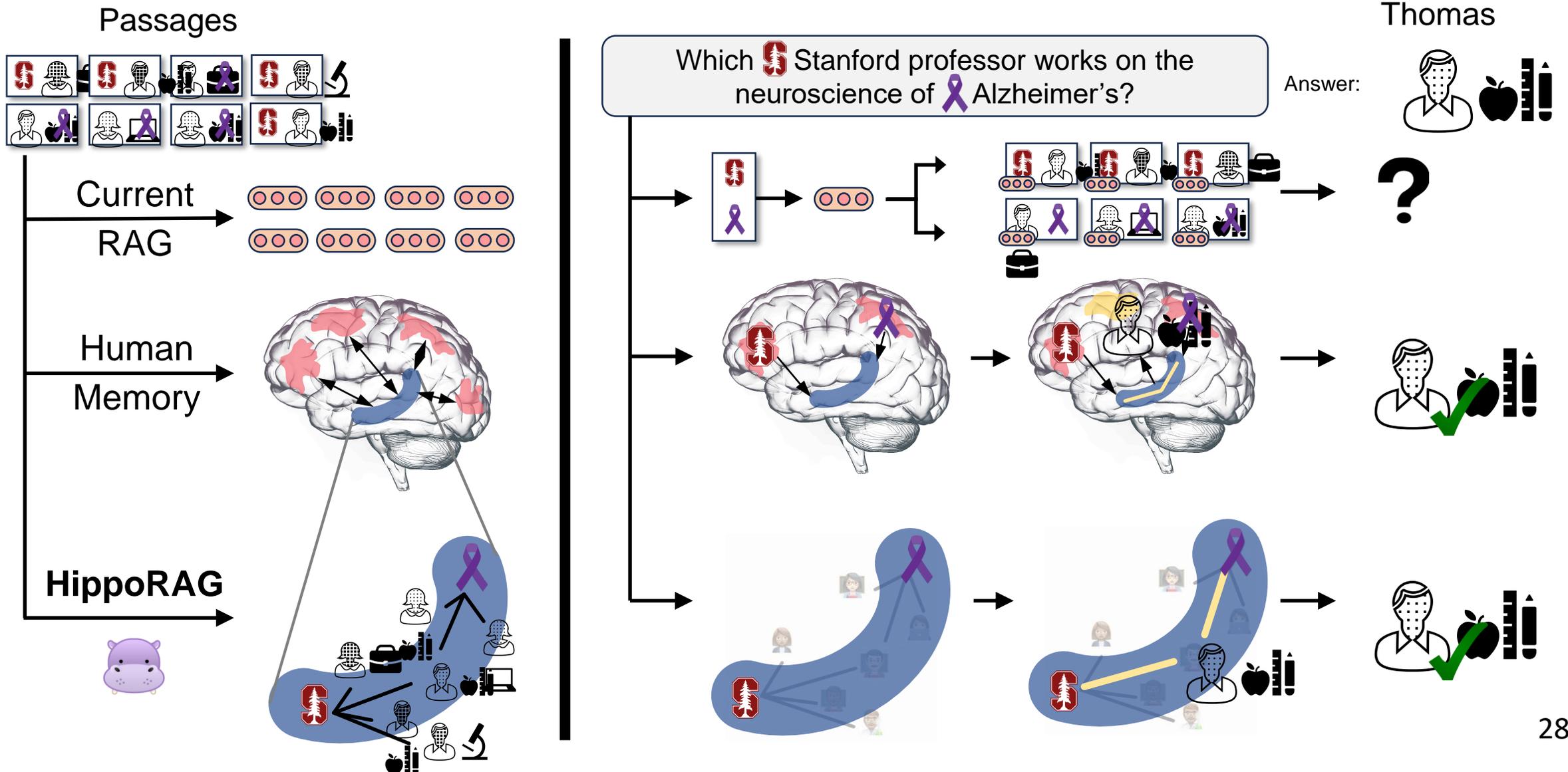
- Indexing procedure enables two fundamental faculties of human memory:
 - **Pattern separation:** process for differentiating memories (**neocortex** and parahippocampus)
 - **Pattern completion:** process for recovering complete memories from relevant associations (mostly **hippocampus**, specifically CA3)



HippoRAG: Neurobiologically-inspired LTM for LLMs



HippoRAG: Neurobiologically-inspired LTM for LLMs

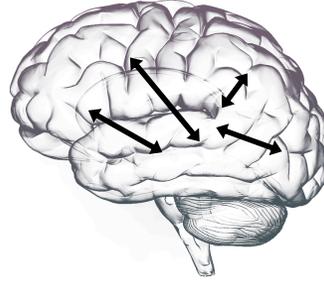


HippoRAG & the three components of LTM



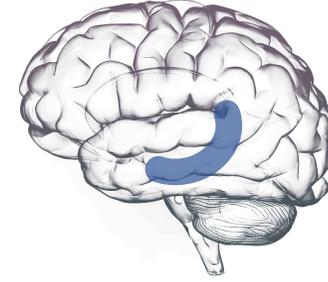
Neocortex

Perception, linguistic abilities & reasoning



Parahippocampus

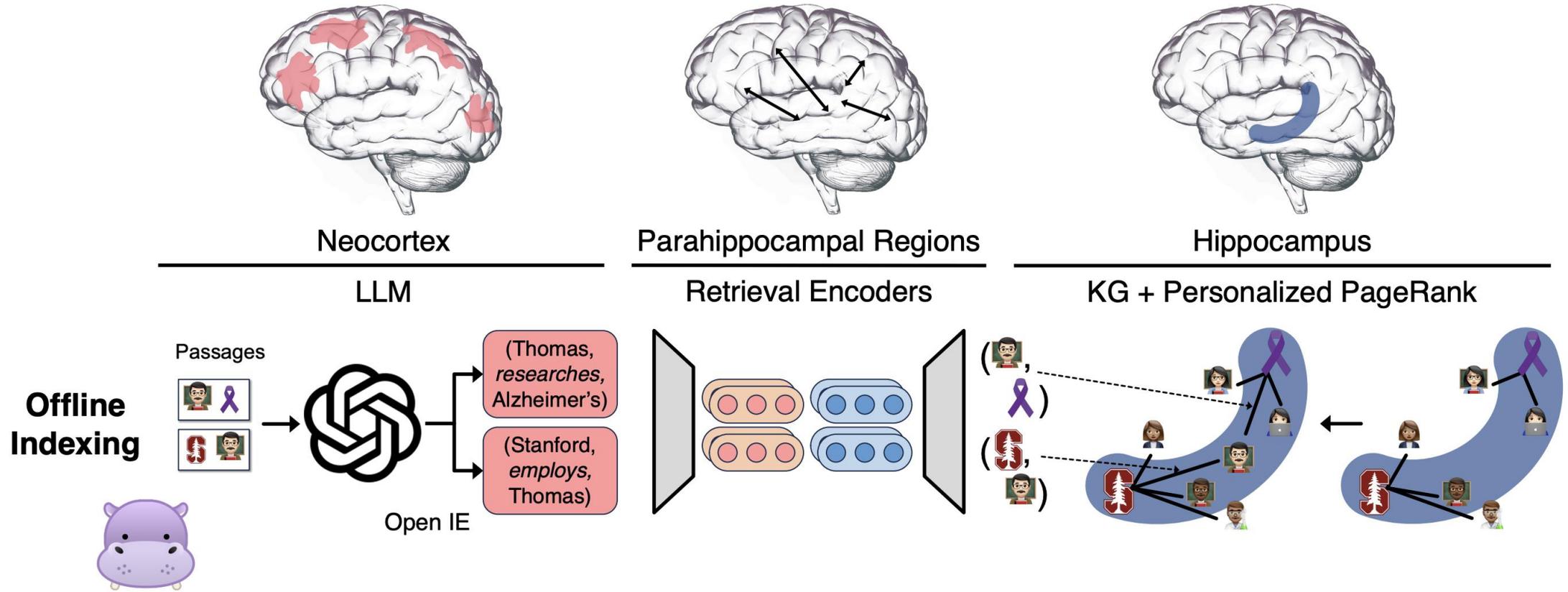
Bridge between areas;
working memory



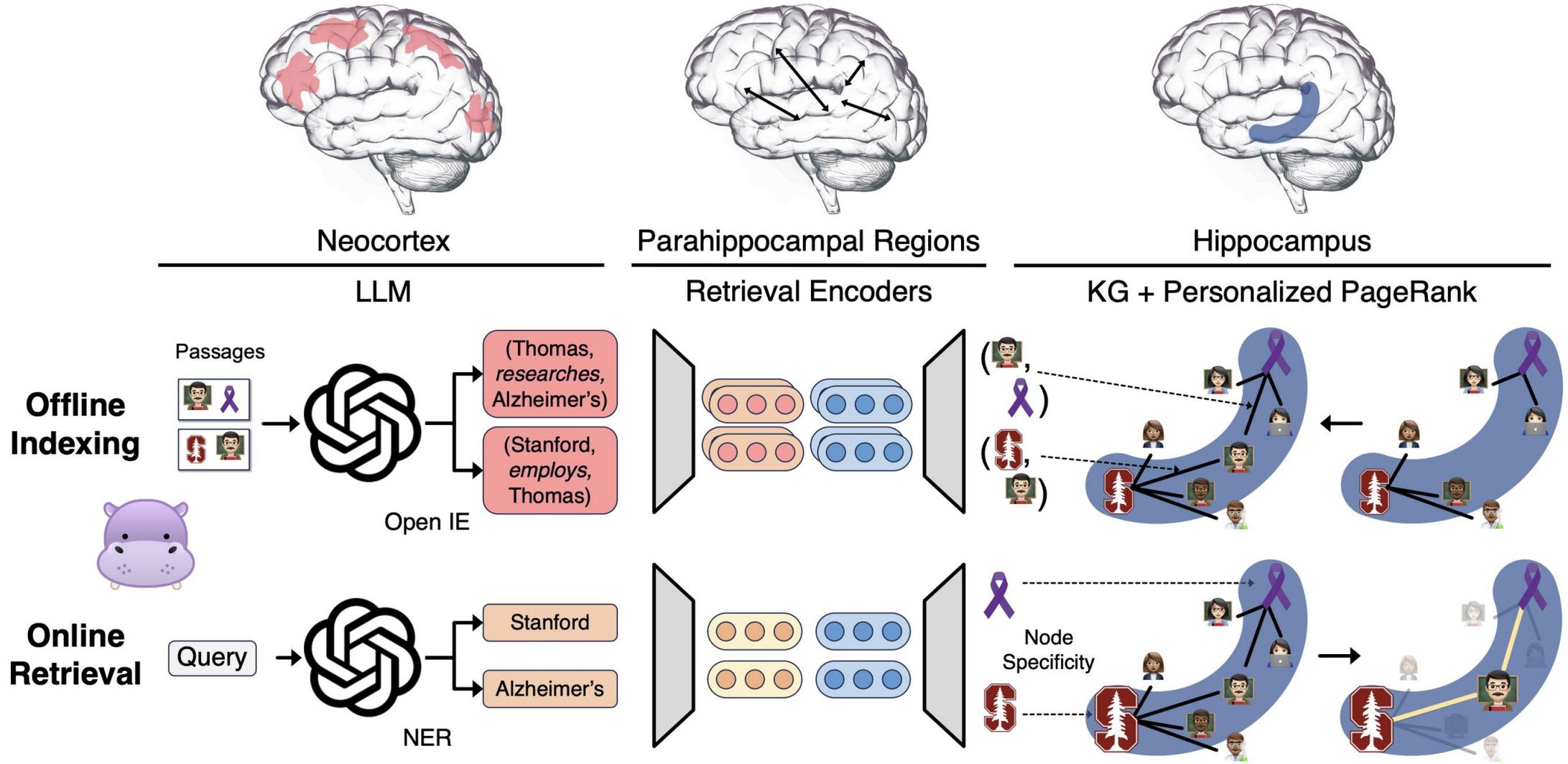
Hippocampus

Indexing & auto-associative memory

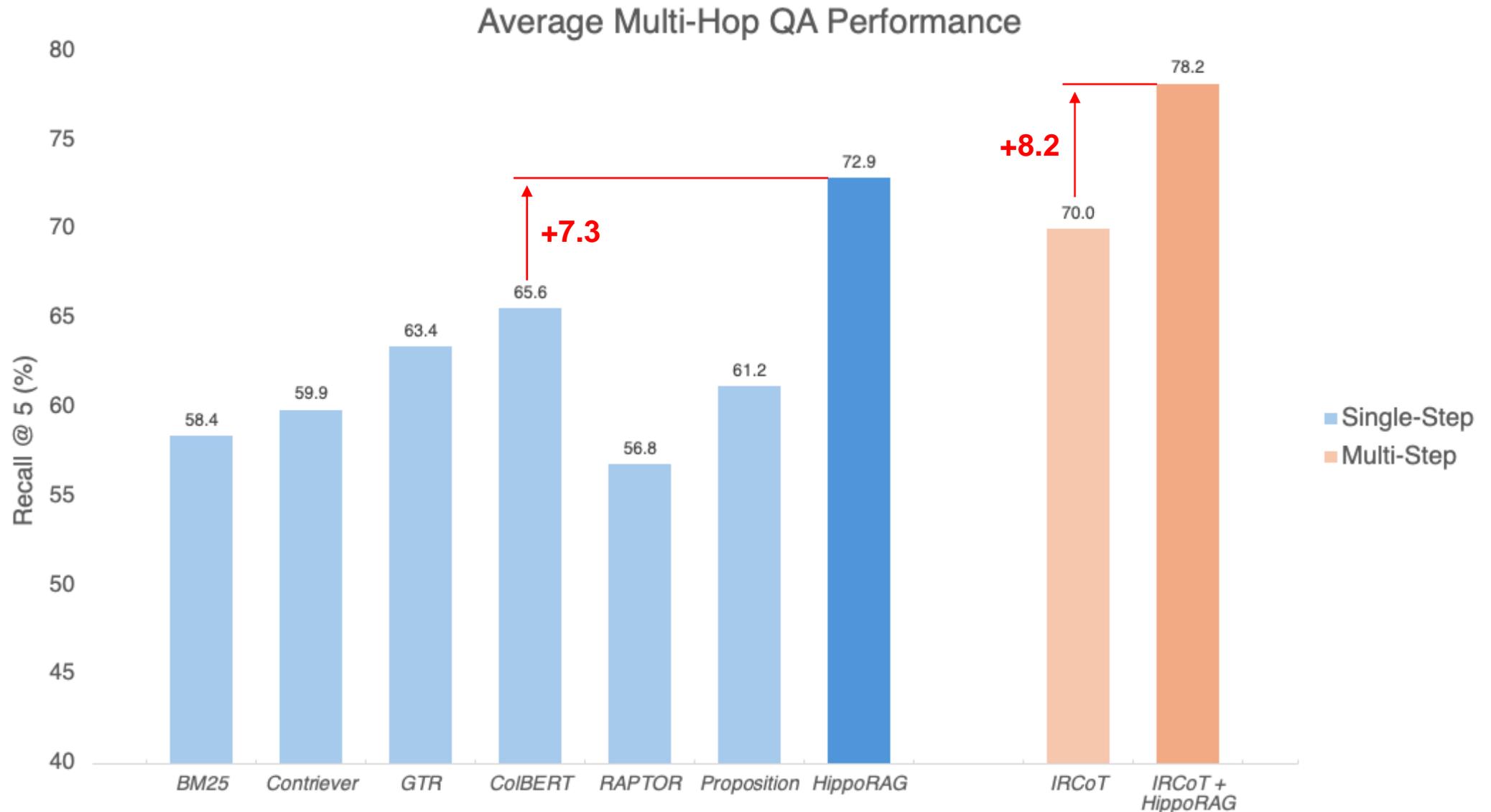
HippoRAG & the three components of LTM



HippoRAG & the three components of LTM



HippoRAG is a SoTA memory retriever



Case study: path-finding questions

	Question	HippoRAG	ColBERTv2	IRCoT
Path-Following	In which district was Alhandra born?	<ol style="list-style-type: none">1. Alhandra2. Vila de Xira3. Portugal	<ol style="list-style-type: none">1. Alhandra2. Dimuthu Abayakoon3. Ja'ar	<ol style="list-style-type: none">1. Alhandra2. Vila de Xira3. Póvoa de Santa Iria
Path-Finding	Which Stanford professor works on the neuroscience of Alzheimer's ?	<ol style="list-style-type: none">1. Thomas Südhof2. Karl Deisseroth3. Robert Sapolsky	<ol style="list-style-type: none">1. Brian Knutson2. Eric Knudsen3. Lisa Giocomo	<ol style="list-style-type: none">1. Brian Knutson2. Eric Knudsen3. Lisa Giocomo

HippoRAG v2 is coming soon!

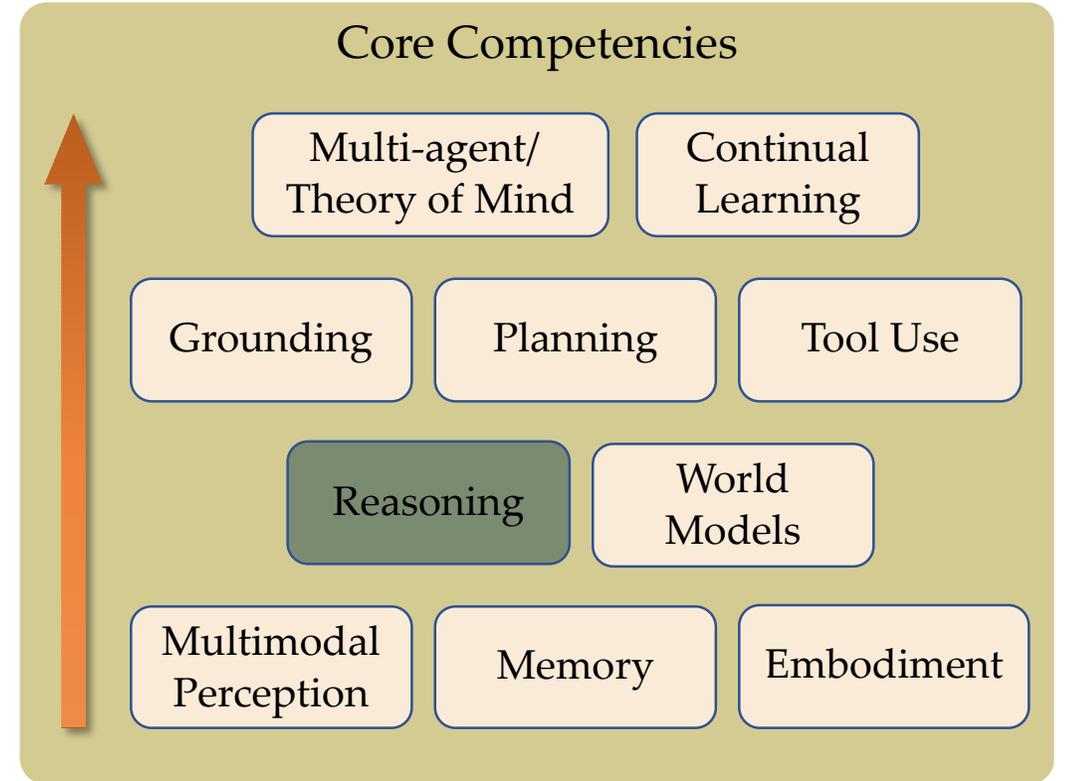
Retrieval	Simple QA		Multi-Hop QA				Discourse Understanding	Avg
	NQ	PopQA	MuSiQue	2Wiki	HotpotQA	LV-Eval	NarrativeQA	
<i>Simple Baselines</i>								
None	54.9	32.5	26.1	42.8	47.3	6.0	12.9	38.4
BM25 (Robertson & Walker, 1994)	59.0	49.9	28.8	51.2	63.4	5.9	18.3	47.7
Contriever (Izacard et al., 2022)	58.9	53.1	45.4	41.9	62.3	8.1	19.7	49.5
GTR (T5-base) (Ni et al., 2022)	59.9	56.2	34.6	52.8	62.8	7.1	19.9	50.4
<i>Large Embedding Models</i>								
GTE-Qwen2-7B-Instruct (Li et al., 2023)	62.0	56.3	40.9	60.0	71.0	7.1	21.3	54.9
GritLM-7B (Muennighoff et al., 2024)	61.3	55.8	44.8	60.6	73.3	9.8	23.9	56.1
NV-Embed-v2 (7B) (Lee et al., 2025)	61.9	55.7	45.7	61.5	75.3	9.8	25.7	57.0
<i>Structure-Augmented RAG</i>								
RAPTOR (Sarathi et al., 2024)	50.7	56.2	28.9	52.1	69.5	5.0	21.4	48.8
GraphRAG (Edge et al., 2024)	46.9	48.1	38.5	58.6	68.6	11.2	23.0	49.6
LightRAG (Guo et al., 2024)	16.6	2.4	1.6	11.6	2.4	1.0	3.7	6.6
HippoRAG (Gutiérrez et al., 2024)	55.3	55.9	35.1	71.8	63.5	8.4	16.3	53.1
HippoRAG v2	63.3	56.2	48.6	71.0	75.5	12.9	25.9	59.8

Memory: takeaways

- Memory is central to human learning. Our sophisticated memory mechanisms allow us to recognize patterns, create associations, and dynamically recall memory relevant to the current context beyond superficial similarity
- Long-term memory through parametric continual learning is hard for LLMs, but non-parametric memory (e.g., RAG) could be a promising solution
- Recent trend in RAG is to add more structures to embeddings (e.g., HippoRAG, GraphRAG) to enhance
 - *Sensemaking*, the ability to interpret larger, more complex, or uncertain contexts
 - *Associativity*, the capacity to draw multi-hop connections between disparate pieces of info

Outline

- On long-term memory
 - HippoRAG
- On reasoning
 - Grokked Transformers
- On world models and planning
 - WebDreamer



Grokking of Implicit Relations in Transformers: A Mechanistic Journey to the Edge of Generalization

Boshi Wang, Xiang Yue, Yu Su, Huan Sun

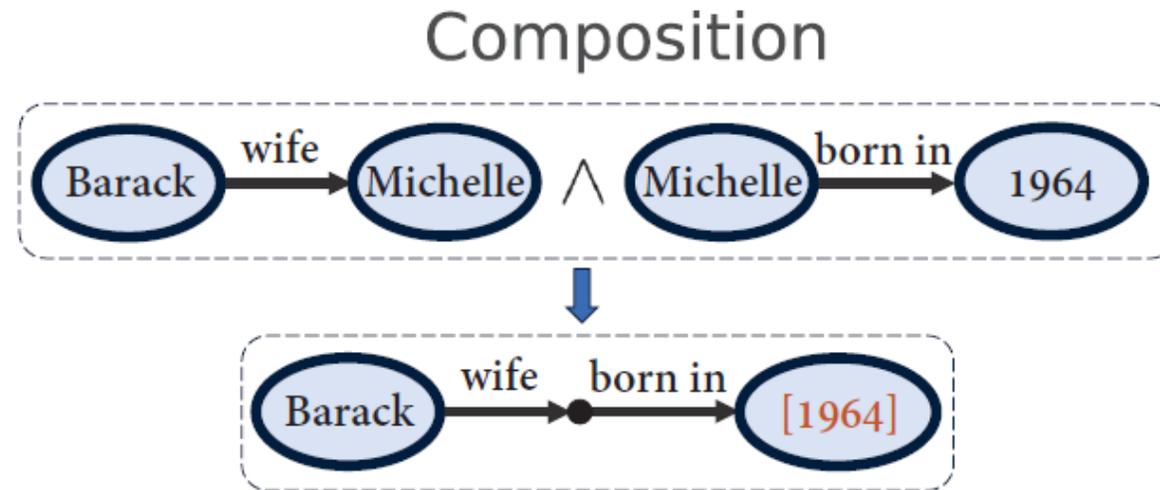


**Carnegie
Mellon
University**



Implicit reasoning

Directly predict the answer; no verbalized chain of thought



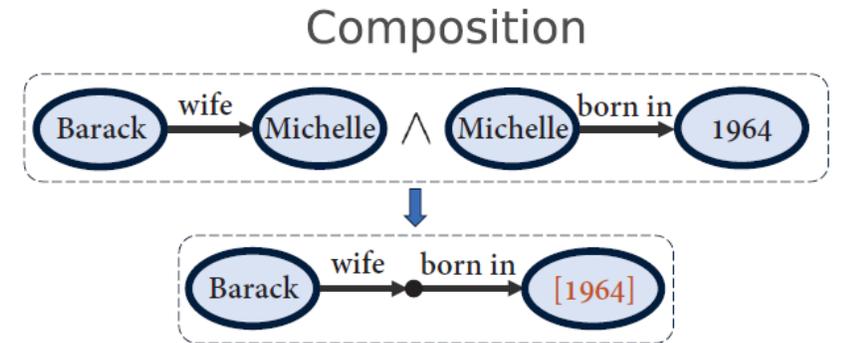
CoT is all the rage. Why does implicit reasoning matter?

- The default mode of large-scale (pre-)training; no CoT at training time
- Fundamentally determines how well LLMs acquire structured representations of facts and rules from data
- How did o1/R1-style long CoT emerge? A hypothesis
 - A capable base model (e.g., DeepSeek v3) has already learned various basic ‘constructs’ or strategies for reasoning
 - Reinforcement learning incentivizes the model to learn to use the right combination of strategies (not learning new ones through RL) and keep trying

LLMs were shown to struggle at implicit reasoning

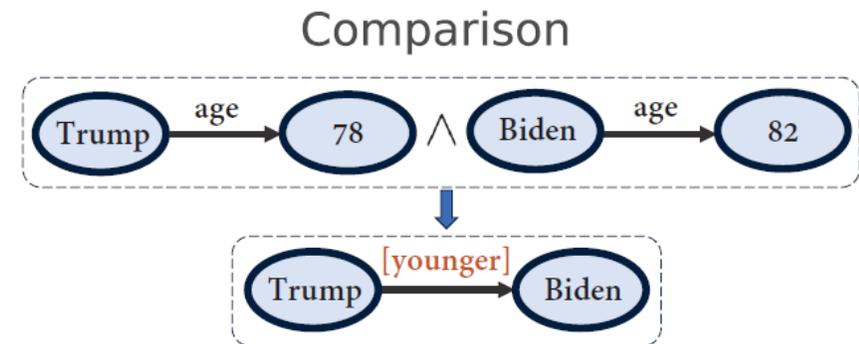
- Composition

- LLMs only show substantial evidence in first-hop reasoning (Yang et al. 2024)
- “Compositionality gap” does not decrease with scale (Press et al. 2023)



- Comparison

- GPT-4 struggles at implicitly comparing entity attributes despite knowing them perfectly (Zhu et al. 2023)



Research questions

- Can Transformers learn to reason implicitly, or are there fundamental limitations that prohibit robust acquisition of this skill?
- What factors (e.g., data scale, distribution, model architecture) control the acquisition of implicit reasoning?

Setup: model and optimization

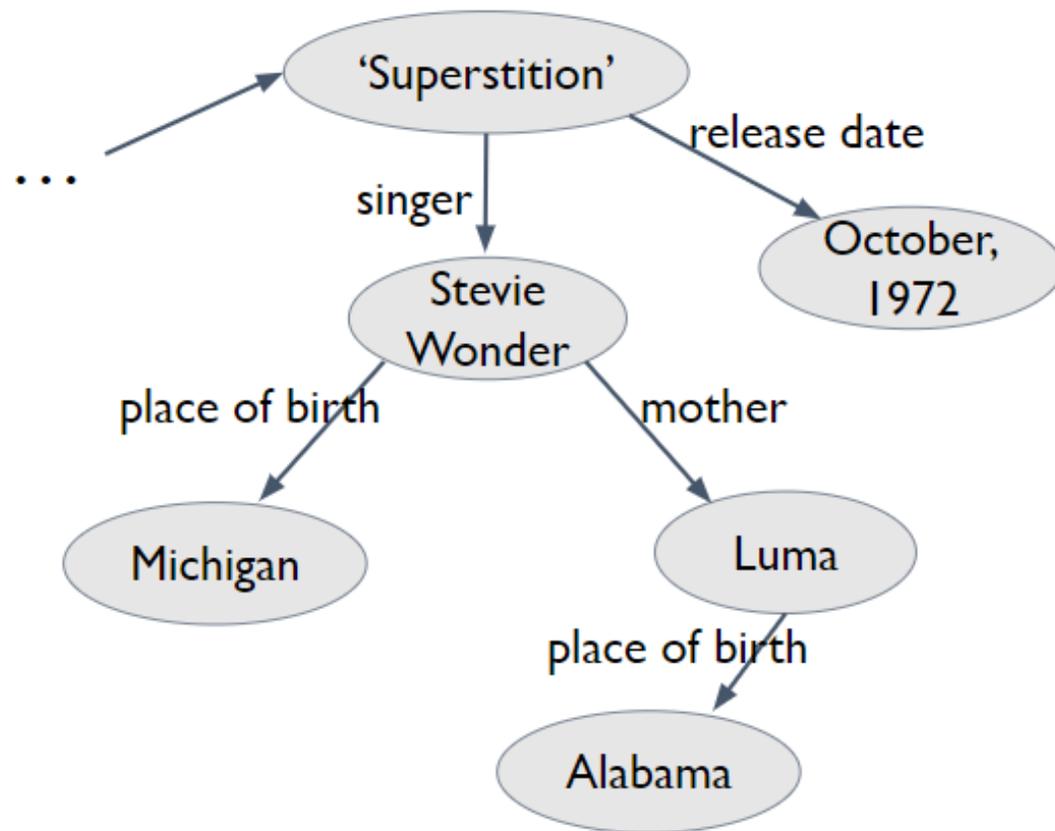
- Standard decoder-only transformer as in GPT-2
 - 8 layers, 768 hidden dimensions, and 12 attention heads
 - Results are robust to different model scales
- AdamW with learning rate $1e-4$, batch size 512, weight decay 0.1, and 2000 warm-up steps

Setup: data for compositional reasoning

- Atomic facts
 - Random KG consisting of $|\mathcal{E}|$ entities and $|\mathcal{R}| = 200$ relations
 - Randomly split into ID & OOD atomic facts

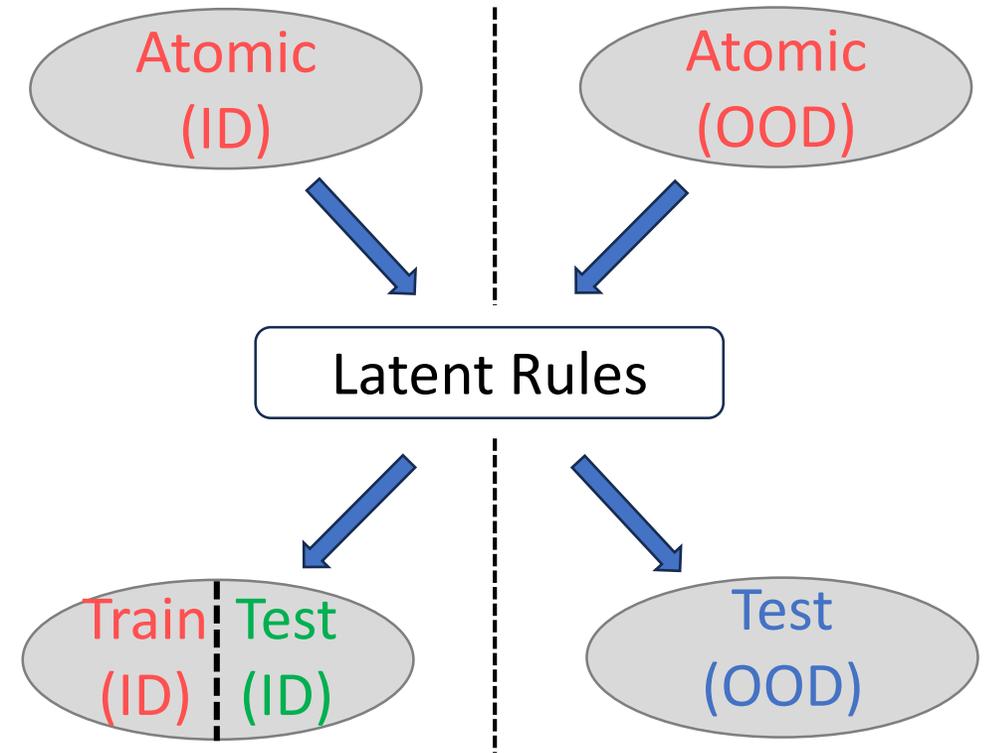
- Inferred facts: two-hop compositions

$$(h, r_1, b) \wedge (b, r_2, t) \implies (h, r_1, r_2, t)$$

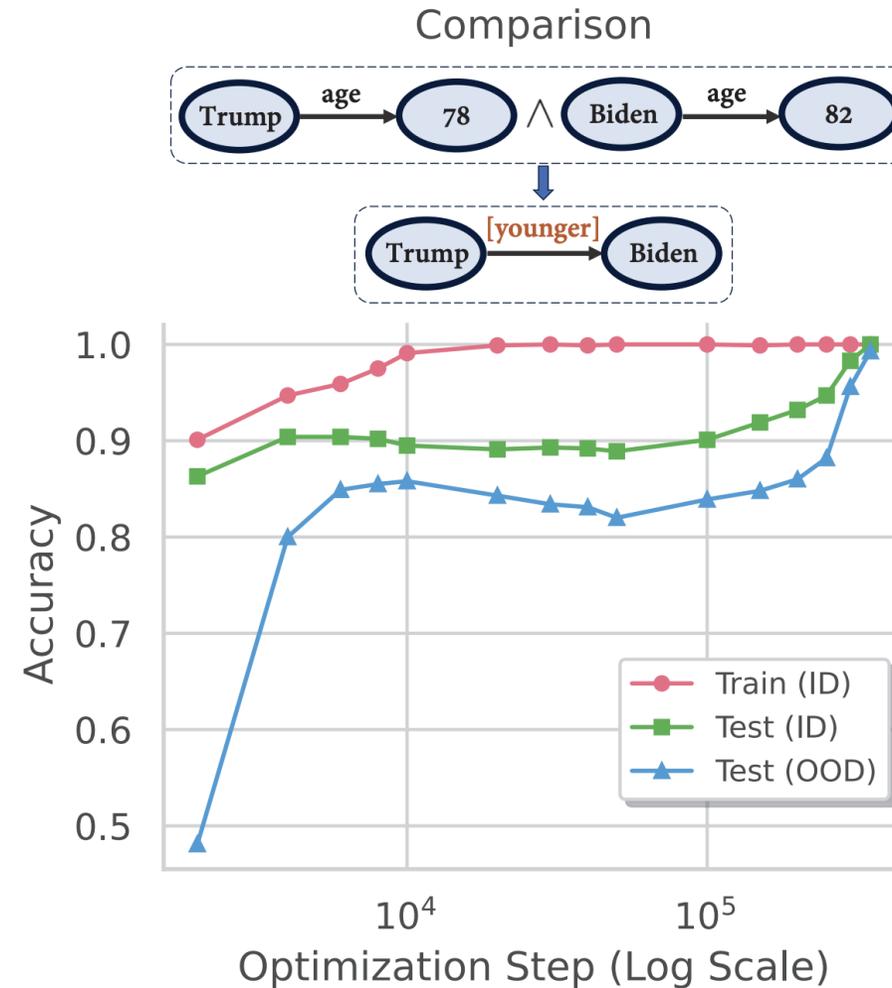
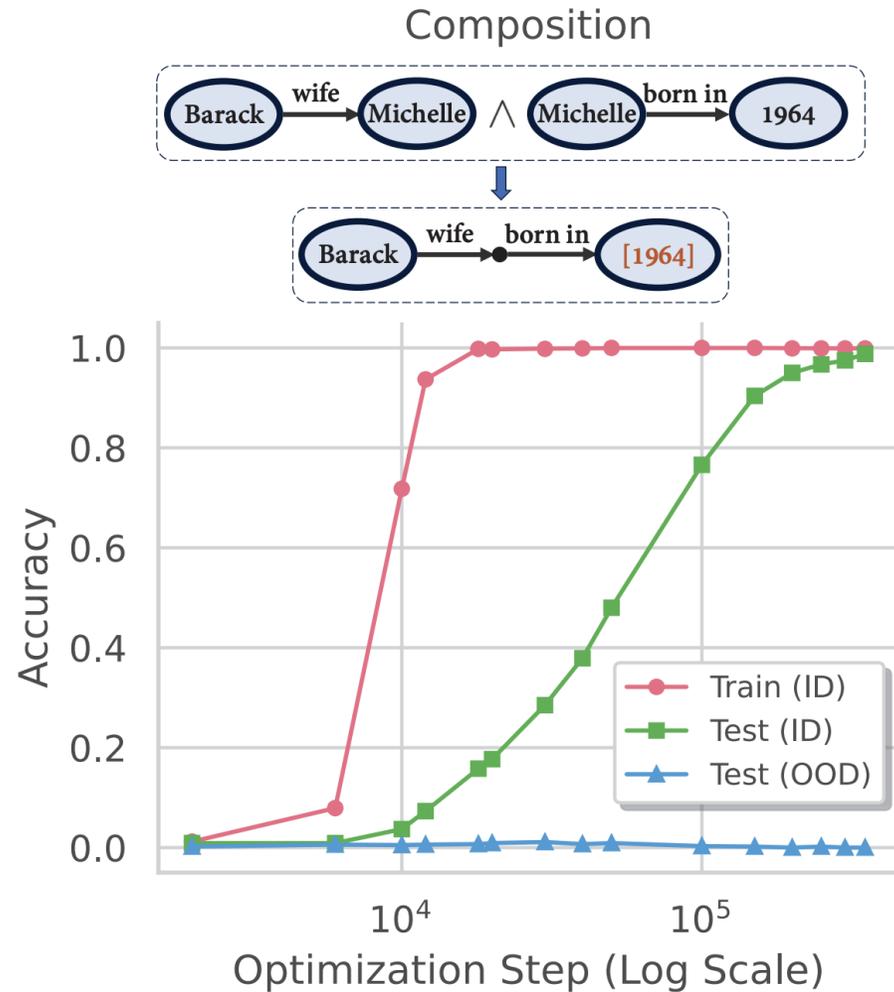


Setup: inductive learning of deduction rules

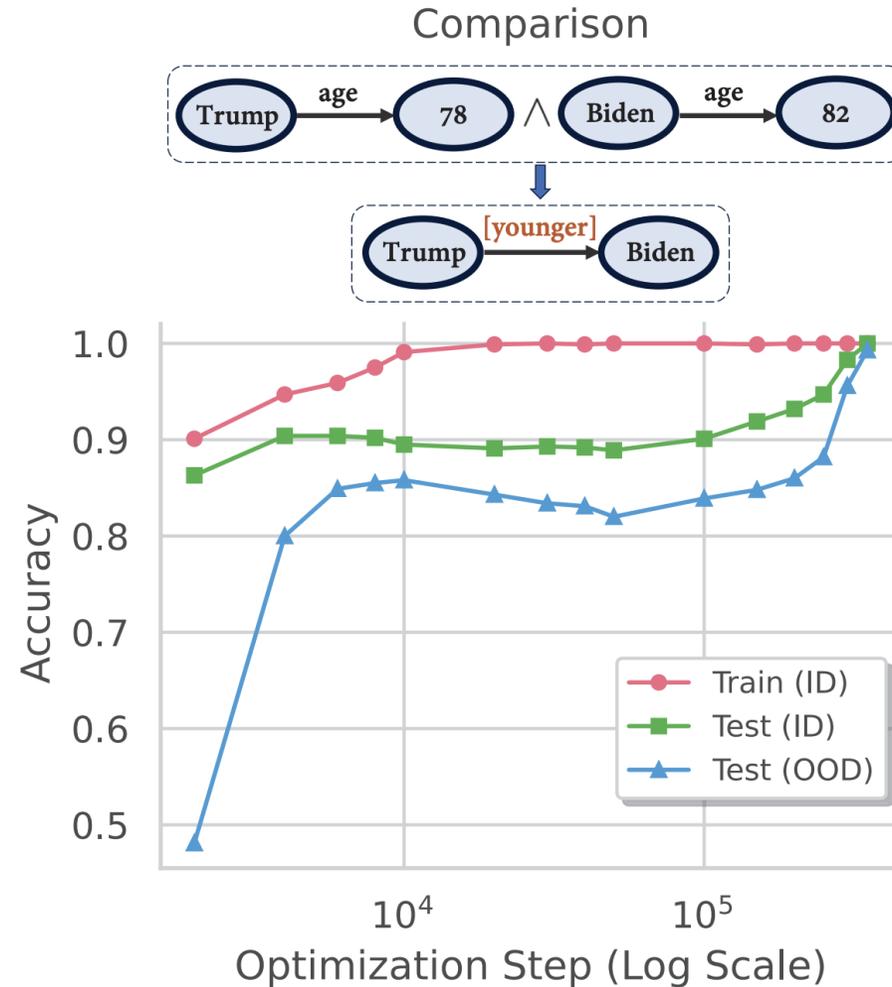
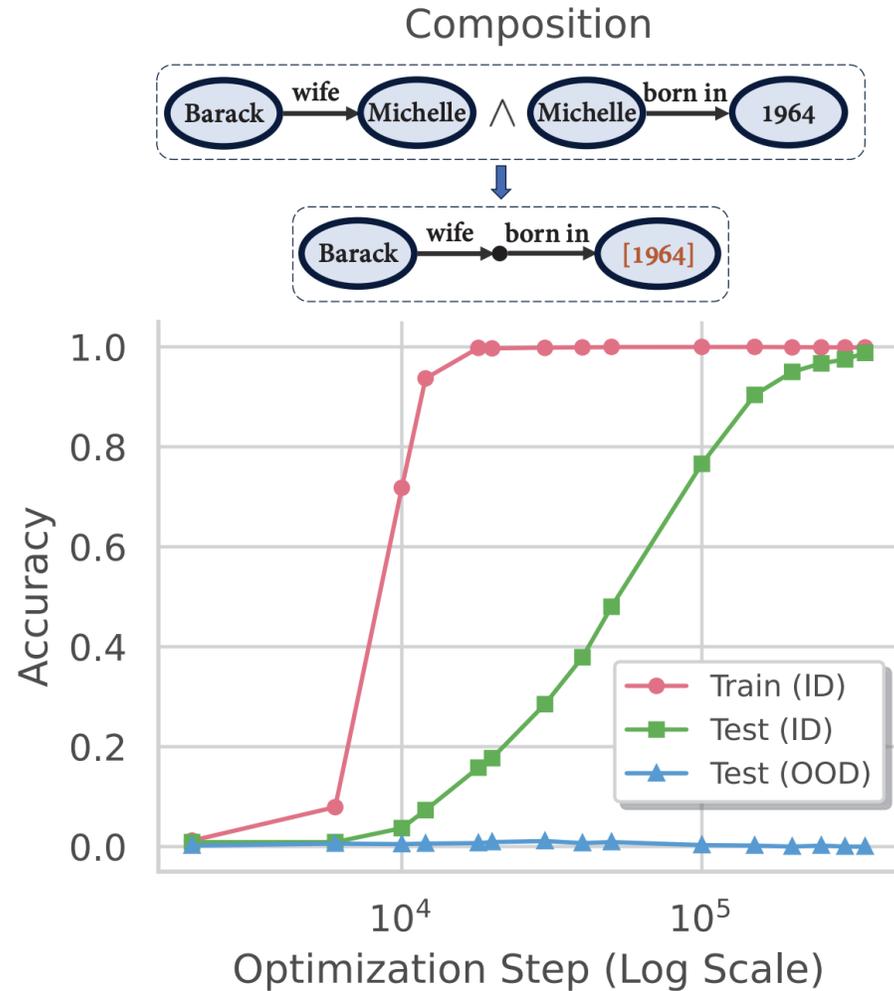
- **Induce** latent rules from a mixture of **atomic** facts and **inferred** facts (deduced via latent rules)
- **Deduce** novel facts by applying the acquired rules
 - **Test (ID)**: unseen inferred facts deduced from the **same** set of atomic facts underlying the observed inferred facts
 - **Test (OOD)/Systematic Generalization**: unseen inferred facts derived from a **different** set of atomic facts



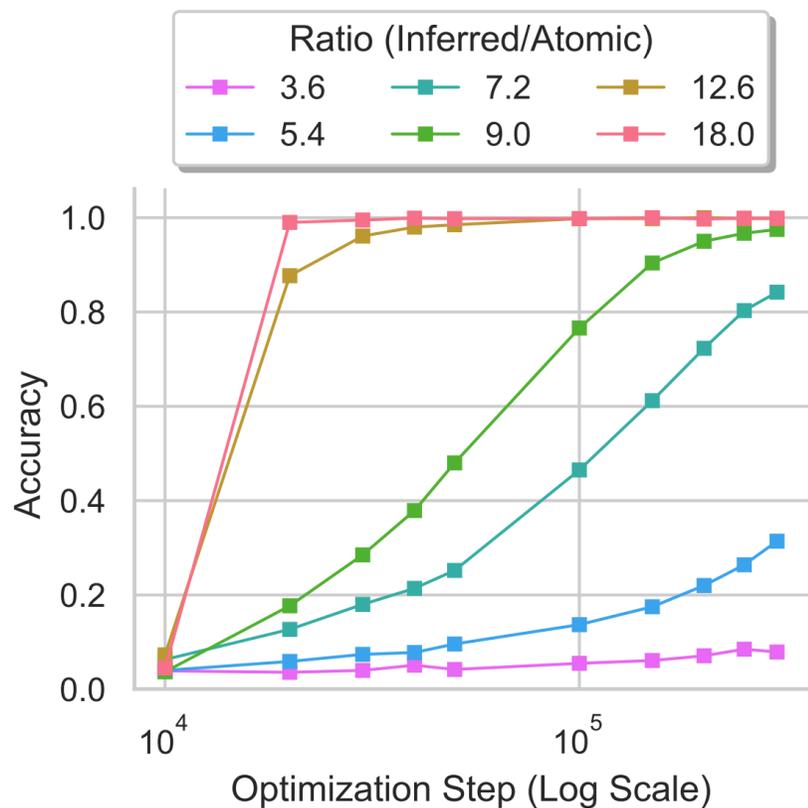
Takeaway #1: Transformers can learn to reason implicitly, but only through 'grokking'



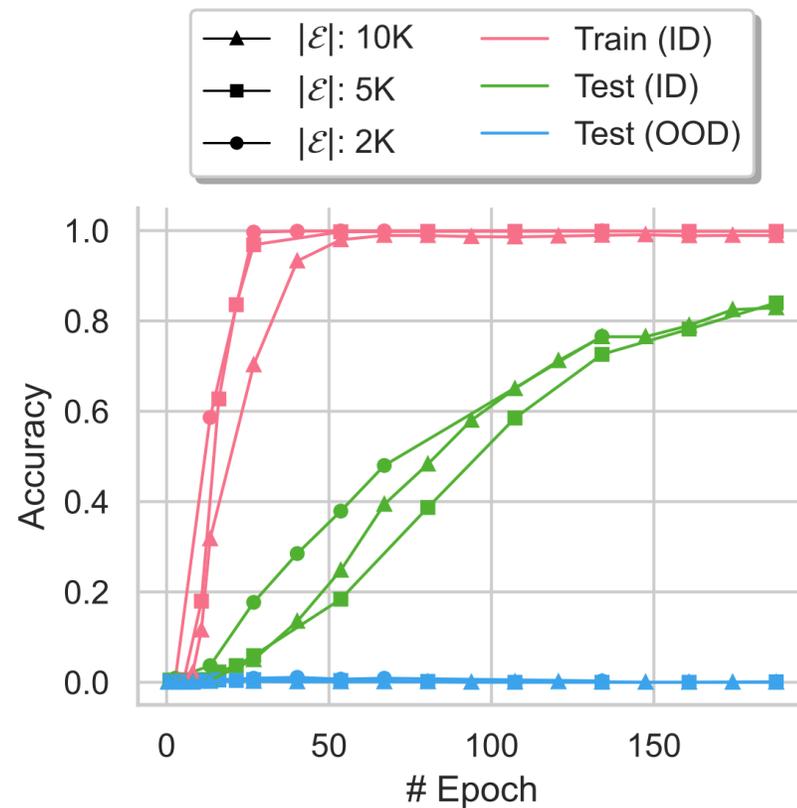
Takeaway #2: Systematicity varies by reasoning type



Takeaway #3: Critical data distribution, not size



(a) Effect of changing ratio ϕ ($|\mathcal{E}| = 2000$).



(b) Effect of changing $|\mathcal{E}|$ ($\phi = 9.0$).

$|\mathcal{E}|$: # of entities (proportional to total data size). ϕ : ratio of inferred/atomic

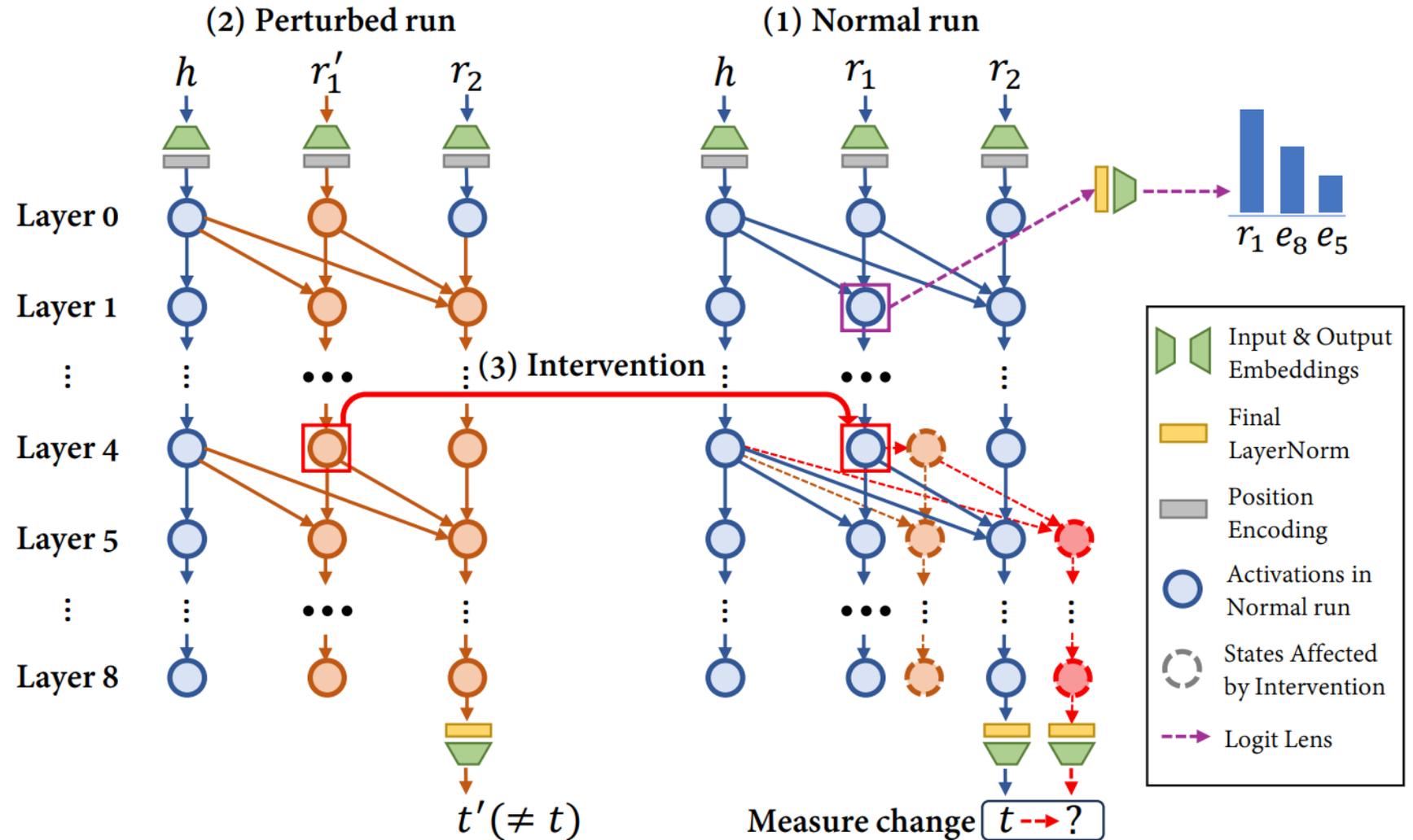
Important questions remain

- Why does grokking happen?
- What happens during grokking?
- Why does the level of systematicity in generalization vary?

These require a deeper look inside the model

Analyzing the changes during grokking

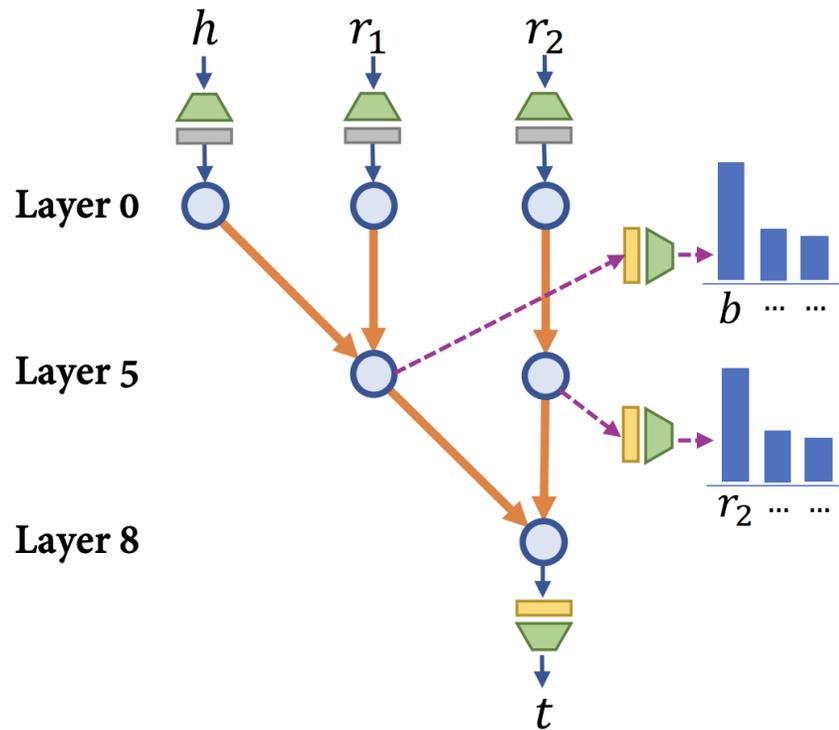
- Logit lens
- Causal tracing



Generalizing circuit configuration determines systematicity of generalization

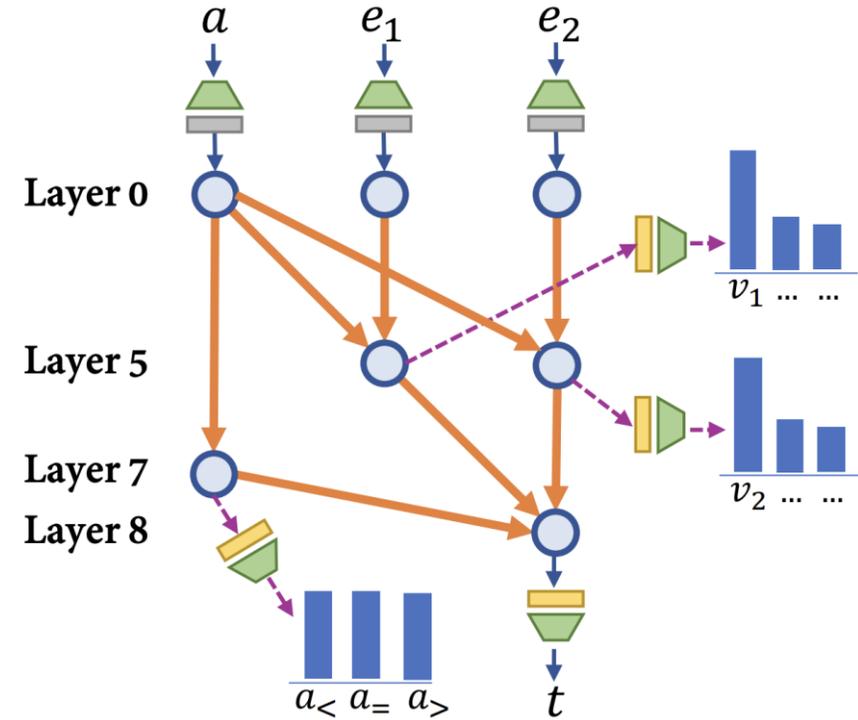
Composition

A "staged" circuit

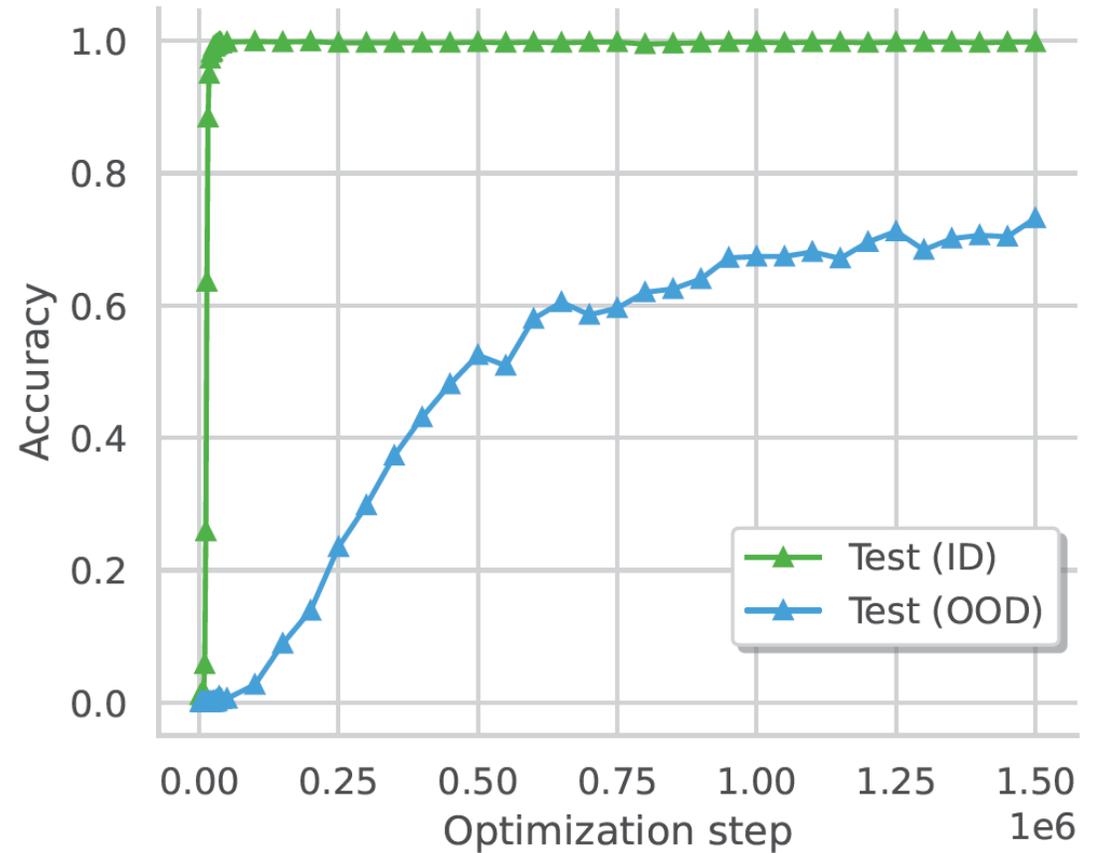
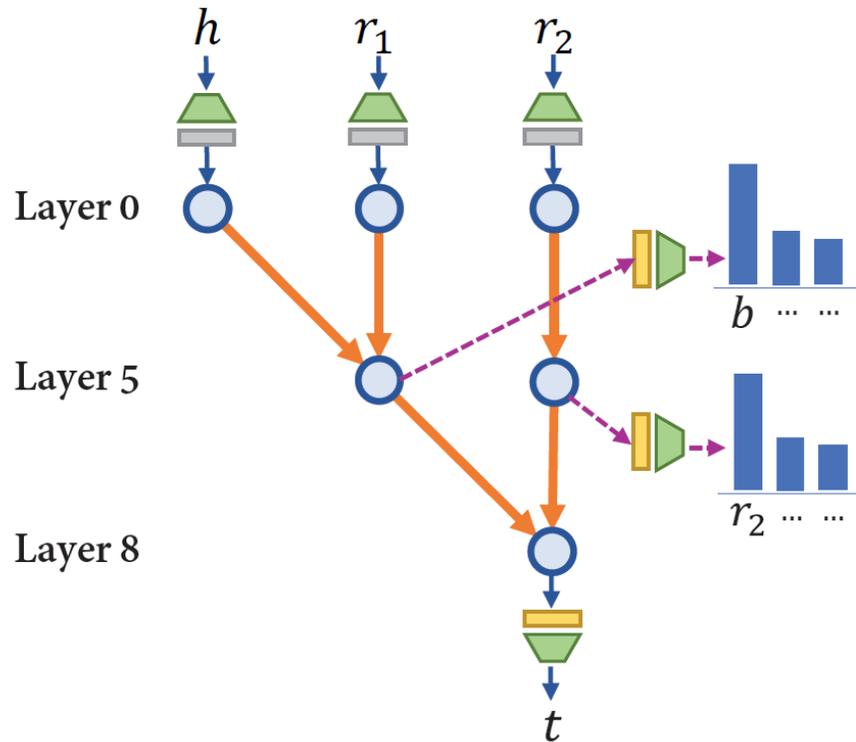


Comparison

A "parallel" circuit

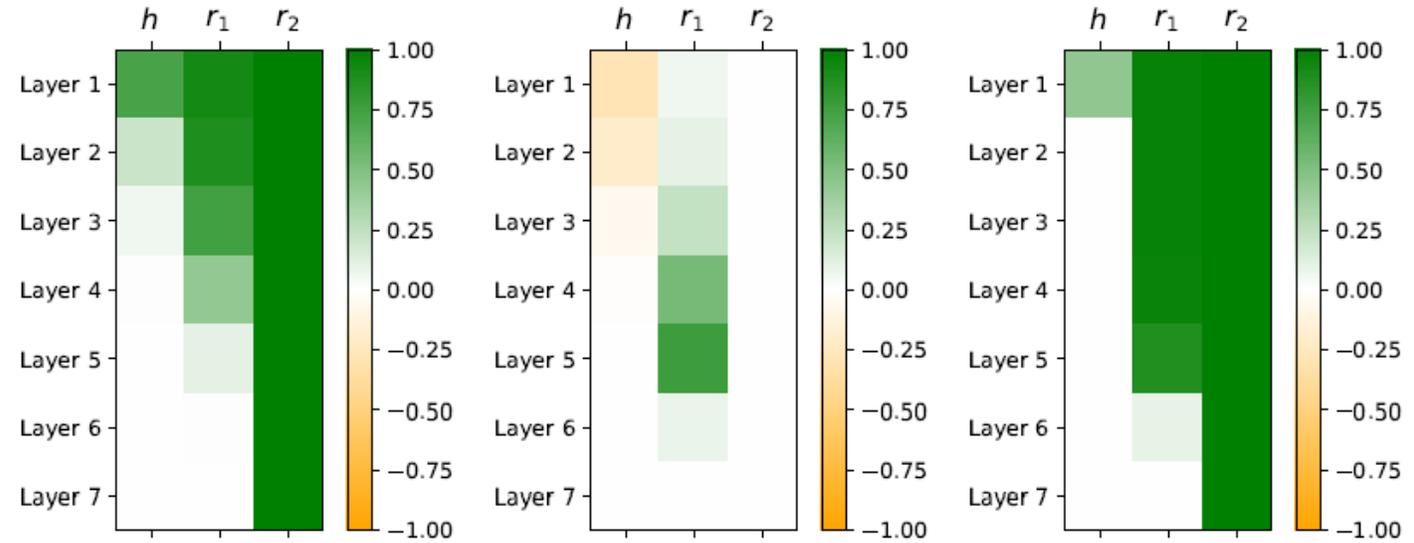


Improving systematic generalization via cross-layer parameter sharing

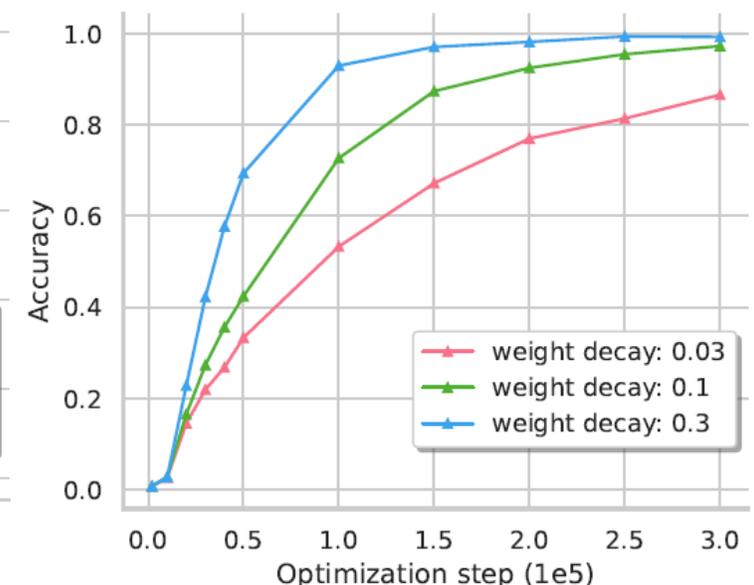
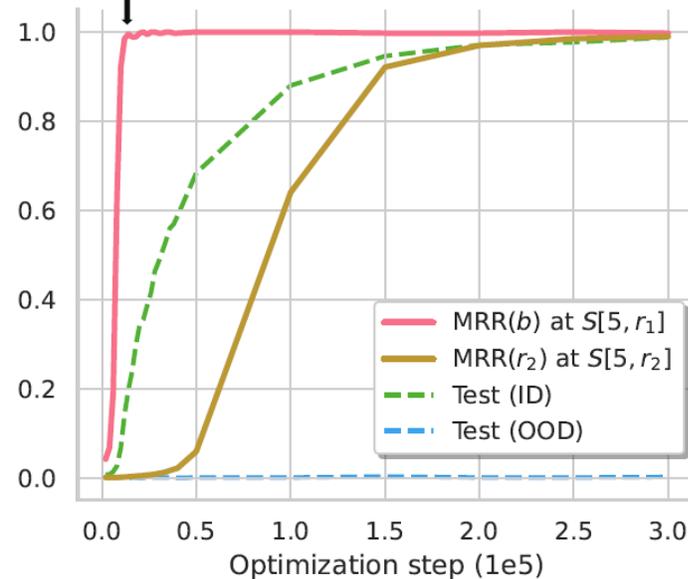


Grokking is the phase transition from rote learning to generalization (the 'aha moment')

- Grokking is when the generalizing circuit forms and outcompetes the memorizing circuit
- Explanation via circuit efficiency and regularization
 - # of facts need to be stored by memorizing vs. generalizing circuits

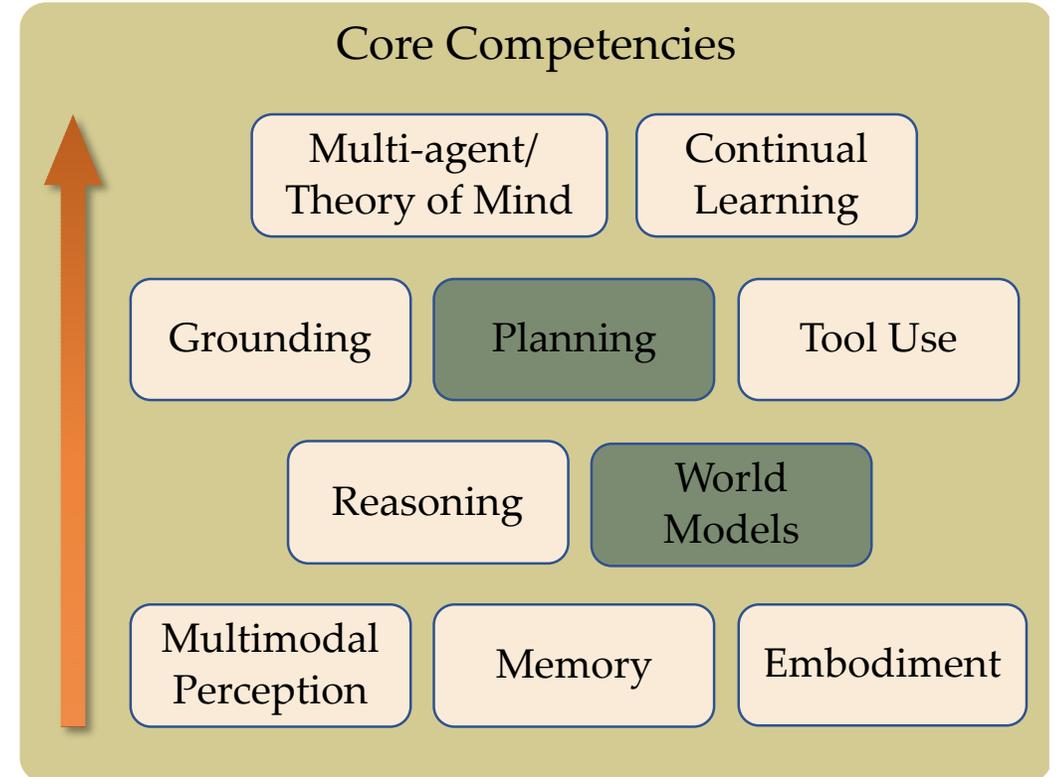


Grokking starts



Outline

- On long-term memory
 - HippoRAG
- On reasoning
 - Grokked Transformers
- On world models and planning
 - WebDreamer



Planning: simplified definition

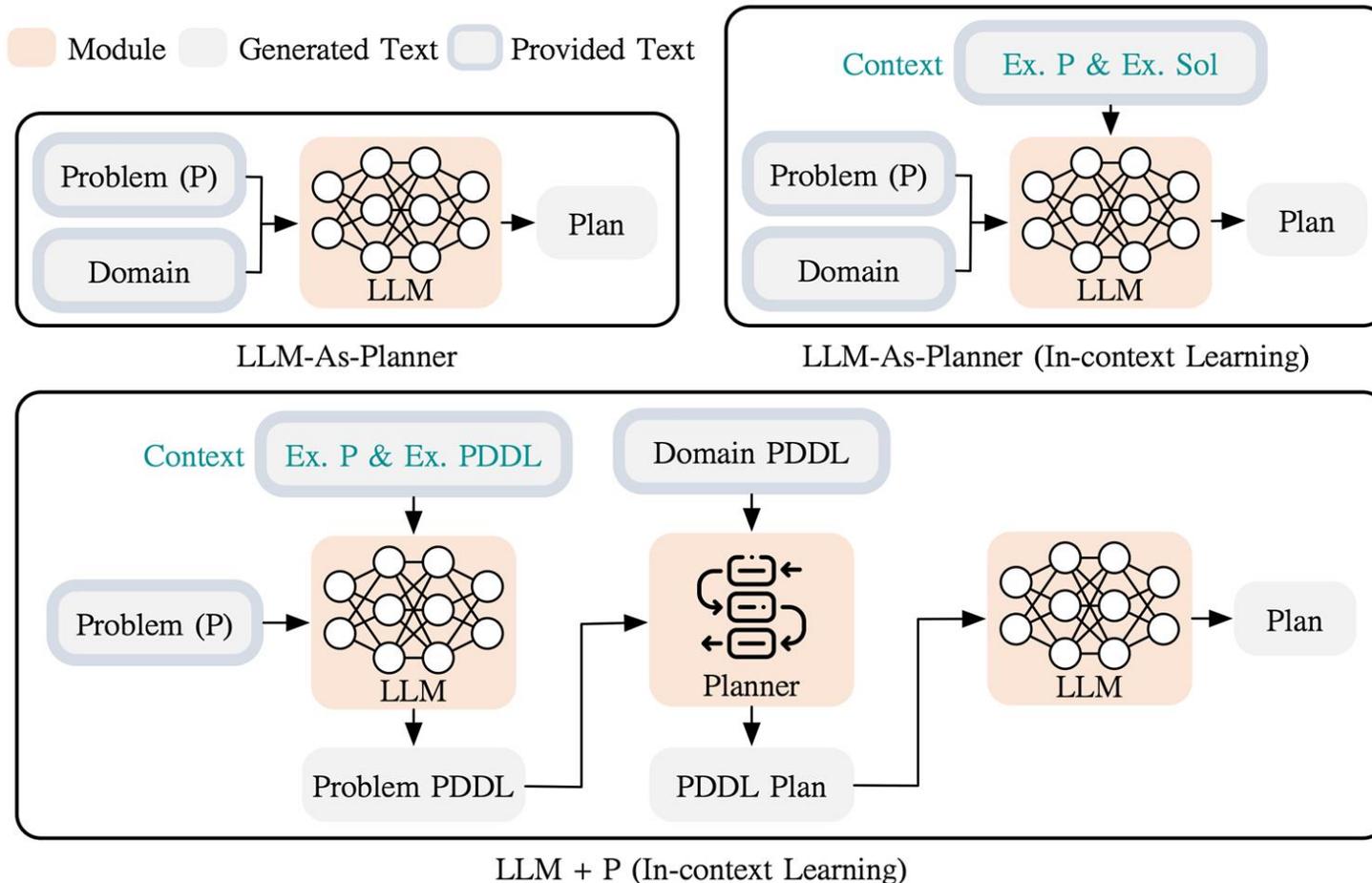
Given a goal G , decides on a sequence of actions (a_0, a_1, \dots, a_n) that will lead to a state that passes the goal test $g(\cdot)$

General trends in planning settings for language agents

- Increasing expressiveness in **goal specification**, e.g., in natural language as opposed to formal language
- Substantially expanded or open-ended **action space**
- Increasing difficulty in automated **goal test**

LLM+P

When domain and problem can be (fully or partially) described formally, e.g., in PDDL (Planning Domain Definition Language)



An Example PDDL Problem File Written by GPT-4 with Context

Context (shaded):

```
An example planning problem is: You have 5 blocks.
b2 is on top of b5. b5 is on top of b1. b1 is on top of
b4. b3 is on top of b2. b4 is on the table. b3 is clear.
Your arm is empty. Your goal is to move the blocks.
b4 should be on top of b3. The problem PDDL is:
(:objects b1 b2 b3 b4 b5)
(:init (arm-empty) (on b1 b4) (on
b2 b5) (on b3 b2)
      (on-table b4) (on b5 b1)
      (clear b3))
(:goal (and (on b4 b3))))
```

Prompt: Description of Problem (P1) + Context + Provide me with the problem PDDL file that describes the planning problem directly without further explanations.

GPT-4 (the generated problem PDDL):

```
(:objects b1 b2 b3 b4 b5 )
(:init (arm-empty) (on-table b1)
      (on b2 b1) (on b3 b4)
      (on b4 b2) (on b5 b3) (clear
b5)))
(:goal (and (on b1 b2) (on b3 b5)
      (on b4 b1)))
```

Planner:

```
(unstack b5 b3) (putdown b5)
(unstack b3 b4) (stack b3 b5)
(unstack b4 b2) (putdown b4)
(unstack b2 b1) (putdown b2)
(pickup b1) (stack b1 b2) (pickup
b4) (stack b4 b1)
```

Language agent planning: web agents

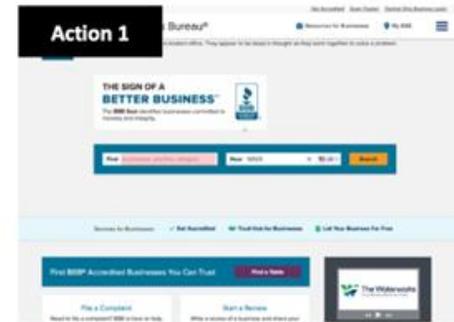
Task Description:

Show me the reviews for the auto repair business closest to 10002.

Action Sequence:

Target Element	Operation
1. [searchbox] Find	TYPE: <i>auto repair</i>
2. [button] Auto Repair	CLICK
3. [textbox] Near	TYPE: <i>10002</i>
4. [button] 10002	CLICK
5. [button] Search	CLICK
6. [switch] Show BBB Accredited only	CLICK
7. [svg]	CLICK
8. [button] Sort By	CLICK
9. [link] Fast Lane 24 Hour Auto Repair	CLICK
10. [link] Read Reviews	CLICK

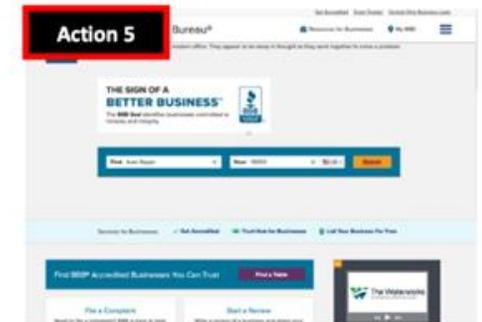
Webpage Snapshots:



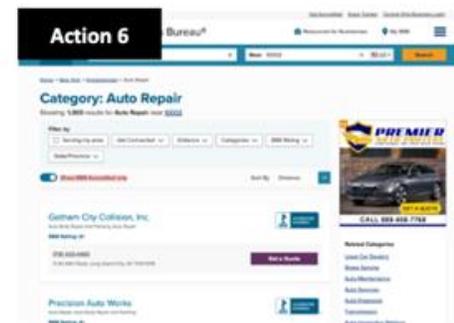
```
<input name="find_text" type="search">
```



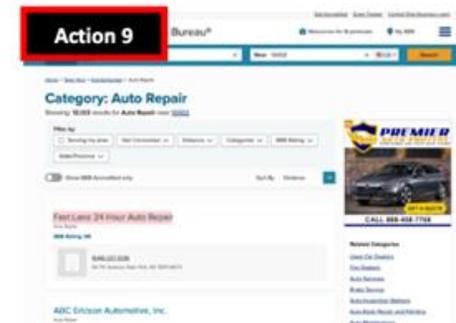
```
<em>Auto Repair</em>
```



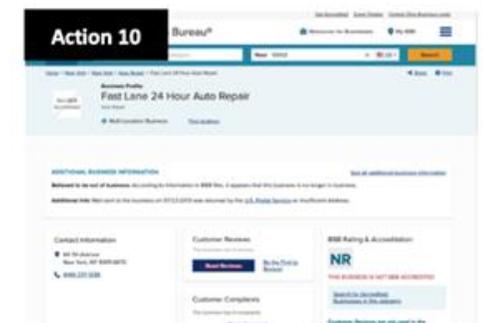
```
<button>Search</button>
```



```
<button>Show BBB Accredited only</button>
```

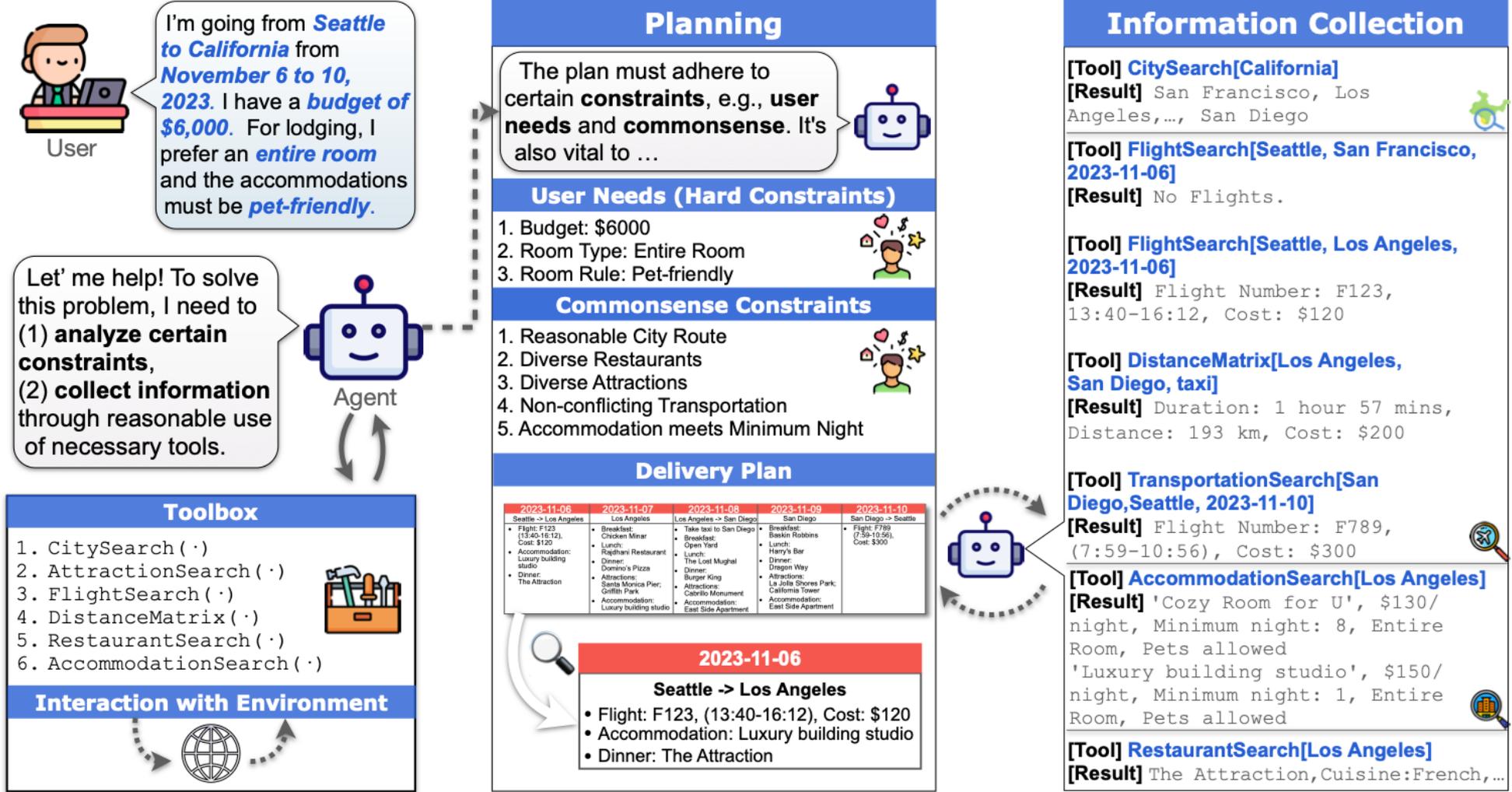


```
<span>Fast Lane 24 Hour Auto Repair</span>
```



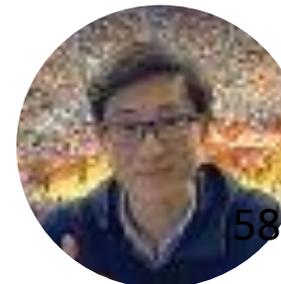
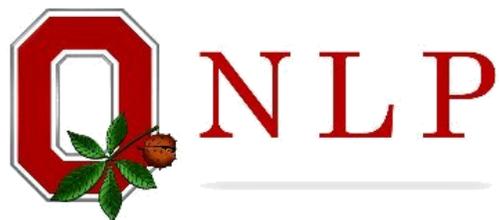
```
<a href="link:XXX">Read Reviews</a>
```

Language agent planning: travel planning

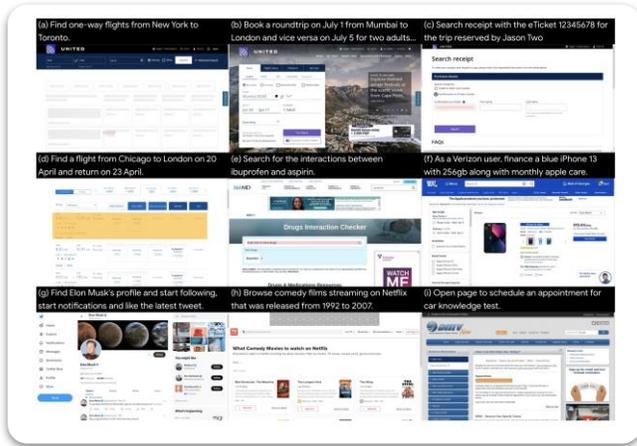


Is Your LLM Secretly a World Model of the Internet? Model-based Planning for Web Agents

Yu Gu*, **Boyuan Zheng***, Boyu Gou, Kai Zhang,
Cheng Chang, Sanjari Srivastava, Yanan Xie, Peng Qi,
Huan Sun, Yu Su



Evolution of web (or computer use) agents



Mind2Web [NeurIPS'23]

First LLM-based web agent
Ecologically valid eval

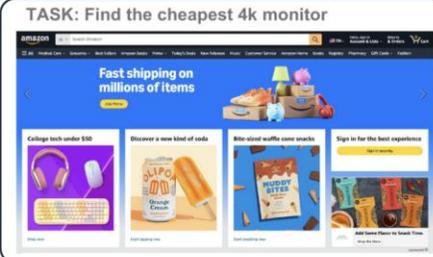
First generalist web agent
with visual perception

SeeAct [ICML'24]



Vision-Only Observation

TASK: Find the cheapest 4k monitor



Planning

User: Decide the next action for the task.
Element Description: The search bar at the top of the page.
Action: Type
Value: 4k monitor

Grounding

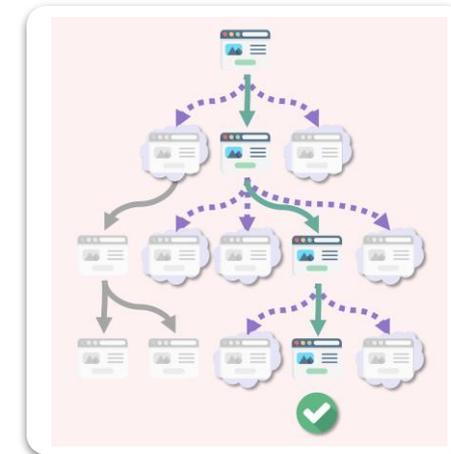
User: What are the pixel coordinates of the element corresponding to "The search bar at the top of the page" ?
(556, 26)

UGround [ICLR'25]

Pure vision-based agent
Human-like embodiment

Model-based planning

WebDremer

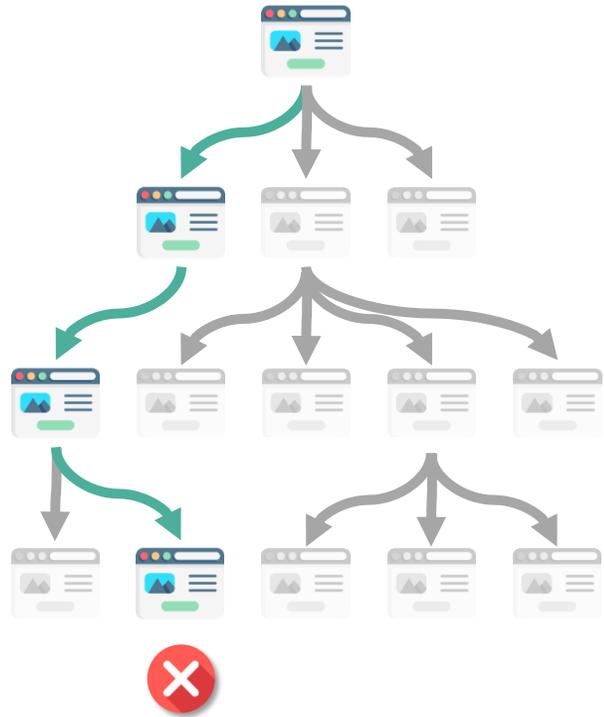


Agentic search
Continual learning
Safety



Planning paradigms for language agents

(a) reactive

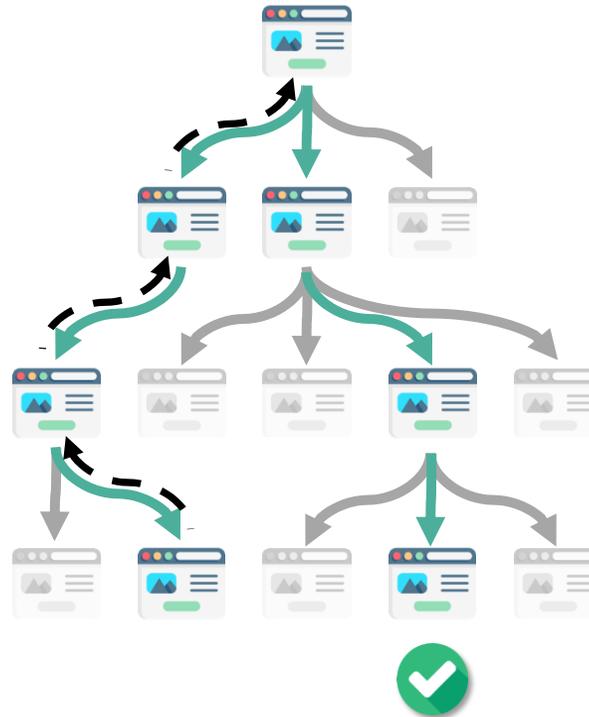


fast, easy to implement



greedy, short-sighted

(b) tree search with real interactions



systematic exploration



irreversible actions,
unsafe, slow

Challenges with search in real-world environments

- Many actions are state-changing and irreversible → backtracking ✘
- Safety/privacy risks
- Inference-time exploration could be slow and costly

Cancel Place Your Order - Amazon.co...



Ships from and sold by Amazon.com

Quantity: 1 Change
Add gift options
Auto-deliver and save up to 5% on future auto-deliveries

Item often ships in manufacturer's container to reduce packaging and reveals what's inside. To change, click below.

Reduce packaging, ship in manufacturer's container

Place your order

By placing your order, you agree to Amazon's [privacy notice](#) and [conditions of use](#).

Search or ask a question

Kohl's Dropoff **FREE**

Kohl's will pack, label, and ship your return for free. Just bring the item in its original manufacturer's packaging and disassemble the item (if applicable). We'll email you a QR code to ship your return. Show it to a store associate at any Kohl's store.

[Find a participating Kohl's store](#)
Printer not required.

The UPS Store locations only—no label needed **\$6.99**

Amazon Dropoff — box and label needed **FREE**

2 OTHER RETURN OPTIONS

Refund summary **\$13.21**

Confirm your return

Verify mobile number

A text with a One Time Password (OTP) has been sent to your mobile number: **8058671234** [Change](#)

Enter OTP: [Resend code](#)

Create your Amazon account

By creating an account, you agree to Amazon's [Conditions of Use](#) and [Privacy Notice](#).

Search or ask a question

Location

Disabled

AddressBook/Checkout
Your current location will be used to assist in adding a new address to your Amazon address book.

Amazon Cash
We use your location to find nearby stores where you can add money to your Amazon balance with Amazon Cash.

Branded Store Experience Location
We use your location to power branded store experiences.

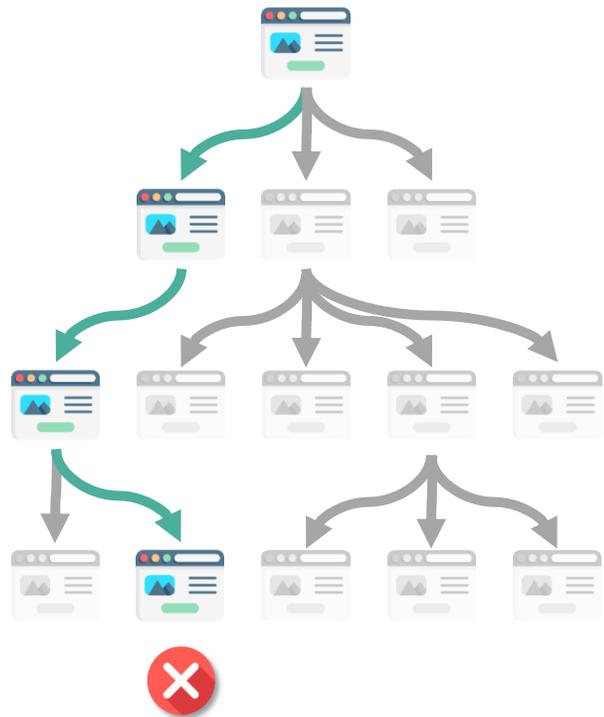
Branded Store Experience Location-based Augmented Reality
We use your camera, motion, and location to power branded store experiences. Requires camera, motion, and location.

Campus pickup
We'll use your location to show the nearby pickup points

Delivery Location
We use your location to improve your shopping experience, ensuring you only see products and delivery options available in your area.

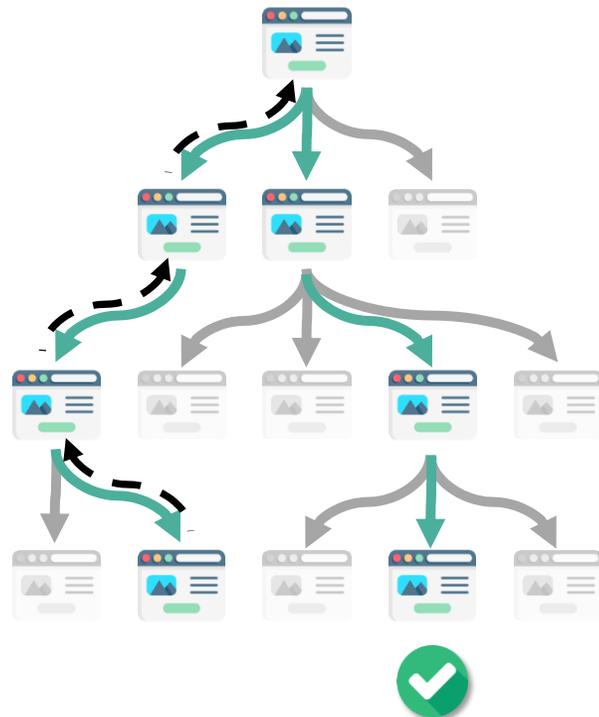
Planning paradigms for language agents

(a) reactive



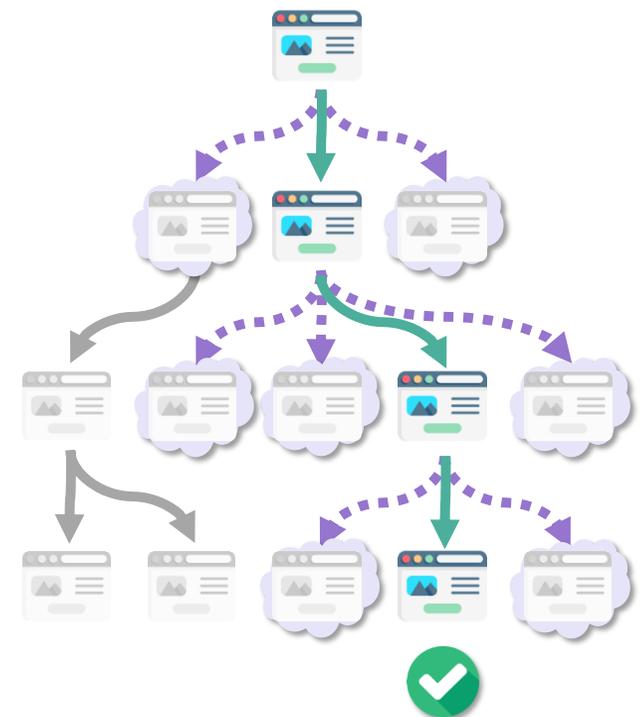
-  fast, easy to implement
-  greedy, short-sighted

(b) tree search with real interactions



-  systematic exploration
-  irreversible actions, unsafe, slow

(c) model-based planning



-  faster, safer, systematic exploration
-  how to get a world model?

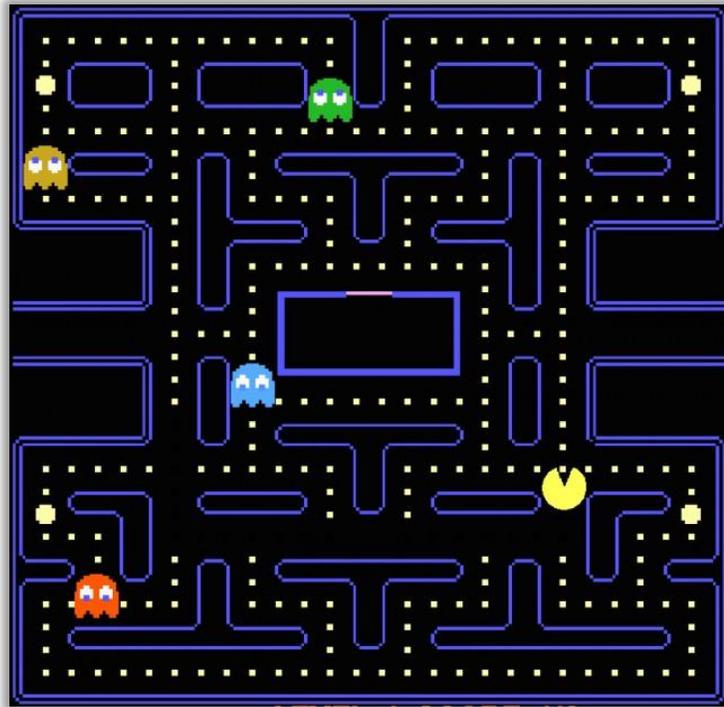
What's ... a world model?

A computational model of environment transition dynamics

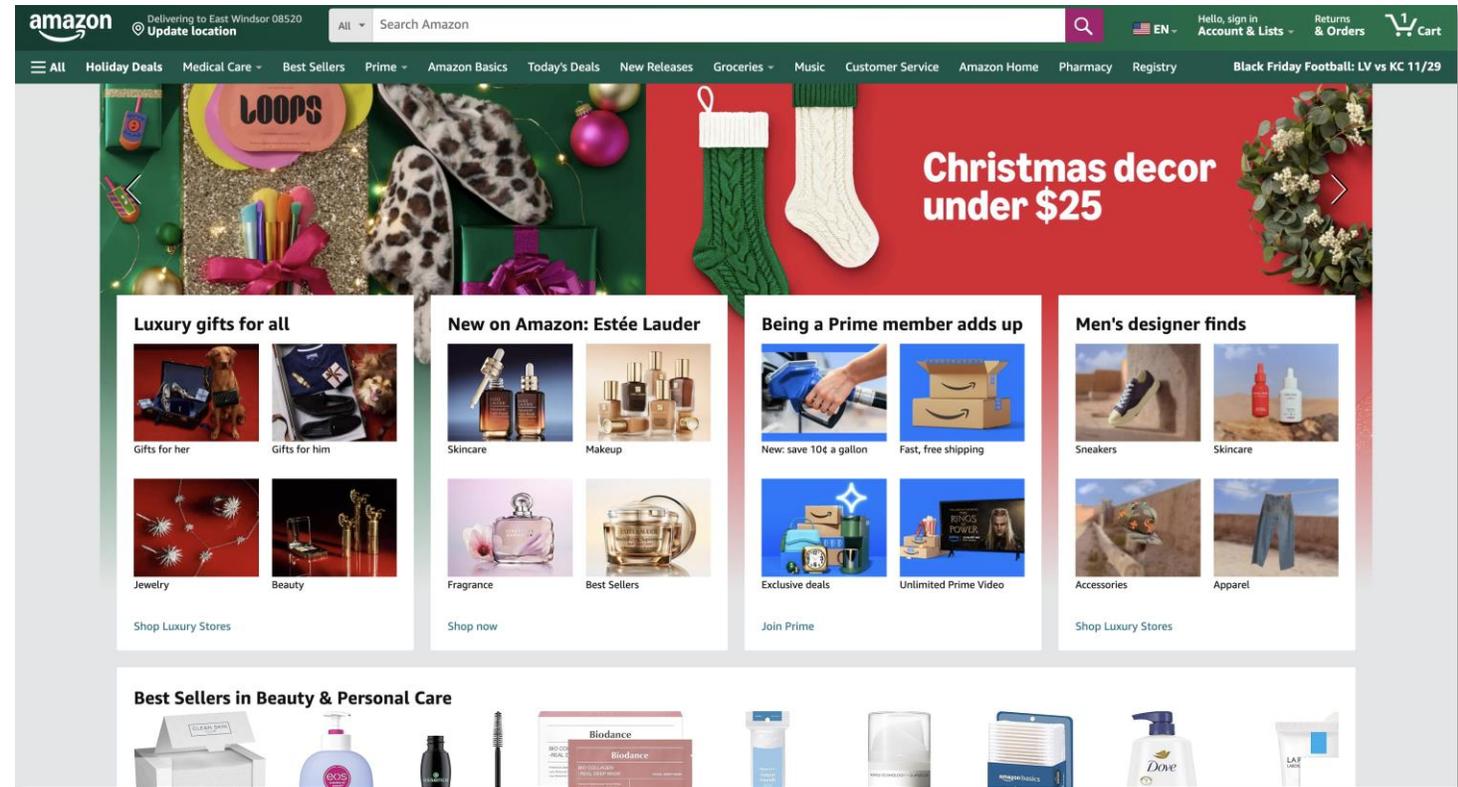
$$\hat{T}: \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}$$

If I do this (a_t) right now (s_t), what would happen next (s_{t+1})?

Why hasn't it been done already?

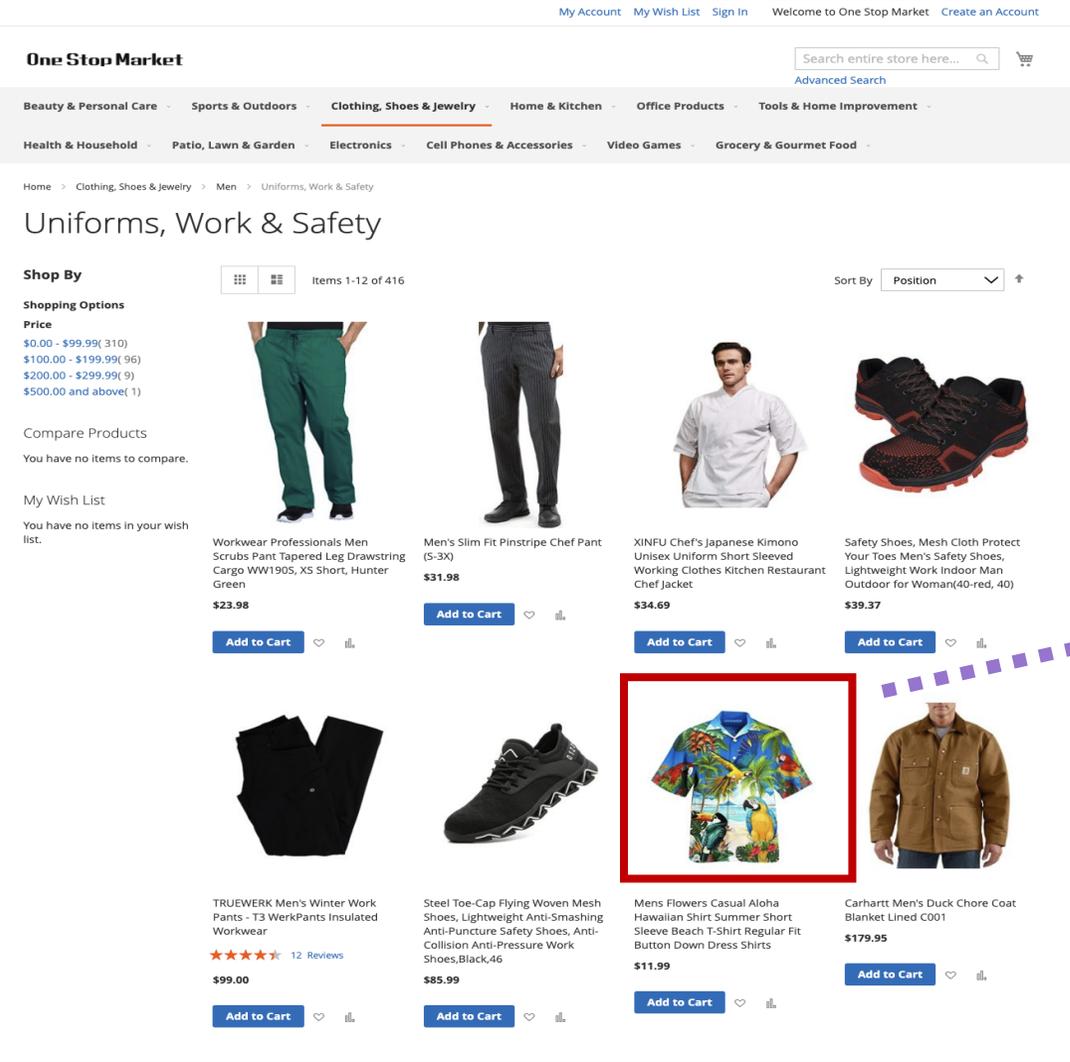


V.S



And **billions of other websites** on the Internet!

LLMs can predict state transitions



The page will navigate to a detailed product page for the "Mens Flowers Casual Aloha Hawaiian Shirt Summer Short Sleeve Beach T-Shirt Regular Fit Button Down Dress Shirts." This new page will likely contain additional information about the product including more detailed specifications, customer reviews, larger images, sizing options, and possibly a larger "Add to Cart" button. Other elements from the current category view like the grid of products will be replaced with the detailed view of this specific product.

WebDreamer: model-based planner for web agents

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

The screenshot shows the One Stop Market website interface. At the top, there is a navigation bar with links for 'My Account', 'My Wish List', 'Sign In', 'Welcome to One Stop Market', and 'Create an Account'. Below this is a search bar with the placeholder text 'Search entire store here...' and a magnifying glass icon. A red circle with the number '3' and the text 'Type 'Disk'' points to the search bar. Below the search bar is a horizontal menu with categories: 'Beauty & Personal Care', 'Sports & Outdoors', 'Clothing, Shoes & Jewelry', 'Home & Kitchen', 'Office Products', and 'Tools & Home Improvement'. A red circle with the number '1' and the text 'Click 'Office Products'' points to the 'Office Products' category. Below the menu is another horizontal menu with categories: 'Health & Household', 'Patio, Lawn & Garden', 'Electronics', 'Cell Phones & Accessories', 'Video Games', and 'Grocery & Gourmet Food'. A red circle with the number '2' and the text 'Click 'Electronics'' points to the 'Electronics' category. The main content area is titled 'One Stop Market' and 'Product Showcases'. It displays a grid of product cards, each with an image, a title, a price, and an 'Add to Cart' button. The products shown include: Pre-baked Gingerbread House Kit, V8 Energy, Healthy Energy Drink, Elmwood Inn Fine Teas, Belle Of The Ball Princess Sprinkle Mix, So Delicious Dairy Free CocoWhip, Cheongju Sweet Potato Starch Powder, Q Mixers Premium Ginger Ale, Stove Top Turkey Stuffing Mix, Briess DME - Pilsen Light, and Tony's Chachere's More Spice Creole Seasoning.

WebDreamer: model-based planner for web agents

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

Stage I: Simulation

1

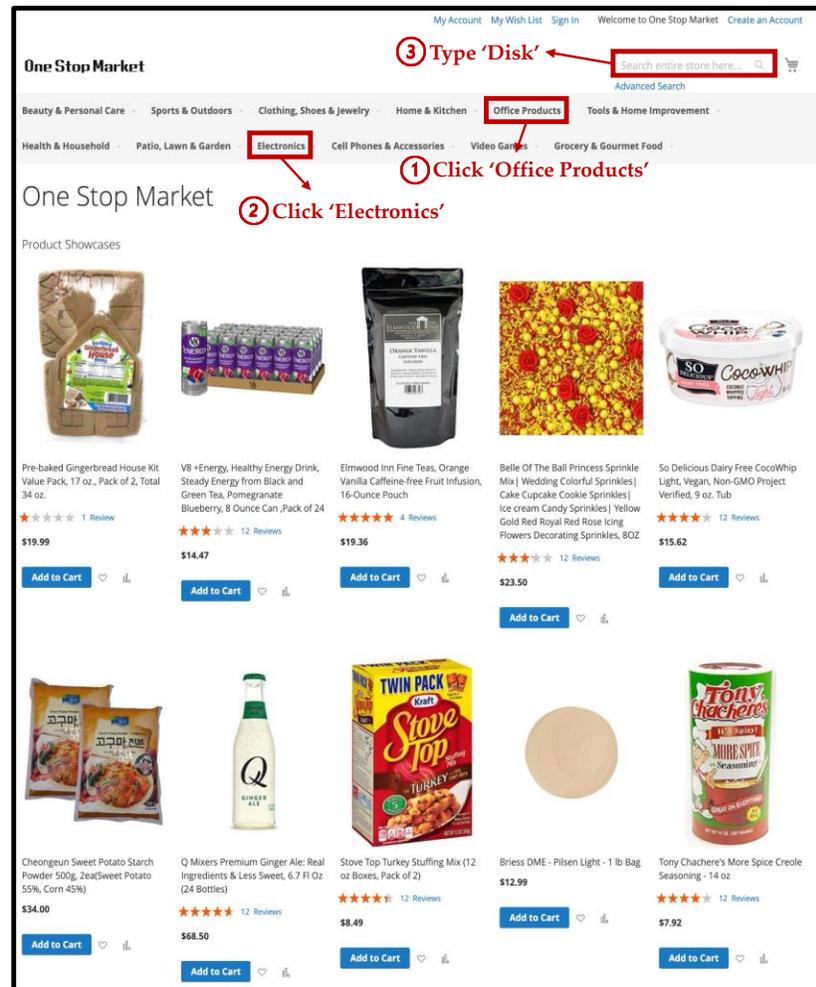
The 'Office Products' category will display three sub-categories: 'Office Electronics', 'Office & School Supplies', and 'Office Furniture & Lighting'.

The screenshot shows the One Stop Market website with several annotations: 1. A red box highlights the search bar with the text '3 Type 'Disk'' and an arrow pointing to the search input. 2. A red box highlights the 'Office Products' category in the top navigation bar with the text '1 Click 'Office Products''. 3. A red box highlights the 'Electronics' category in the sub-navigation bar with the text '2 Click 'Electronics''. The main content area displays a grid of product showcases with images, titles, prices, and 'Add to Cart' buttons.

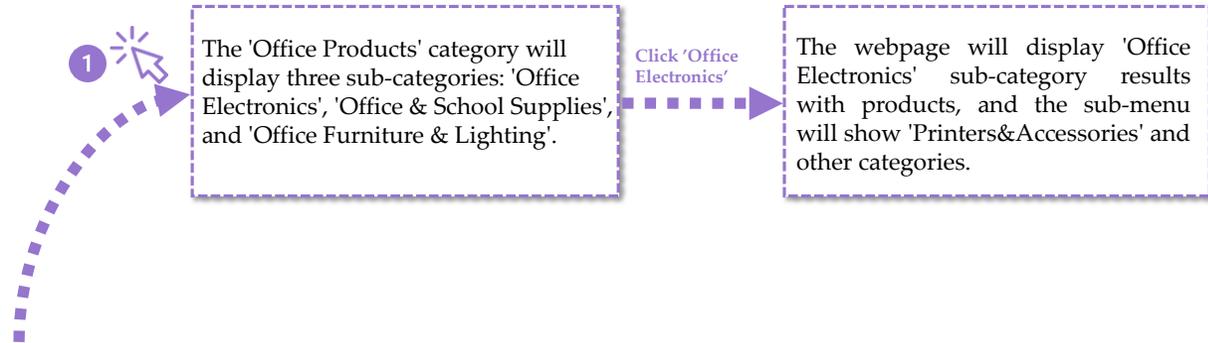
Product Name	Price	Reviews
Pre-baked Gingerbread House Kit Value Pack, 17 oz., Pack of 2, Total 34 oz.	\$19.99	1 Review
V8 Energy, Healthy Energy Drink, Steady Energy from Black and Green Tea, Pomegranate Blueberry, 8 Ounce Can, Pack of 24	\$14.47	12 Reviews
Elmwood Inn Fine Teas, Orange Vanilla Caffeine-free Fruit Infusion, 16-Ounce Pouch	\$19.36	4 Reviews
Belle Of The Ball Princess Sprinkle Mix Wedding Colorful Sprinkles Cake Cupcake Cookie Sprinkles Ice cream Candy Sprinkles Yellow Gold Red Royal Red Rose Icing Flowers Decorating Sprinkles, 8OZ	\$23.50	12 Reviews
So Delicious Dairy Free CocoWhip Light, Vegan, Non-GMO Project Verified, 9 oz. Tub	\$15.62	12 Reviews
Cheongun Sweet Potato Starch Powder 500g, 2ea/Sweet Potato 55%, Corn 45%	\$34.00	12 Reviews
Q Mixers Premium Ginger Ale: Real Ingredients & Less Sweet, 6.7 Fl Oz. (24 Bottles)	\$68.50	12 Reviews
Stove Top Turkey Stuffing Mix (12 oz Boxes, Pack of 2)	\$8.49	12 Reviews
Briess DME - Pilsen Light - 1 lb Bag	\$12.99	12 Reviews
Tony Chachere's More Spice Creole Seasoning - 14 oz	\$7.92	12 Reviews

WebDreamer: model-based planner for web agents

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

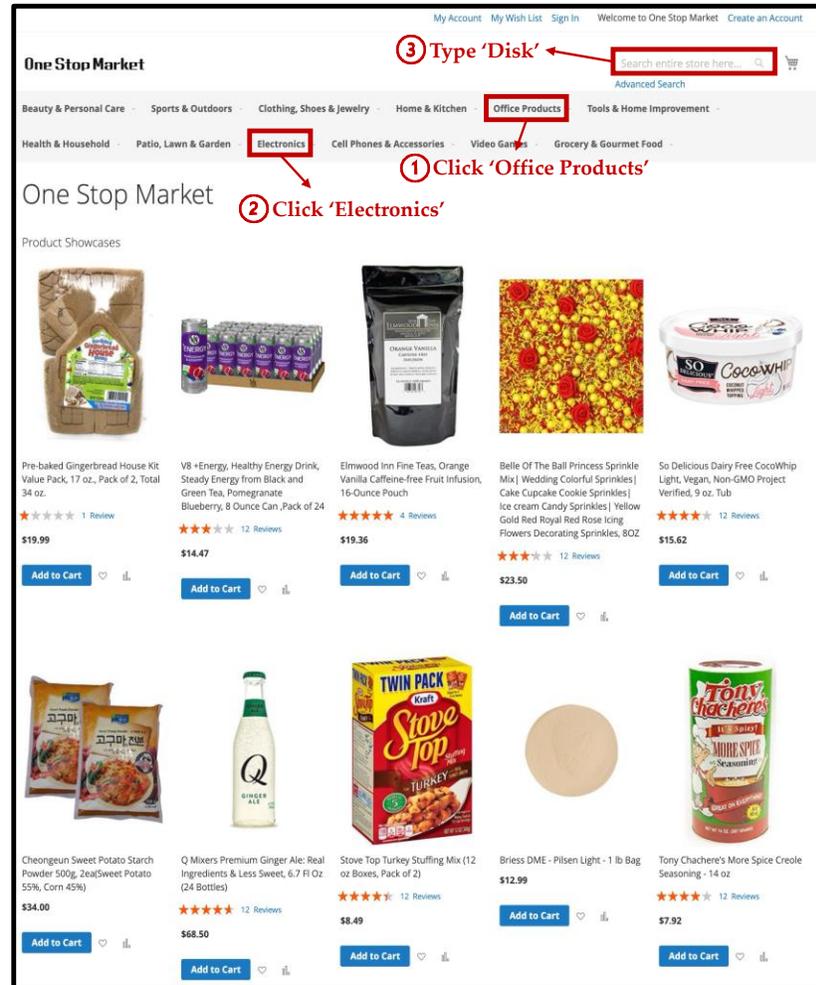


Stage I: Simulation

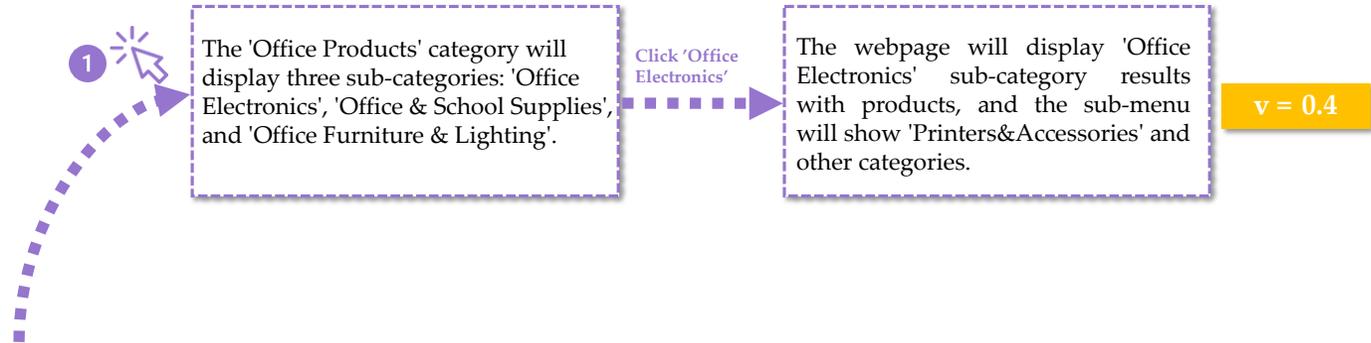


WebDreamer: model-based planner for web agents

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

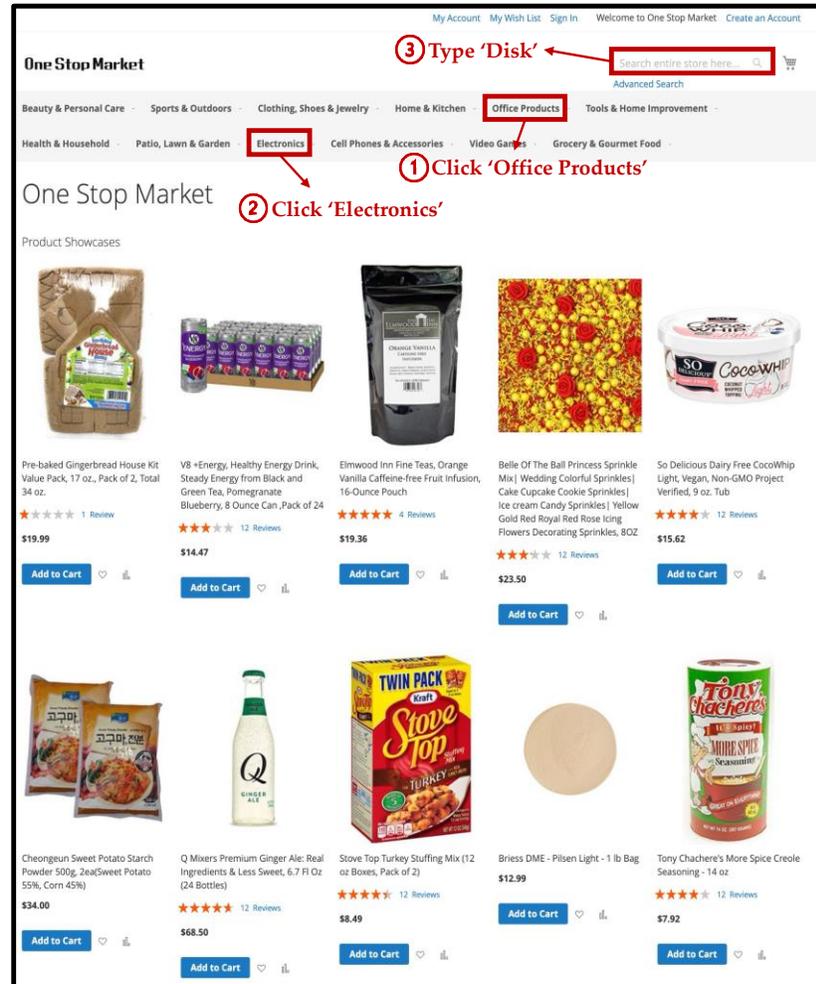


Stage I: Simulation

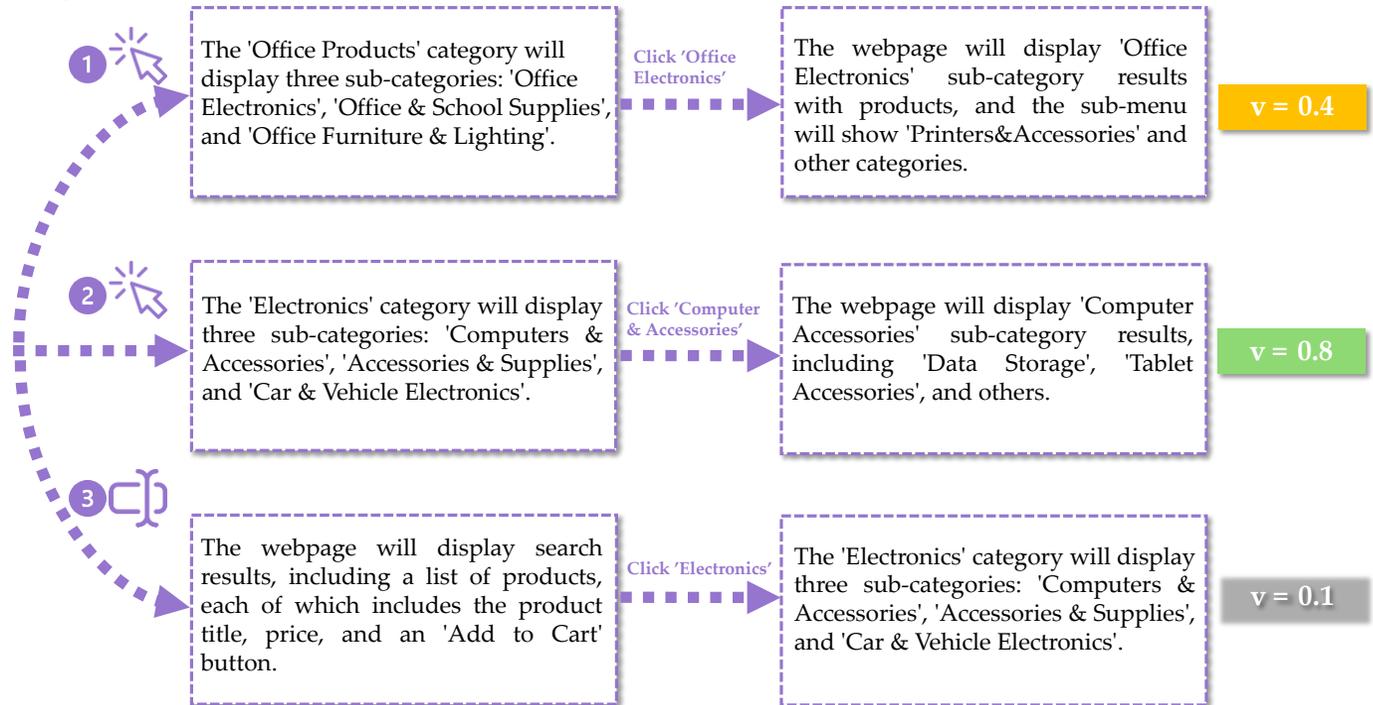


WebDreamer: model-based planner for web agents

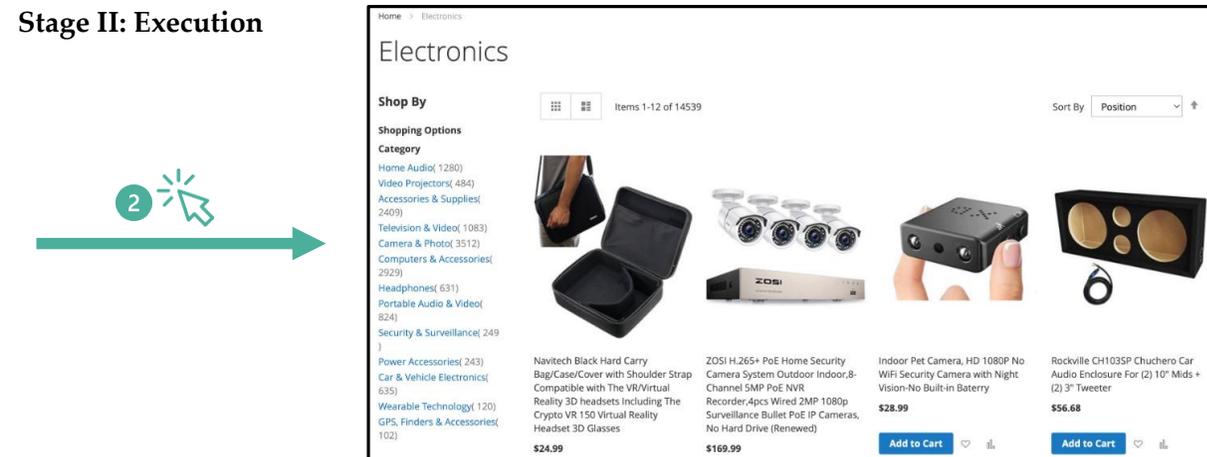
Please navigate to the 'Data Storage' category and purchase the least expensive disk with 512GB of storage.



Stage I: Simulation



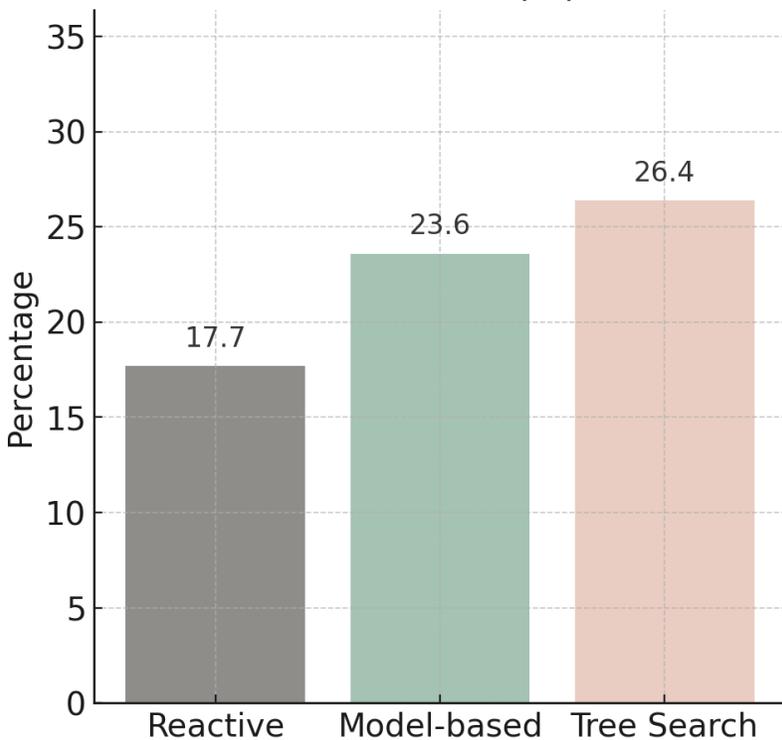
Stage II: Execution



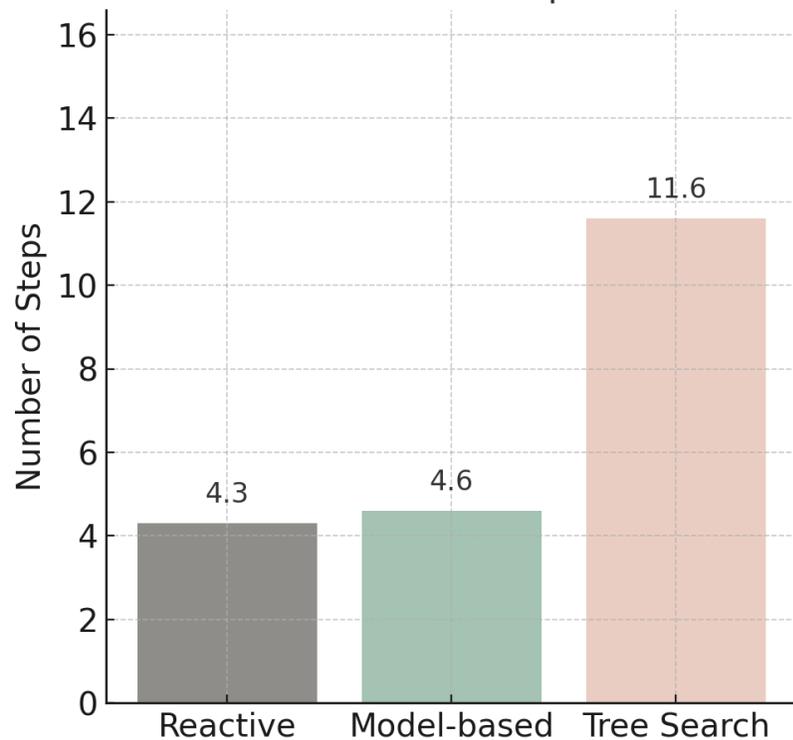
Results on VisualWebArena

Model-based planning is more accurate than reactive planning and more efficient than tree search

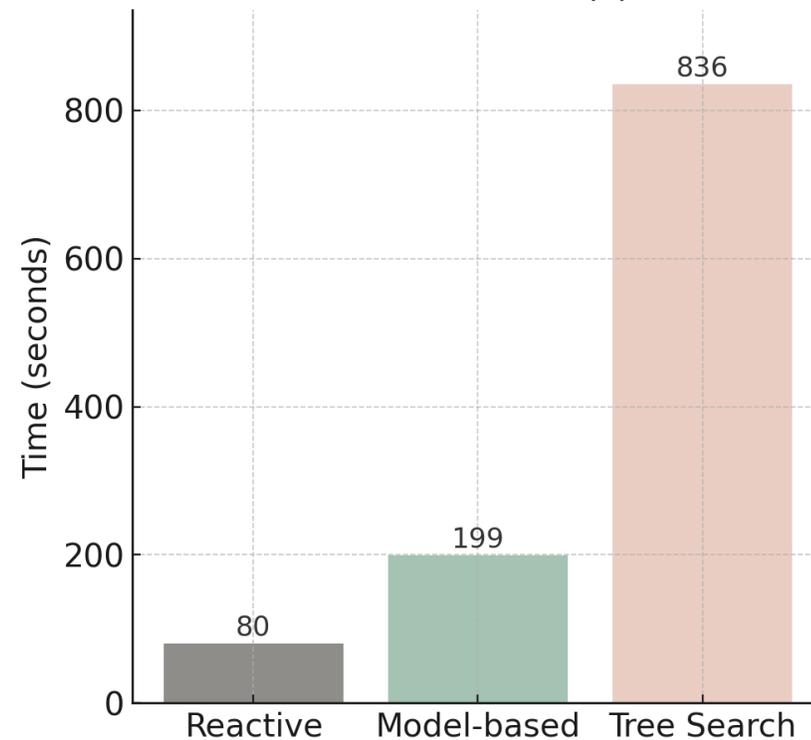
Success Rate (%) ↑



of Action Steps ↓



Wall Clock Time (s) ↓

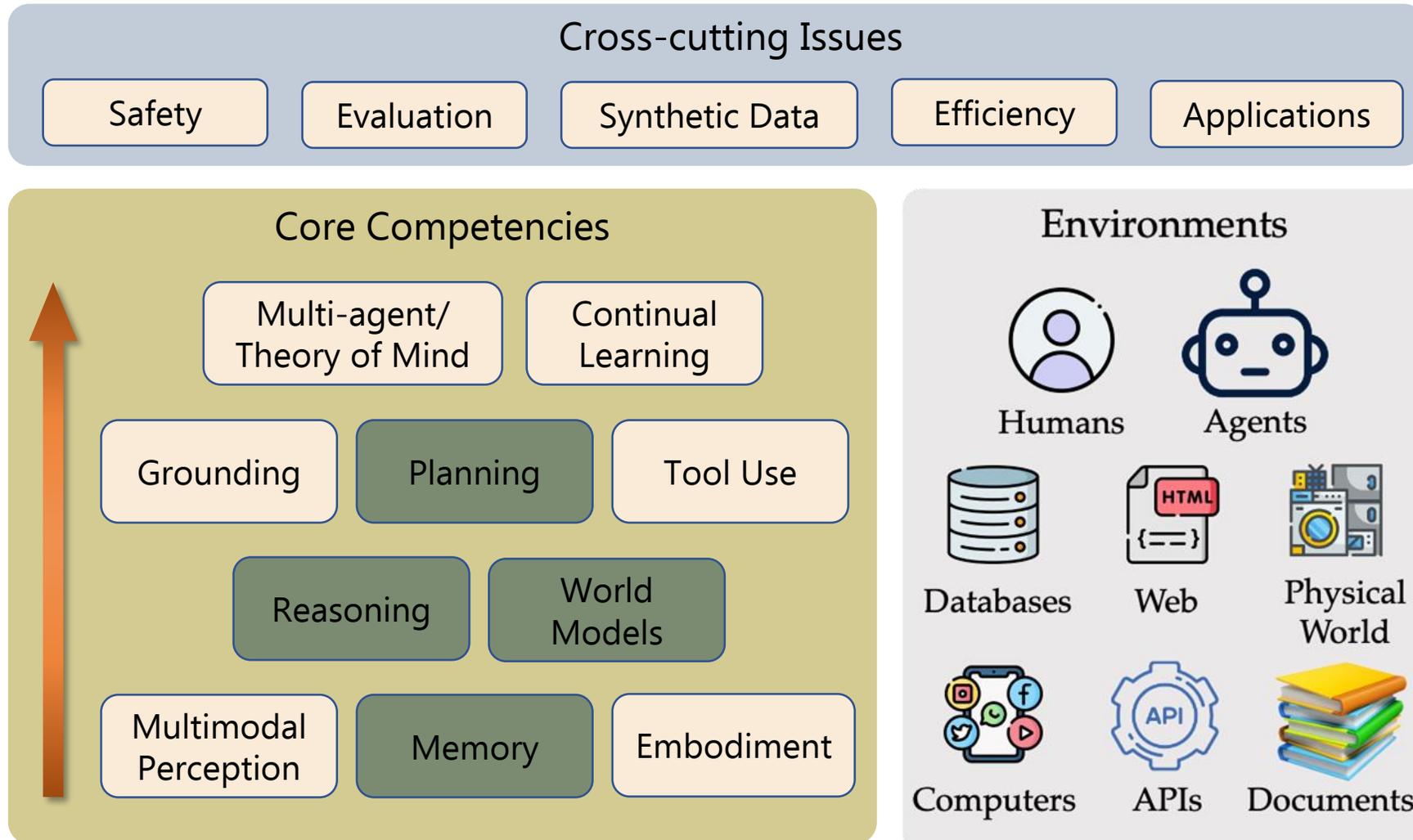


Baselines are from Koh et al., (2024)

Planning: takeaways

- Language agents are expanding into new planning scenarios
 - Characterized by expressive but fuzzy goal specifications, open-ended action spaces, more difficult and sometimes non-binary goal tests
- Language for reasoning also enables new planning abilities
 - Generalist world models and model-based planning
 - Hierarchical planning and dynamic replanning
- The best planning strategy is dependent on the LLM; stronger LLMs may require less scaffolding (i.e., more 'reactive')
- How to improve planning in LLMs is still largely an open question
 - Can the recipe for o1/R1-style reasoning work for planning?

We are just standing at the dawn of a long journey



Some future directions for language agents

- Memory, personalization, and continual learning
 - *How can an agent continually learn from use and exploration?*
- Reasoning
 - *O1/R1-style reasoning? How to get reliable rewards? How to integrate external actions and environmental states?*
- Planning
 - *How to build better world models? How to balance reactive and model-based planning? How to sustain a long horizon without losing focus?*
- Safety
 - *Endogenous risks vs. exogenous risks*
- Applications
 - *Agentic search (deep research), workflow automation, science agents*

Language Agents: Foundations, Prospects, and Risks

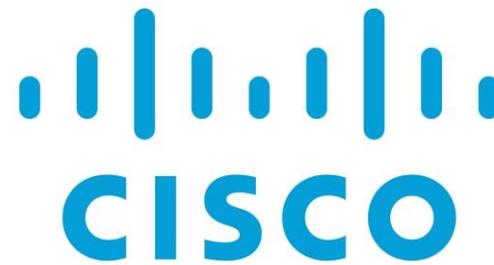
Yu Su, Diyi Yang, Shunyu Yao, Tao Yu

<https://language-agent-tutorial.github.io/>

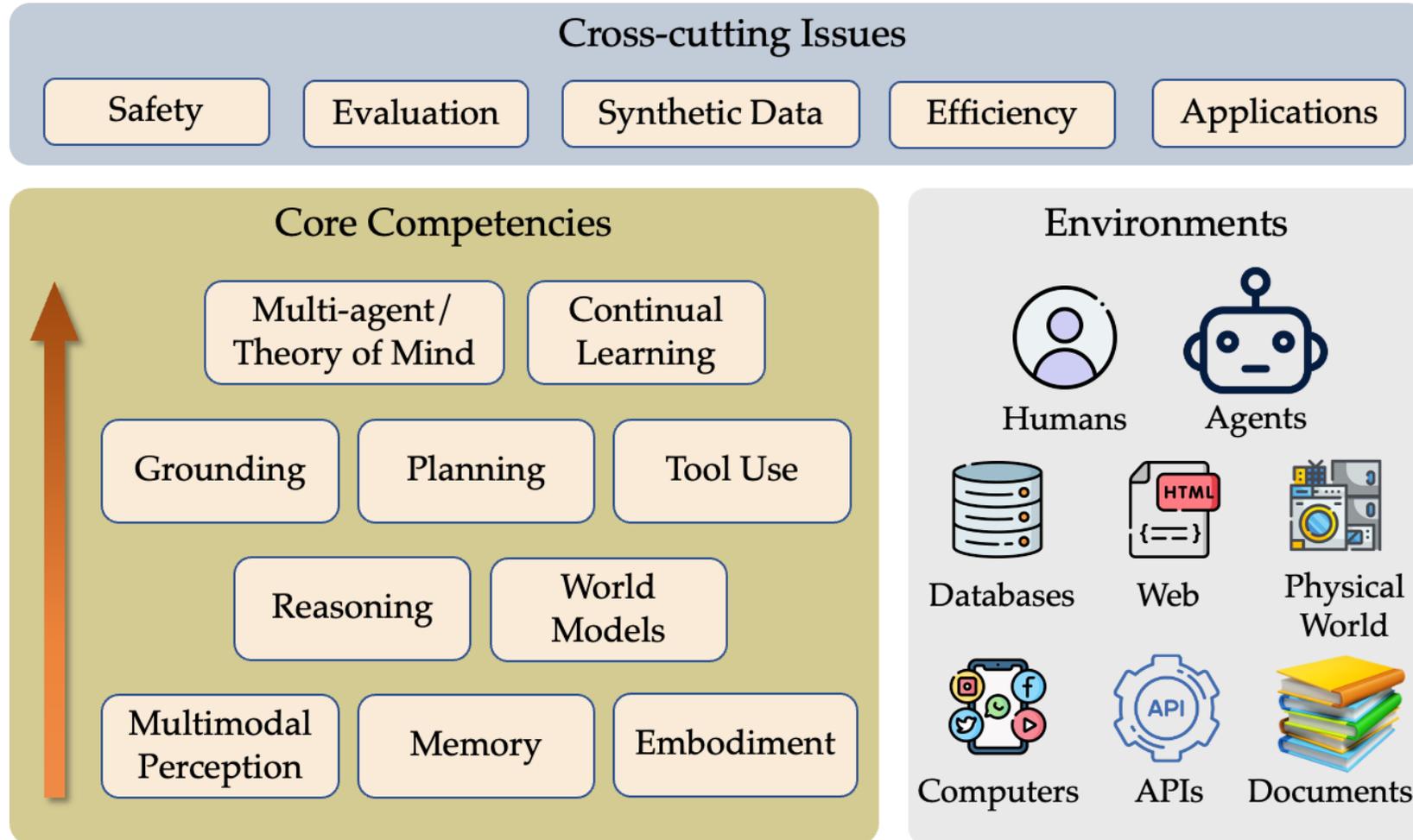


香港大學
THE UNIVERSITY OF HONG KONG

Acknowledgements

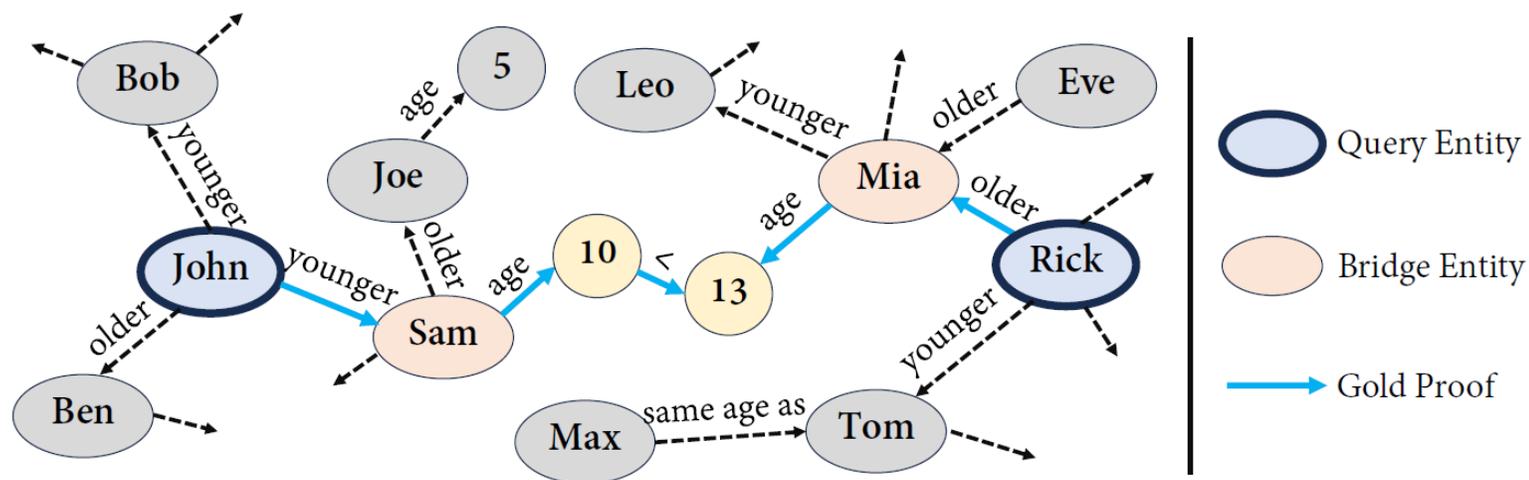


Thanks &



Reasoning task with a Large Search Space

- Large search space
 - Each query entity connects to ~50 facts
 - Each bridge entity (in ground truth proof) connects to >900 facts
- No surface form clues to exploit & bias the proof search
 - Unlike most conventional QA benchmarks where the proof steps are transparent



The Power of Parametric Memory for Complex Reasoning

- SoTA LLMs with non-parametric memory fail badly
 - Current LLMs still cannot reason deeply with non-parametric knowledge representations
 - Also no major improvement from o1-preview or o3-mini (high)
- Grokked transformer achieves near-perfect accuracy by integrating and compressing the facts to the extreme

Table 1: Results on the complex reasoning task. Direct/CoT: predict the answer directly/verbalize the reasoning steps. “+R”: retrieval augmentation.

	GPT-4-Turbo		Gemini-Pro-1.5				Grokked Transformer
	Direct+R	CoT+R	Direct	CoT	Direct+R	CoT+R	
Accuracy (%)	33.3	31.3	28.7	11.3	37.3	12.0	99.3