

Evaluating a ReAct-Based Agent for Agricultural Planning

Anonymous Submission

Anonymous affiliation

Abstract

Planning is a hallmark of intelligence, enabling both humans and artificial agents to navigate complex environments, adapt to dynamic conditions, and achieve intricate goals. Significant emphasis has been placed on enhancing the planning capabilities of large language models (LLMs) using agents. ReAct, an agentic framework, introduces a paradigm that combines reasoning traces and actionable steps in an interleaved, iterative process, allowing LLMs to adaptively plan based on real-time feedback. This paper investigates the application of ReAct-based agents in agricultural planning, focusing on gardening activity planning as a testbed and introduces the **Gardening Planner**, an agent that integrates reasoning capabilities with tools such as a Retrieval-Augmented Generation (RAG) system, a weather forecast API, and a dynamic web search tool. The agent leverages these components to create personalized, context-aware crop plans for the current season. Our agent can provide effective crop plans only 50% of the time. The findings reveal key limitations of ReAct-based agents in dynamic, real-world settings such as agriculture, emphasizing their reliance on retrieval and summarization over genuine planning. This study also contributes to the broader discourse on planning and reasoning in AI by highlighting the challenges of deploying agentic frameworks in practical applications.

Introduction

The ability to effectively plan is often associated with intelligence, particularly in humans, where planning enables individuals to achieve complex goals by anticipating future needs, adapting to dynamic environments, and making informed decisions based on available information. This skill is equally valuable in artificial intelligence (AI), where planning capabilities can significantly enhance the functionality of agents deployed in real-world tasks. In recent years, large language models (LLMs) have gained attention for their potential in various decision-making and reasoning tasks. (Giadikiaroglou et. al. 2024). Chain-of-Thought (CoT) (Wei et. al. 2022) prompting has demonstrated impressive versatility, successfully tackling both rule-based puzzles (deterministic and stochastic) and rule-less challenges including riddles, programming tasks, and commonsense reasoning. The Reactive Action (ReAct) framework further builds on LLMs capabilities by

combining reasoning traces with actionable steps in an iterative process. When integrated with CoT and Self-Consistency methods, ReAct demonstrates enhanced performance on complex tasks, achieving significant improvements in question-answering benchmarks like HotpotQA and Fever (Yao et. al. 2022).

Impressed by the performance of the ReAct agents in the domain of planning and reasoning tasks, we sought to test it -out on real world planning problem in the agriculture domain. In the context of agricultural planning, particularly gardening activity planning, these capabilities and limitations become especially relevant. The complex, dynamic nature of agricultural systems, influenced by numerous environmental variables and requiring real-time adaptation, provides an ideal testing ground for evaluating the practical utility of ReAct-based agents. Our findings suggest that the ReAct based agents can help in retrieval of information from multivarious sources and summarize it well to aid the task such as developing a crop plan for current season, but it was not able to develop a concrete plan consistently. This study will help understand the shortcomings of the agents in dynamic environment settings such as agriculture and will help to further to advance the efforts on planning and reasoning using agents in this domain.

Related Work

The emerging interest in leveraging Large Language Models (LLMs) for planning and reasoning domains is well-documented, with various frameworks questions on reasoning capabilities of LLMs. aiming to expand their applicability (Yao et al. 2022; Xu et al. 2023). However, a growing body of research raises reasonable Valmeekam, Marquez, and Kambhampati (2023) investigated that whether planning and reasoning abilities of LLMs can improve using self-critiquing setup and suggested that it further degrades in performance. Furthermore, in another study by Stechly, Marquez, and Kambhampati (2023) found that GPT-4 cannot self-critique itself and hence it is not reasoning effectively. Even in deterministic tasks like cipher

encoding, GPT-4's accuracy varies notably between high-probability (21%) and low-probability (11%) input sentences further supporting the pattern-matching hypothesis. (McCoy et. al. 2023). Research has shown that these models are more likely to change correct answers to incorrect ones than to fix their mistakes (Huang et. al. 2024). To sum up the argument Kambhampati (2024) suggests that LLMs may not stem from genuine cognitive reasoning or structured planning, instead these abilities are increasingly viewed as manifestations of "approximate retrieval". Finally, Kambhampati et. al. (2024) proposed a framework "LLM-Modulo" which demonstrates the how LLMs can help in planning domain. However, these studies do not explore the agentic frameworks from the lens of planning and reasoning specific tasks in a real-world scenario.

Architecture

The Gardening Planner is an agent based on ReAct framework designed to provide personalized crop plans and aids in decision-making processes. The architecture consists of two major components, Reasoner and Tools interconnected through the ReAct loop. Figure 1 demonstrates the implementation of Gardening Planner's ReAct Loop using LangGraph a composable AI workflow framework. (LangChain Inc. 2024)

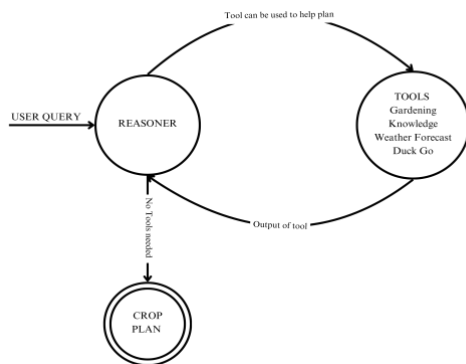


Figure 1: State diagram of Gardening Planner

Reasoner

The reasoner in the ReAct framework acts as a brain of the system. It enables the agent to analyze user inputs, formulate a logical plan, and execute tasks effectively through dynamic interactions with external tools. The Reasoner operates in a step-by-step manner, ensuring accurate and context-aware task completion. This part of the ReAct system usually incorporate an LLM. In the implementation of Gardening Planner, GPT-4o has been used with a system prompt: "You are an expert Gardener tasked with creating an optimal crop plan for current season." Based on the

output of the tool used, Reasoner i.e. the LLM decides whether to further call another tool to gather more information to curate an effective crop plan for the current season or it can create a plan based on all the previous outputs of the tools.

Tools

In the gardening planning agent, we have integrated three tools to enable it to develop a contextual aware and effective crop plan

Gardening Knowledge Tool

The Gardening Knowledge Tool is a Retrieval-Augmented Generation (RAG) based query system that serves as the backbone for developing comprehensive crop plans. At its core is an indexed Gardening Resources website that provides detailed guidance on garden preparation and seeding. (UC Master Gardener Program, 2024)

To build this tool, the process involved converting a collection of HTML files from the website into Markdown (MD) files, which allowed efficient chunking of data. These chunks were embedded using OpenAI's text-embedding-3-large embeddings, a model designed for generating high-quality vector representations of textual data. The embeddings were then stored in ChromaDB, a high-performance vector database optimized for scalable retrieval.

A retriever operates on top of ChromaDB, enabling the system to query dynamically and retrieve context-specific information. This includes details such as optimal planting schedules, soil preparation techniques, plant spacing requirements, and seed recommendations. Although the knowledge is specifically tailored to California's agricultural conditions, it is versatile enough to be applied across the United States. The use of ChromaDB ensures rapid query response times, making the tool highly effective for real-time applications. This tool forms the foundation for precise, actionable crop plans that adapt seamlessly to agricultural needs and seasonal variations.

Weather Forecast Tool

This functional tool enhances the agent's capabilities by enabling it to fetch a 7-day weather forecast for the garden's location using the zip code via the OpenWeatherMap API. By accessing real-time weather data, the tool provides the agent with essential insights into the dynamic environmental conditions of the specified location. This information is crucial for the agent to carefully recommend vegetables that are best suited to the current and upcoming weather patterns, ensuring optimal yield and a productive gardening season. The integration of weather forecasting into the system allows the agent to deliver context-aware and climate-sensitive crop planning, further improving the relevance and accuracy of its recommendations.

DuckDuckGo Search Tool

This tool is an essential component that empowers the agent to fetch additional relevant information tailored to the location of the garden and the best vegetables for the current season. By performing real-time searches, this tool complements the agent's existing knowledge base by retrieving up-to-date and location-specific insights. It can dynamically gather data such as regional gardening tips, seasonal vegetable recommendations, and localized planting techniques. sensitive advice to optimize gardening outcomes.

Evaluation and Analysis

To evaluate the reasoning and planning capabilities of the gardening planner, we have asked it to generate crop plan for ten different locations across the United States. We evaluated the agents based on effective use of tools, reasoning on tool output and correction in approach. The summary of the evaluation results is in Table 1.

Location	Effective use of Tools	Correction in approach
Fremont, CA	Yes	Not Required
Salt Lake City, UT	Yes	Not Required
Madison, WI	No	No
Reno, NV	No	No
Detroit, MI	No	No
Portland, OR	Yes	Not Required
Orlando, FL	No	No
Dallas, TX	Yes	Not Required
Cleveland, OH	Yes	Not Required
Lewiston, ME	No	No

Table 1: Summary of evaluation results of different cities in the United States

Based on the runs, we have identified two distinct patterns in the approach it took to develop a plan. In 1st pattern, agent tries to fetch weather information as well as best crop selection as well as how to plant it using gardening knowledge tool, this approach worked well when the location is in California. Although the gardening knowledge is broadly applicable, it contains detailed information about crops that thrive specifically in California's climate and soil nutrient conditions. Therefore, to develop an effective plan for locations outside of California, the agent must not only

retrieve relevant data and summarize but also reason critically about the retrieved content. It should recognize the necessity of first identifying the best crop for the location

Pattern 1: Simple retrieval and summarization without Duck Duck Go Search

Tool: get_7_day_forecast_by_zip

Parameter: zip code of the location

Output: Weather forecast of 7 days

Tool: gardening knowledge

Parameter: best crop to plant in <current month> in <location> for high yield

Output: Retrieves chunks related to how to plant best crop

Summarizes the content and comes with plan

using a search tool like DuckDuckGo. Subsequently, it can employ the gardening knowledge base to extract appropriate planting methods tailored to that crop.

Figure 2 depicts the retrieved chunk when the location was Madison, WI and shows how trivial it was to detect the content is tailored towards California.

Pattern 2: DuckDuckGo Search and then retrieval and summarization

Tool: duckduckgo_search

Parameter: best crops to plant in <location> in <month> for high yield

Output: Retrieves the potential crops

Tool: get_7_day_forecast_by_zip

Parameter: zip code of the location

Output: Weather forecast of 7 days

Tool: gardening knowledge

Parameter: how to prepare garden for planting <best crop>

Output: Retrieves best practices related to the crop
Summarizes the content and comes with plan

```

gardening_knowledge with query: best crops to plant in November
in Madison, Wisconsin

Output:

California has a unique vegetable growing climate, we can grow
many different types of vegetable crops throughout the year....

The California Master Gardener Handbook provides great
recommendations on when and when to plant by seed (row) or
young transplants (plants) and how much to plant for a family of
four:

| Vegetables | Recommended planting dates by location in
California | Crop type | Amount to plant (4-persons) |
| artichoke | North and north coast: Aug-Dec South coast: May-
July Interior Valley: July Desert Valley: Sept | cool | 3-4 plants |
| asparagus | North and north coast: Jan-March South coast: Jan-
Feb Interior Valley: Jan-Feb Desert Valley: Feb-April | cool | 30-40
plants |
| beets | North and north coast: Feb-Aug South coast: Jan-Sept
Interior Valley: Feb-April; Aug Desert Valley: Sept-Jan

```

Figure 2: Gardening Knowledge Tool and retrieves content related to California

In 2nd pattern, the agent tries to first search for the best crop in terms of yield using DuckDuckGo tool. After fetching that input, it uses weather tool and gardening tool to get information about the crop that was retrieved by using the search tool. This pattern allows the agent to plan efficiently for the any location. Figure 3 shows the sequence of the tool calls in the case of Salt Lake City, Utah.

```

Tool Calls:

duckduckgo_search with best crops to plant in Salt Lake City Utah in November for high yield

Output

Below we learn the Utah vegetable planting calendar, month-by-month Utah vegetable planting
schedule, the best vegetables to grow in Utah, and a seasonal vegetable planting guide for Utah
home gardens. ... Most people don't consider the best time to plant their crops, so they usually do
it all at once in May. ... Salt Lake City: Brigham City ... This planting information is for the Salt
Lake City, Utah region, Hardiness Zones 6b-7a. If you are in a different Zone of our region,
please check out our Planting Guides page to find an option that applies! Planting environments
in the Intermountain West vary greatly including Arid High Desert, Urban, and High Mountain
regions. The soil should never dry out, or the celery may become stringy or bolt (go to seed)
prematurely. It's best to start celery indoors in November and transplant out when they have two
sets of true leaves. Plant 6-8" apart in rows 24-36" apart and have a lot of patience. A relative of
the bulbing onion, leeks are also an ideal vegetable to plant in November. Leave a large crop in
the garden over winter and pull them as you need to add your cozy soups and stews. ' King
Richard ' is a great cold-hardy option. November is later than usual to plant leeks, particularly in
zones below 9. Lettuce is one of the most popular crops for hydroponic farming due to its rapid
growth and consistent market demand. The U.S. lettuce market is substantial, with Americans
consuming approximately 14 pounds of lettuce per capita annually. The demand for lettuce
remains steady year-round, making it a reliable crop for hydroponic growers.

get_7_day_forecast_by_zip with args zip_code: 84101

Output:

[{"date": "2024-11-18", "avg_temp": 38.49, "avg_humidity": 55.83}, {"date": "2024-11-19",
"avg_temp": 31.06, "avg_humidity": 65.75}, {"date": "2024-11-20", "avg_temp": 32.51,
"avg_humidity": 59.0}, {"date": "2024-11-21", "avg_temp": 42.86, "avg_humidity": 47.75},
{"date": "2024-11-22", "avg_temp": 47.67, "avg_humidity": 35.88}, {"date": "2024-11-23",
"avg_temp": 49.26, "avg_humidity": 32.5}]

gardening_knowledge with query: how to prepare garden for planting celery and leeks

Output:

# Early start on growing vegetables
## Starting Plants Indoors

Growing your own transplants indoors can extend your gardening season by several weeks,
reduce your gardening costs, and allow you to grow transplants of some of the hard-to-find
varieties of crops. In general, seed should be sown from 6 to 8 weeks before the date that you wish
to set the plants in the garden.Successful plant starting requires:
suitable soil mix, suitable containers, proper moisture, temperature and aeration, adequate light,
conditioning plants to the outdoors

To get a head start on a vegetable garden, you can also start seed in a hotbed or cold frame in
order to transplant young plants early in the season.....

### Row covers, hot caps, and similar devices

Another method of promoting early plant growth and production is to cover seeds or young
plants in the field with polyethylene row covers, paper hot caps, discarded plastic gallon water or
milk containers with bottoms removed, or similar heat-trapping devices. These coverings keep
daytime air temperature higher but have a limited effect on soil or night temperatures. If you use
them, you must harden the plants before completely removing the covering.....

```

Figure 3: Sequence of tool calls with their outputs in the case of Salt Lake City, Utah

Figure 4 depicts the plan developed the agent for Salt Lake City, Utah using the output from the tools. By analyzing these two patterns, a crucial insight emerges that the agent is not genuinely reasoning with the outputs of the tools but rather summarizing them effectively to draft a plan. When the output of a tool fails to provide a proper solution, the agent demonstrates a lack of self-correction or significant adjustments in the planning process. This limitation hinders the agent's ability to dynamically adapt and refine its approach, impacting its effectiveness in completing tasks.

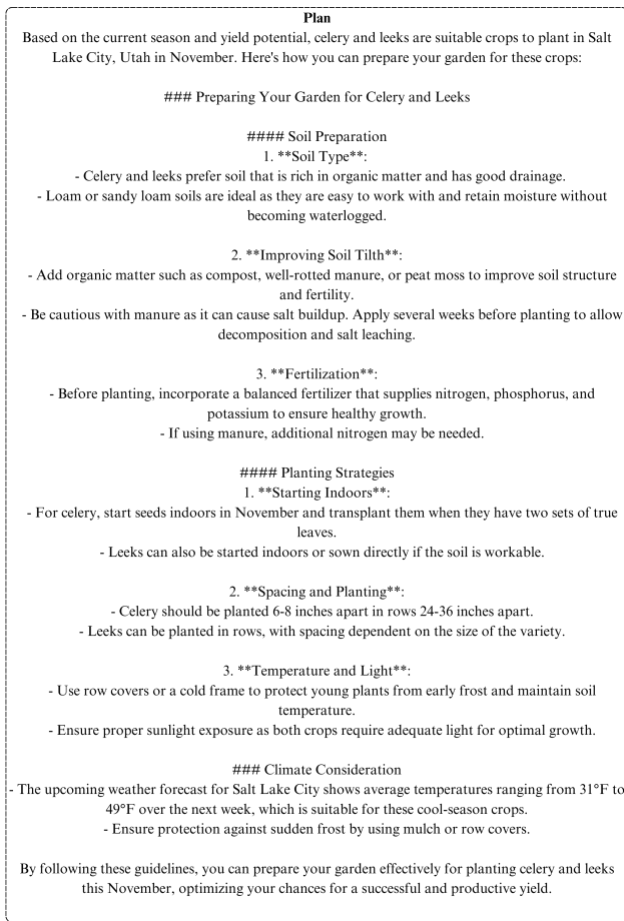


Figure 4: Plan developed by Gardening Planner for Salt Lake City, Utah

Conclusion and Future Work

Our study explored the use of agents in dynamic environments such as agriculture, specifically for developing crop plans. While the agent did not exhibit genuine reasoning on the outputs of the tools yet, but it demonstrated advanced capabilities in efficiently searching for relevant content and summarizing it effectively. This aligns with Kambhampati, Valmeekam, and Guan (2024) observation that LLMs primarily function as powerful summarizers, a characteristic evident in agentic frameworks as well. Furthermore, the findings emphasize the necessity of keeping humans in the loop for critical tasks, underscoring the current limitations of agents in autonomous decision-making and reasoning. For the future work, we intend to run the similar experiments on other agentic frameworks such as ReWOO. Additionally, we aim to further research methods to enable true reasoning on the outputs of tools, leveraging reinforcement learning techniques to enhance the agent's decision-making and reasoning capabilities.

References

- Giadikiaroglou, P.; Lymperaioi M.; Filandrianos G.; Stamou G.; 2024. Puzzle solving using reasoning of large language models: A survey. arXiv preprint arXiv:2402.11291.
- Wei, J.; Wang X.; Schuurmans D.; Bosma M.; Ichter B.; Xia F.; Chi E.; Le Q.; Zhou D.; 2022. Chain-of-Thought Prompting Elicits Reason- ing in Large Language Models. arXiv:2201.11903.
- Yao, S.; Zhao, J.; Yu, D.; Du, N.; Shafran I.; Narasimhan K.; Cao Y.; 2023. ReAct: Synergizing Reasoning and Acting in Language Models. In *Eleventh International Conference on Learning Representations*, ICLR '23.
- Xu, B.; Peng, Z.; Lei, B.; Mukherjee, S.; Liu Y.; Xu D.; 2023. ReWOO: Decoupling reasoning from observations for efficient augmented language models. arXiv preprint arXiv:2305.18323.
- Valmeekam, K.; Marquez, M.; Kambhampati, S; 2023. Can large language models really improve by self-critiquing their own plans? arXiv preprint arXiv:2310.08118.
- Stechly, K.; Marquez, M.; Kambhampati, S; 2023. Gpt-4 doesn't know it's wrong: An analysis of iterative prompting for reasoning problems. arXiv preprint arXiv:2310.12397.
- McCoy, R. T.; Yao, S.; Friedman, D.; Hardy, M.; Griffiths T. L.; 2023. Embers of autoregression: Understanding large language models through the problem they are trained to solve. arXiv preprint arXiv:2309.13638.
- Huang, J.; Chen, X.; Mishra, S.; Zheng, H. S.; Yu A. W.; Song X.; Zhou D.; 2024 Large Language Models Cannot Self-Correct Reasoning Yet. In *The Twelfth International Conference on Learning Representations*, ICLR '24.
- Kambhampati, S. 2024. Can large language models reason and plan? *Annals of the New York Academy of Sciences*, 1534(1): 15–18.
- Kambhampati, S.; Valmeekam, K.; Guan, L.; Verma, M.; Stechly K.; Bhambri S.; Saldyt L.; Murthy A.; 2024. Position: LLMs Can't Plan, But Can Help Planning in LLM-Modulo Frameworks. In *Forty-first International Conference on Machine Learning*. ICML '24.
- LangChain Inc. 2024. LangGraph: Build Resilient Language Agents as Graphs. <https://github.com/langchain-ai/langgraph>
- UC Master Gardener Program. 2024. Gardening Advice and Resources. <https://mg.ucanr.edu/Gardening/>. Accessed: November 16, 2024.
- Kambhampati, S.; Valmeekam, K.; and Guan, L; 2024. On the Role of Large Language Models in Planning. Tutorial presented at The 38th Annual AAAI Conference on Artificial Intelligence, Vancouver. <https://aaai.org/aaai-conference/aaai-24-tutorial-and-lab-list/#th20>.

