

Planning



Planning: (simplified) definition

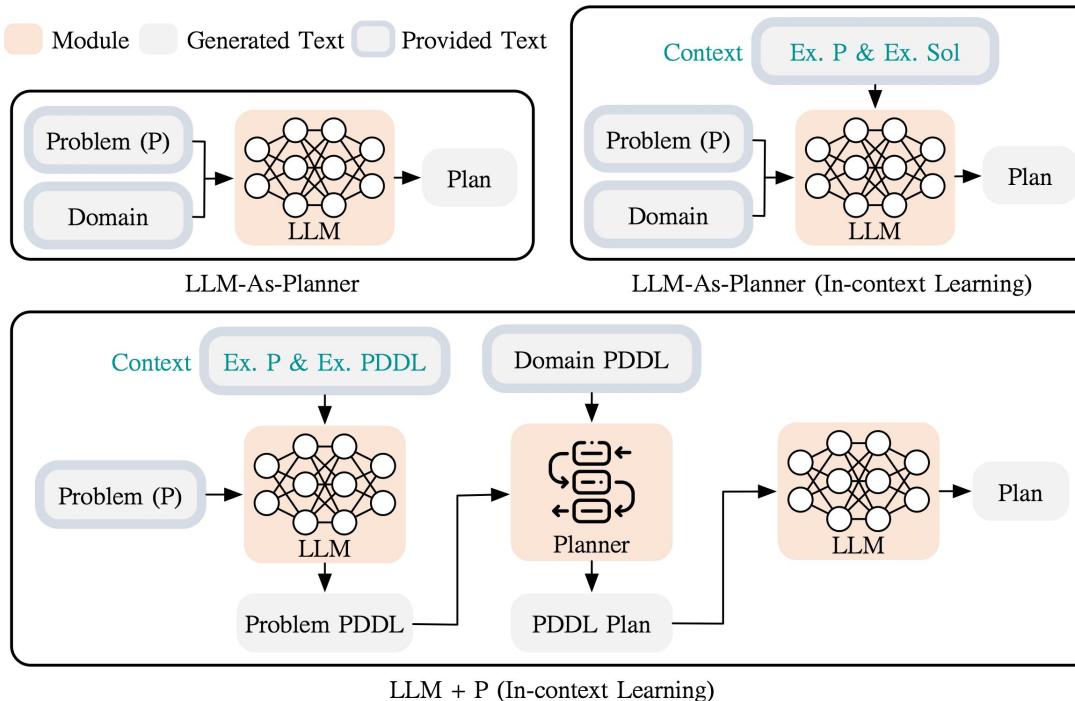
Given a goal G , decides on a sequence of actions (a_0, a_1, \dots, a_n) that will lead to a state that passes the goal test $g(\cdot)$

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 <i>TYPE: auto repair</i>	CLICK
2. [button] Auto Repair	CLICK
3. [textbox] Near <i>10002</i>	CLICK
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:

Action 1



Action 2



Action 5



Action 6



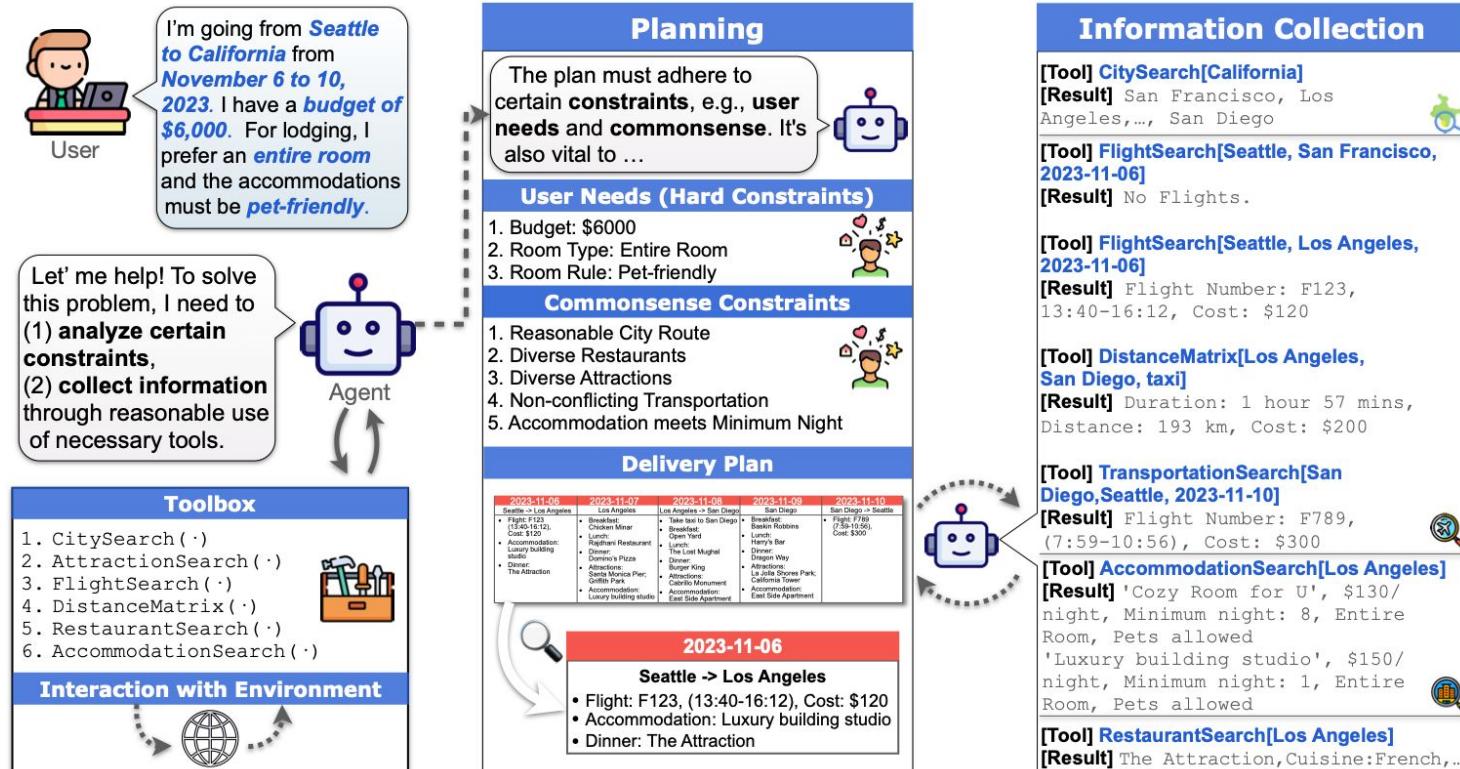
Action 9



Action 10

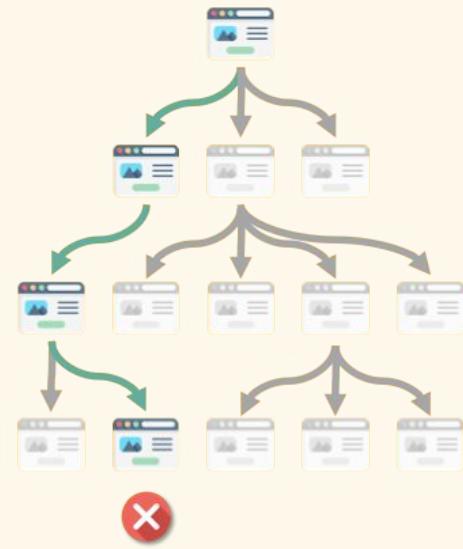


Language agent planning: travel planning

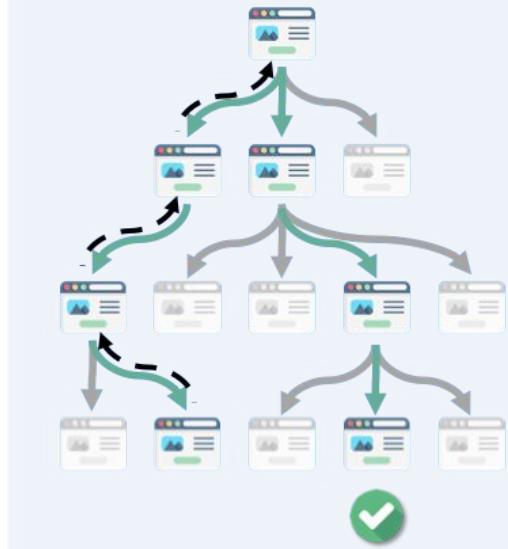


Planning paradigms for language agents

(a) reactive



(b) tree search with real interactions

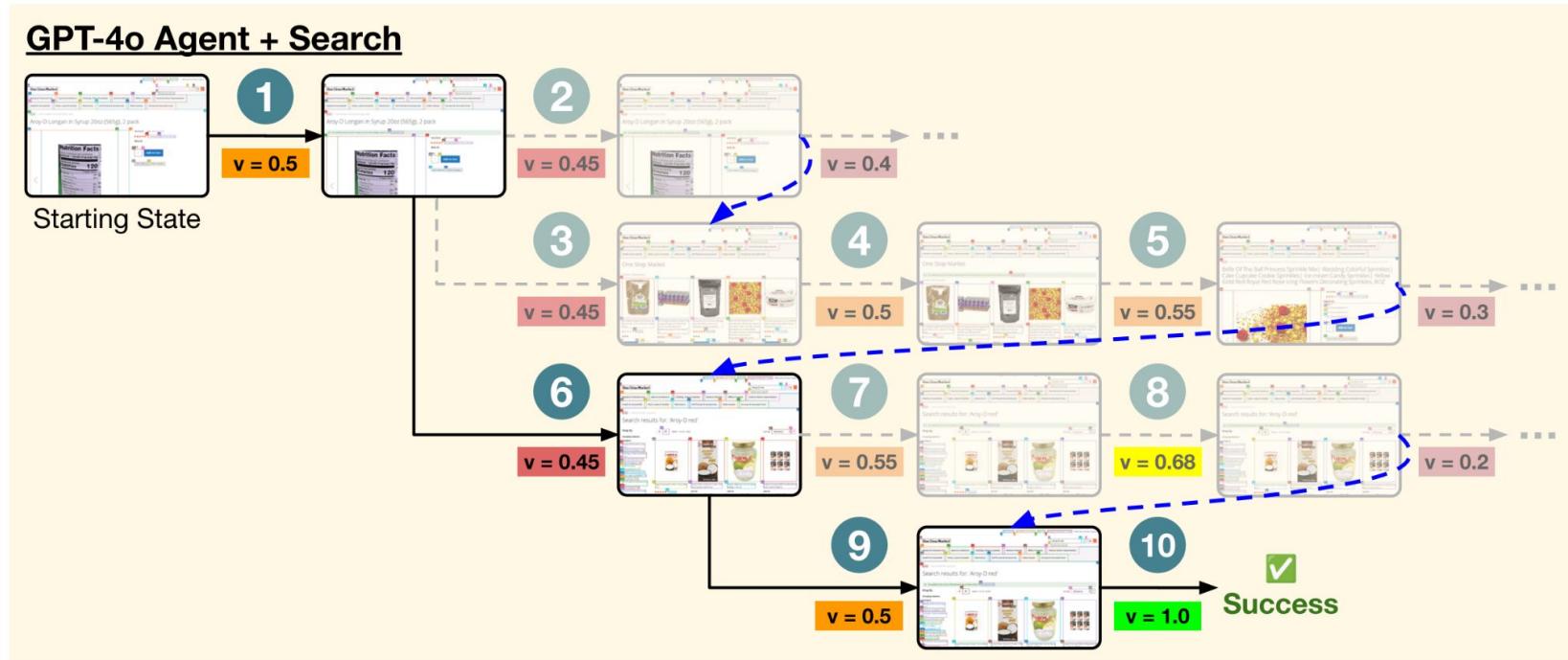


fast, easy to implement



greedy,
short-sighted

Tree search with real interactions



Jing Yu Koh, Stephen McAleer, Daniel Fried, Ruslan Salakhutdinov. "Tree Search for Language Model Agents." arXiv preprint arXiv:2407.01476 (2024).
Andy Zhou, Kai Yan, Michal Shlapentokh-Rothman, Haohan Wang, Yu-Xiong Wang. "Language Agent Tree Search Unifies Reasoning Acting and Planning in Language Models." ICML (2024).

Challenges with tree search in the real world

- 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...

Nespresso Capsules Vertuo, Variety Pack, Medium and Dark Roast Coffee, 30 Count Coffee Pods, Brews 7.8 oz. \$37.50 (\$1.25 / Count)

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](#).

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](#).

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.

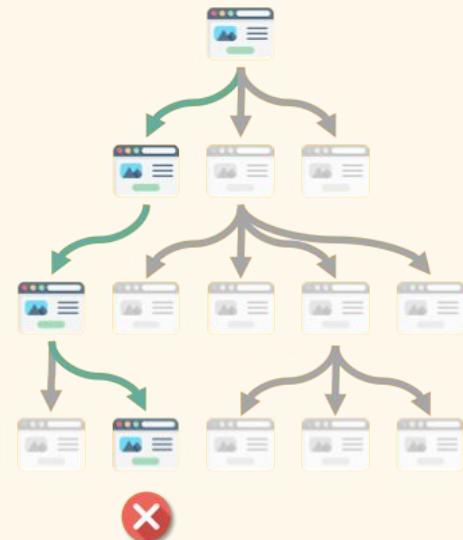
Branded Store Experience Location-based Augmented Reality
We use your camera, motion, and location to power branded store experiences.
Requires camera, motion, and location.

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

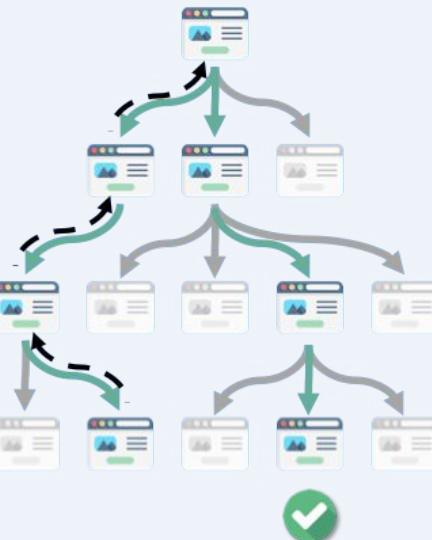
Delivery Location
We use your location to improve your shopping experience, ensuring you only see products and delivery options available in your area.

Planning paradigms for language agents

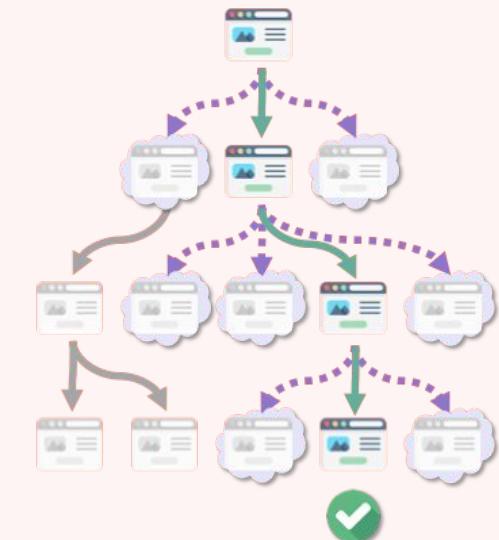
(a) reactive



(b) tree search with real interactions



(c) model-based planning



fast, easy to implement



greedy,
short-sighted



systematic exploration



irreversible actions,
unsafe, slow



faster, safer,
systematic exploration



how to get a world model?

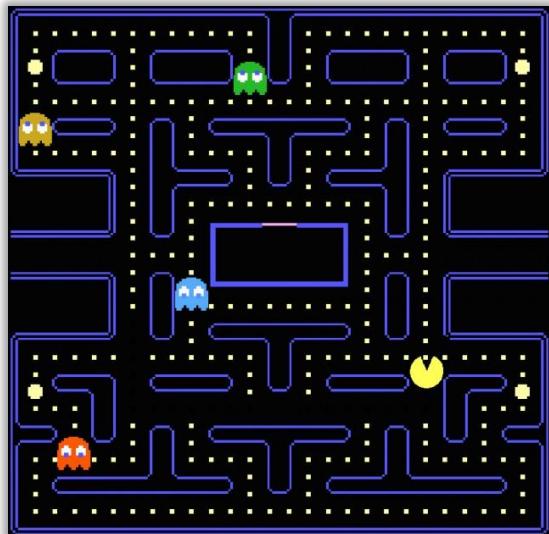
What's ... a world model?

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 yet?



V.S

The screenshot shows the Amazon homepage during the Black Friday sale. The top navigation bar includes links for 'Hello, sign in', 'Account & Lists', 'Returns & Orders', and a shopping cart icon. The main banner features a red background with the text 'Christmas decor under \$25' and various holiday gift items. Below the banner are several promotional sections: 'Luxury gifts for all' (with sub-links for 'Gifts for her', 'Gifts for him', 'Skincare', 'Makeup', 'Jewelry', and 'Beauty'), 'New on Amazon: Estée Lauder', 'Being a Prime member adds up' (with sub-links for 'New: save 10¢ a gallon', 'Fast, free shipping', 'Exclusive deals', and 'Unlimited Prime Video'), 'Men's designer finds' (with sub-links for 'Sneakers', 'Skincare', 'Accessories', and 'Apparel'), and 'Best Sellers in Beauty & Personal Care' (showing various skincare and makeup products). The bottom of the page shows a row of best-selling beauty products.

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

LLMs can predict state transitions (reasonably well)

My Account My Wish List Sign In Welcome to One Stop Market Create an Account

One Stop Market

Search entire store here... Advanced Search

Beauty & Personal Care - Sports & Outdoors - Clothing, Shoes & Jewelry - Home & Kitchen - Office Products - Tools & Home Improvement -

Health & Household - Patio, Lawn & Garden - Electronics - Cell Phones & Accessories - Video Games - Grocery & Gourmet Food -

Home > Clothing, Shoes & Jewelry > Men > Uniforms, Work & Safety

Uniforms, Work & Safety

Shop By

Shopping Options

Price

- \$0.00 - \$99.99 (310)
- \$100.00 - \$199.99 (96)
- \$200.00 - \$299.99 (9)
- \$500.00 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 4161

Sort By Position

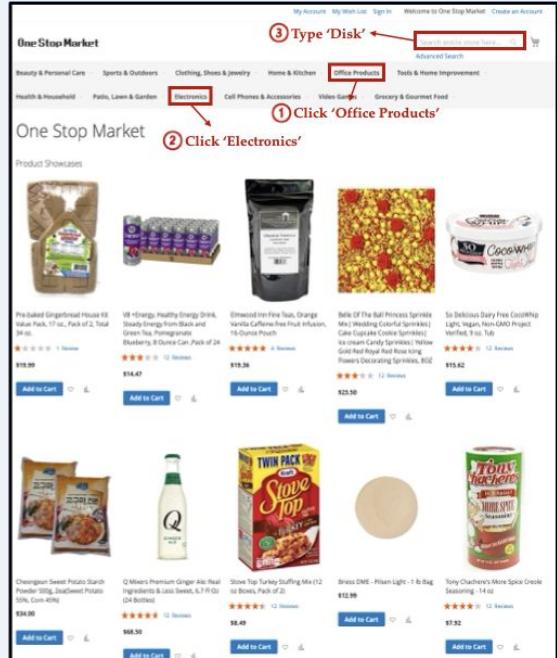
Image	Product Name	Price	Add to Cart
	Men's Slim Fit Pinstripe Chef Pant (S-3X)	\$31.98	Add to Cart
	XINFU Chef's Japanese Kimono Uniform Short Sleeved Working Clothes Kitchen Restaurant Chef Jacket	\$34.69	Add to Cart
	Safety Shoes, Mesh Cloth Protect Work Toes Men's Safety Shoes, Light Weight Work Indoor Man Outdoor for Woman 40-red (40)	\$38.37	Add to Cart
	TRUEWERK Men's Winter Work Pants T3 WorkPants Insulated Workwear	\$23.98	Add to Cart
	Hestra No-Cap Flying V Women Mesh Gloves, Light Weight Anti-Smelling Anti-Puncture Safety Shoes, Anti-Collision Anti-Pressure Work Shoes, Black, 46	\$85.99	Add to Cart
	Mens Flowers Casual Aloha Hawaiian Summer Short Sleeve Beach T-Shirt Regular Fit Button Down Dress Shirts	\$11.99	Add to Cart
	Carhartt Men's Duck Chore Coat Blanket Lined C001	\$179.95	Add to Cart



The page will navigate to a detailed product page for the "Mens Flowers Casual Aloha Hawaiian 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.

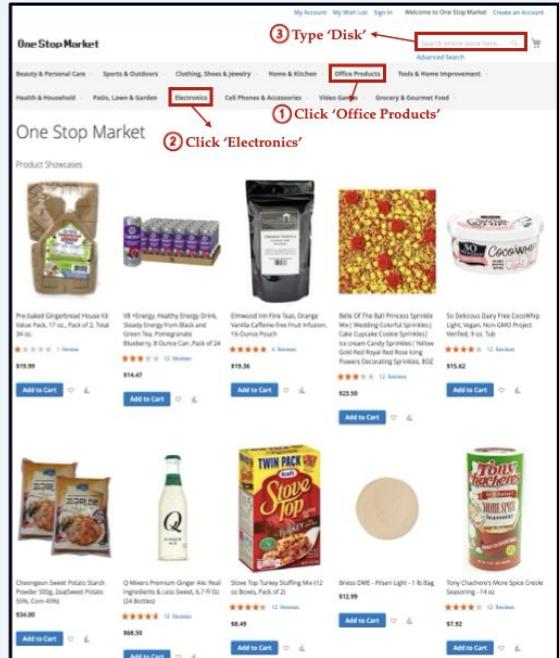
WebDreamer: model-based planner

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

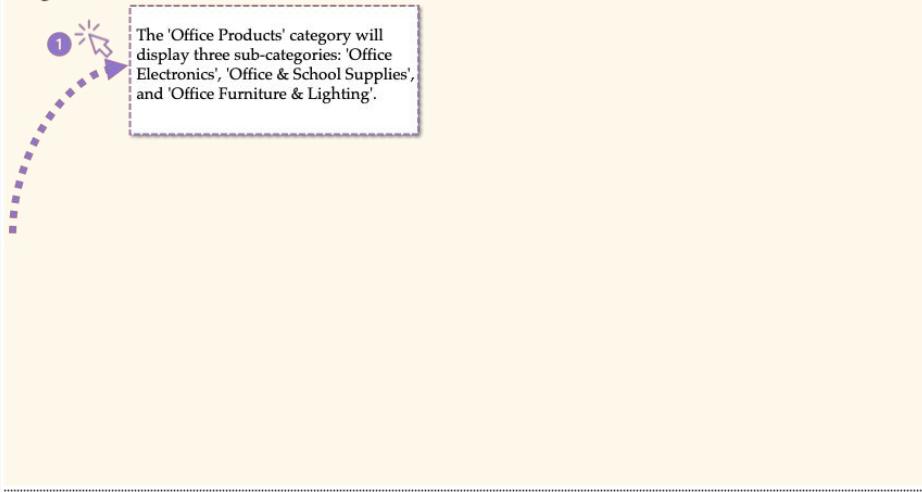


WebDreamer: model-based planner

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

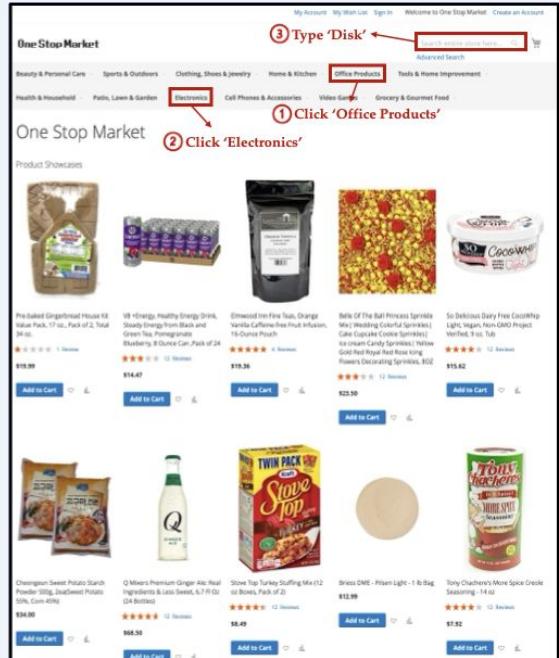


Stage I: Simulation

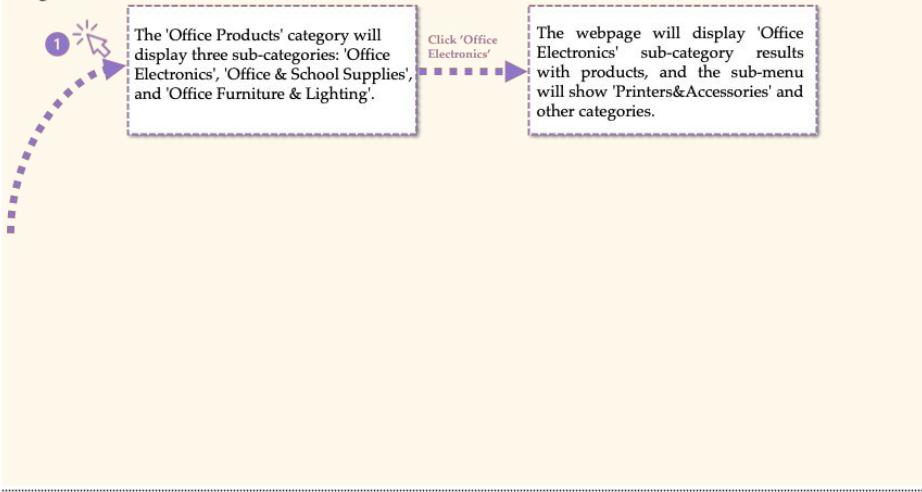


WebDreamer: model-based planner

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

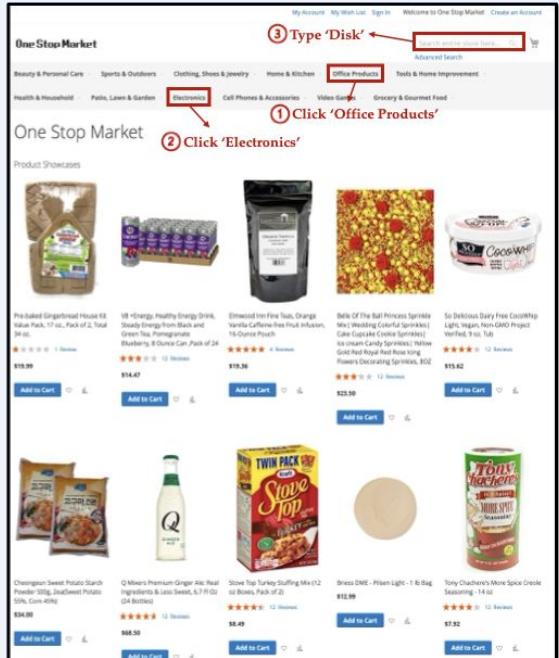


Stage I: Simulation

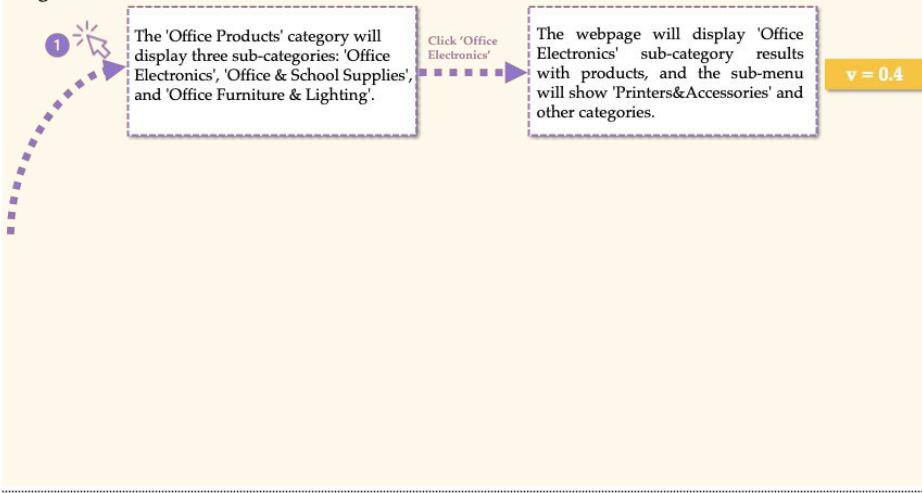


WebDreamer: model-based planner

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

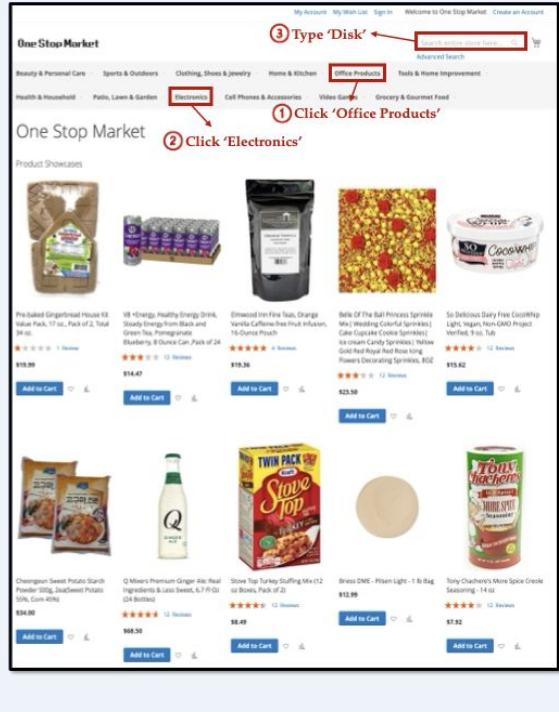


Stage I: Simulation

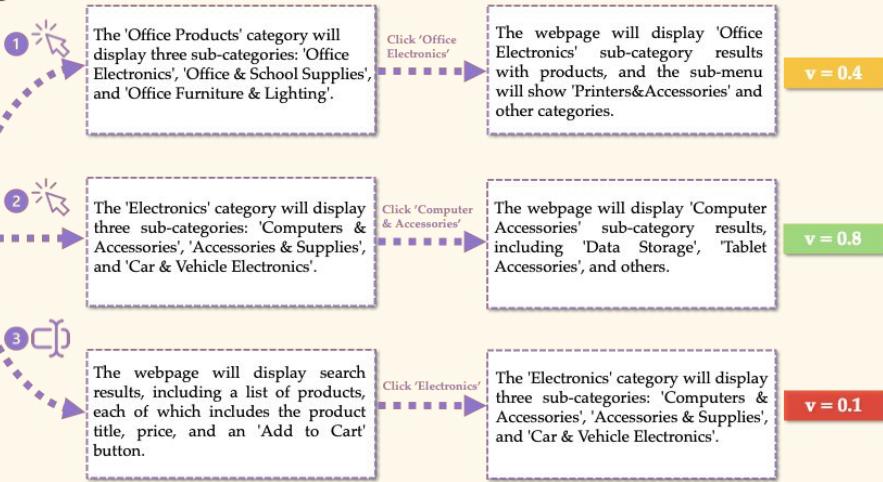


WebDreamer: model-based planner

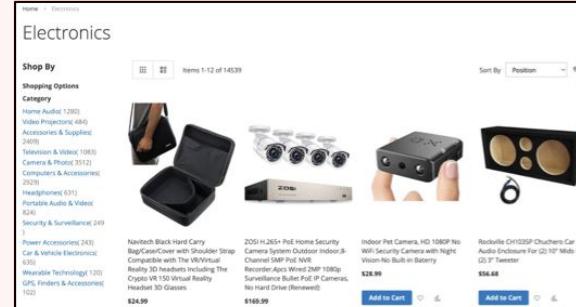
Please navigate to the 'Data Storage' category and purchase the least expensive disk with 512GB of storage.



Stage I: Simulation

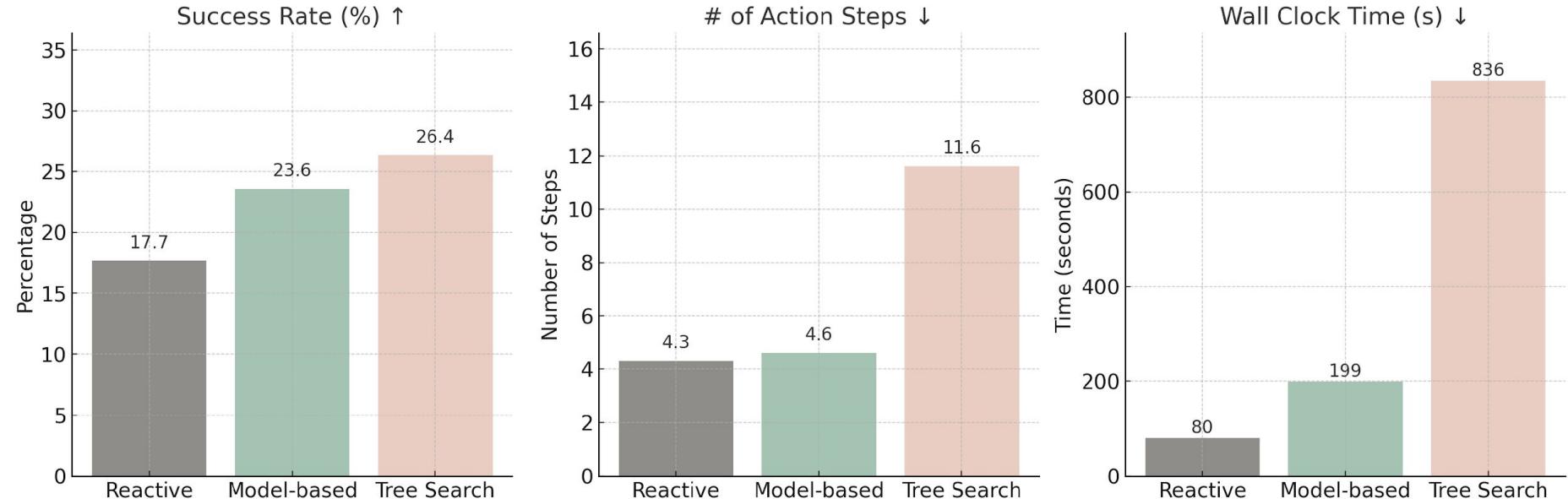


Stage II: Execution



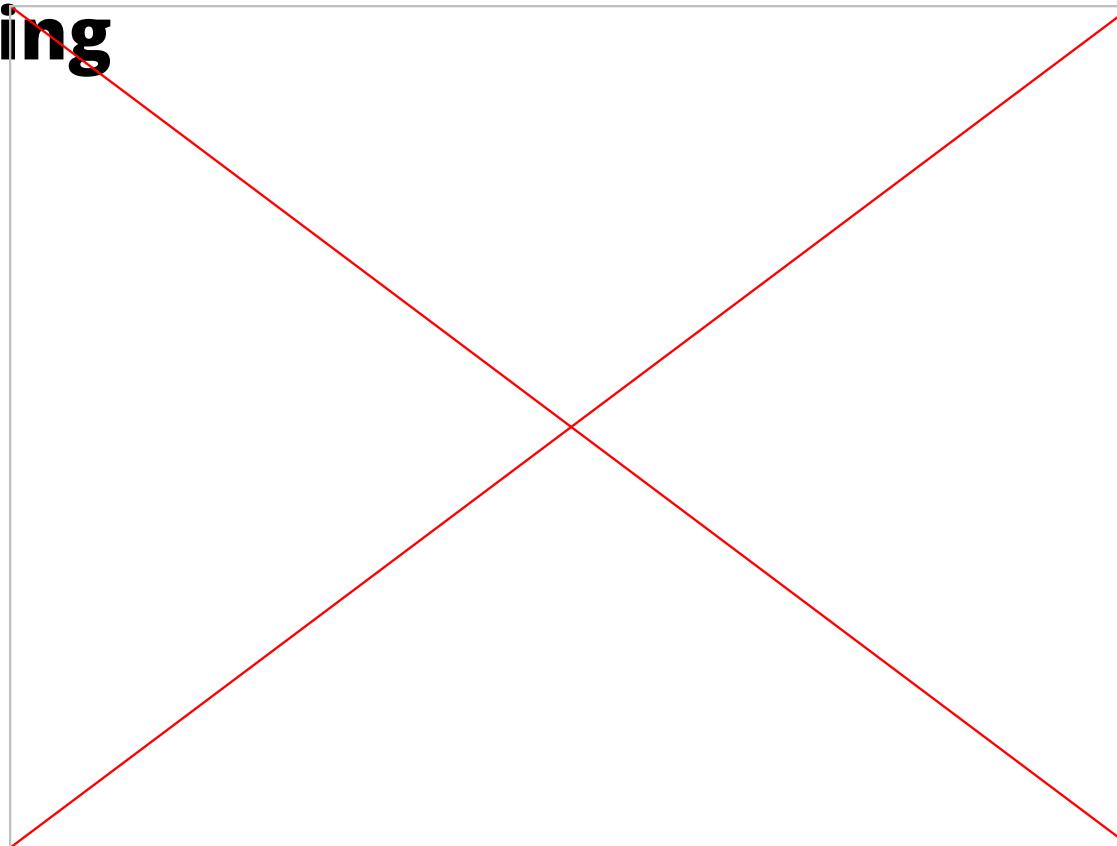
Results on VisualWebArena

Model-based planning is more accurate than reactive planning and more efficient than tree search
(also recall the challenges of tree search in real-world environments)



Reactive and tree search results are from Koh et al., (2024)

Hierarchical planning and dynamic replanning



Hierarchical planning and dynamic rep

The image shows a screenshot of the Devin AI application interface. On the left, there's a chat window titled "Benchmark Llama 2 Providers" between users Scott and Devin. Scott asks Devin to benchmark Llama 2 on three providers. Devin responds, saying he'll look into the API formats and create a script to benchmark Llama 2 across these services. A note at the bottom of the chat says Devin is researching API documentation. At the bottom, there's a message input field and a footer stating Devin is in private beta and may make mistakes.

On the right, the "Devin's Workspace" tab is active. It shows a "Planner" section with a red box highlighting a list of tasks under "Current Focus". The tasks are:

- Research the API documentation for Replicate, Together, and Perplexity to understand their request formats and parameters.
- Write a Python script that constructs and sends HTTP requests to each API with the same prompt and parameters.
- Implement response time measurement and accuracy assessment within the script to benchmark each API's performance.
- Test the script to ensure it works correctly and gathers the necessary data from each API.
- Analyze the collected data to compare the performance of the three APIs.
- Document the script and the benchmarking process, including any dependencies or environment setup needed.
- Report the findings back to the user with the collected data and any relevant visualizations or summaries.

An arrow points from the text "High-level Plan" to the red box around the planner tasks.

Hierarchical planning and dynamic representation

The screenshot shows a dark-themed interface for AI development. On the left, a sidebar labeled "Devin's Workspace" contains a message from Scott asking Devin to ensure model names are all llama 2 70b. Devin responds that they will ensure all model names used in the benchmarking script are Llama 2 70b. Below this, a status message indicates Devin is debugging a script to handle API response errors. In the center, a terminal window titled "default" shows a command-line session where Python is not found, leading to an error. A red box highlights this error output. On the right, a code editor titled "benchmark.py" displays Python code for a benchmarking script, including imports for os, time, requests, and API keys. The code defines constants for API endpoints and keys. At the bottom, a browser window shows the "Chat Completions" documentation for together.ai.

```
Command 'python' not found, did you mean:
  command 'python3' from deb python3
  command 'python' from deb python-is-python3
ubuntu@ip-172-31-27-196:~$ python3 benchmark.py
Error: 422
Traceback (most recent call last):
  File "/home/ubuntu/benchmark.py", line 65, in <module>
    together_tps = benchmark_api(TOGETHER_ENDPOINT, TOGETHER_HEADERS, together_d
ata)
  File "/home/ubuntu/benchmark.py", line 39, in benchmark_api
    tokens = len(response.json()['choices'][0]['text'].split())
KeyError: 'text'
ubuntu@ip-172-31-27-196:~$ _
```

Ran into exception when carrying out a subgoal

Hierarchical planning and dynamic representation

The screenshot shows a browser window with the URL `preview.devin.ai`. The main content is an "Expanded Editor" showing a file named `benchmark.py`. The editor displays two versions of the code side-by-side, with the right version being the current state. A red box highlights the changes made to the `tokens_per_second` function and its surrounding logic. The code uses a dark theme and includes imports from `requests` and `time`.

```
29 # Extended prompt to send to each API to generate more tokens
30 PROMPT = "Please provide a detailed summary of the history of France,
31
32 # Function to measure tokens per second
33 def benchmark_api(endpoint, headers, data):
34     start_time = time.time()
35     response = requests.post(endpoint, headers=headers, json=data)
36     end_time = time.time()
37
38     if response.status_code == 200:
39
40         tokens = len(response.json()['choices'][0]['text'].split())
41         time_taken = end_time - start_time
42         tokens_per_second = tokens / time_taken
43         return tokens_per_second
44     else:
45         print(f"Error: {response.status_code}")
46         return None
47
48 # Data payloads for each API
49 replicate_data = {
50     "version": "0e2509c789964a7ea8736978a43525956ef40397be9033abf9fd2",
51     "input": {"text": PROMPT, "max_tokens": 512}
52 }
53 together_data = {
54     "model": "mistralai/Mixtral-8x7B-Instruct-v0.1",
55     "messages": [{"role": "user", "content": PROMPT}],
56     "max_tokens": 512
57 }
58 perplexity_data = {
59     "model": "gpt-3.5-turbo",
60     "messages": [{"role": "user", "content": PROMPT}],
61     "max_tokens": 512
62 }
63
64 # Function to measure tokens per second
65 def benchmark_api(endpoint, headers, data):
66     start_time = time.time()
67     response = requests.post(endpoint, headers=headers, json=data)
68     end_time = time.time()
69
70     if response.status_code == 200:
71
72         # Print the response for debugging
73         print(f"Response JSON: {response.json()}")
74         try:
75             tokens = len(response.json()['choices'][0]['text'].split())
76             time_taken = end_time - start_time
77             tokens_per_second = tokens / time_taken
78             return tokens_per_second
79         except KeyError as e:
80             print(f"KeyError: {e} in response JSON")
81             return None
82     else:
83         print(f"Error: {response.status_code}, Response: {response.text}")
84         return None
85
86
87 # Data payloads for each API
88 replicate_data = {
89     "version": "0e2509c789964a7ea8736978a43525956ef40397be9033abf9fd2",
90     "input": {"text": PROMPT, "max_tokens": 512}
91 }
92
93 together_data = {
94     "model": "mistralai/Mixtral-8x7B-Instruct-v0.1",
95     "messages": [{"role": "user", "content": PROMPT}],
96     "max_tokens": 512
97 }
98
99 perplexity_data = {
100     "model": "gpt-3.5-turbo",
101     "messages": [{"role": "user", "content": PROMPT}],
102     "max_tokens": 512
103 }
```

Re-plan
and add
print()
around the
exception
to get
debug info

Hierarchical planning and dynamic reporting

The screenshot shows a developer workspace interface with the following components:

- Devin's Workspace**: A sidebar containing a message from Devin (the developer) and a list of tasks:
 - Got it, Scott. I'll create a website showcasing the average tokens per second for each LLM Cloud Service with an interactive element. The title will be "LLM Cloud Service Perf Benchmark" focusing on Llama 2 70b throughput. I'll get started on this now.
 - Initialized a vanilla React app with Chakra UI and charting libraries installed.
 - Designed the frontend to display LLM Cloud Service performance metrics in an interactive chart using React and Chart.js.
 - Resolved Chart.js configuration errors by registering required components and updating chart options.
- default**: A main panel displaying build logs:

```
Compiled successfully!
You can now view llm-perf-app in the browser.
Local: http://localhost:3000
On Your Network: http://172.31.31.156:3000
Note that the development build is not optimized.
To create a production build, use npm run build.
webpack compiled successfully
```
- App.js**: A code editor showing the source code for App.js:

```
1 import React from 'react';
2 import { ChakraProvider, Box, Text, VStack, theme } from '@chakra-ui/react';
3 import { Bar } from 'react-chartjs-2';
4 import {
5   Chart as ChartJS,
6   CategoryScale,
7   LinearScale,
8   BarElement,
9   Title,
10  Tooltip,
11  Legend,
12 } from 'chart.js';
13
14 ChartJS.register(
```
- Preview**: A preview window showing a blank white page at <http://localhost:3000/>.

Solved the exception based on the printed out debug info!

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