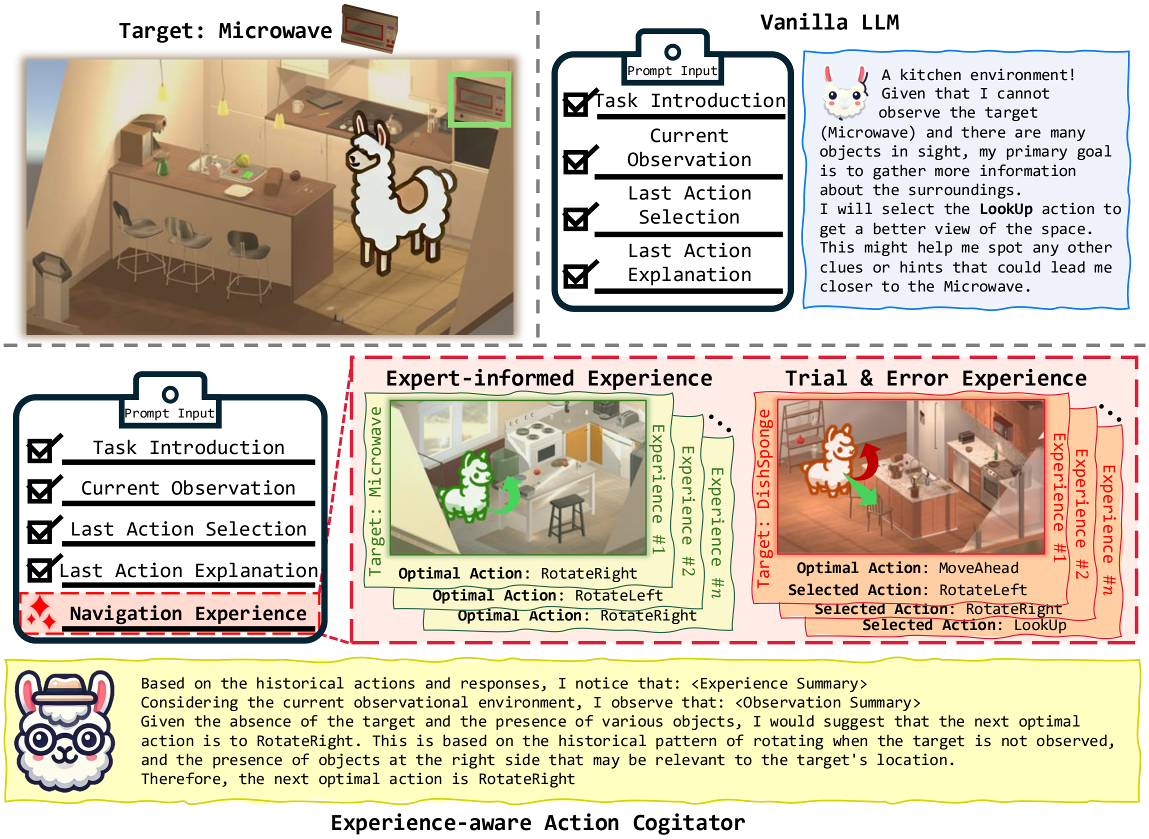
2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Déjà Vu: Unlocking Transparent Action Reasoning for Object-Goal Navigation via Large Language Models

Heming Du[1], Sen Wang[1], Li Xue[1], Yu Xin[2]*
[1]The University of Queensland [2]The University of Adelaide *Corresponding author



A kitchen environment! Given that I cannot observe the target (Microwave) and there are many objects in sight, my primary goal is to gather more information about the surroundings. I will select the **LookUp** action to get a better view of the space. This might help me spot any other clues or hints that could lead me closer to the Microwave.

Based on the historical actions and responses, I notice that: <Experience Summary>
Considering the current observational environment, I observe that: <Observation Summary>
Given the absence of the target and the presence of various objects, I would suggest that the next optimal action is to RotateRight. This is based on the historical pattern of rotating when the target is not observed, and the presence of objects at the right side that may be relevant to the target's location.
Therefore, the next optimal action is RotateRight

Fig. 1: Our table tennis robot playing against a professional coach. The green dots show the trajectory of the ball during the rally. The table tennis robot is a 6 DoF ABB 1100 arm mounted on top of two Festo linear gantries, enabling motion in the 2d plane. The x gantry, which moves side to side across the table, is 4m long and the y gantry, which moves towards and away from the table, is 2m long. A 3d printed paddle handle and paddle with short pips rubber is attached to the arm.

***Abstract*— The remarkable interaction and reasoning capabilities of Large Language Models (LLMs) make them promising in collaborative Embodied AI tasks, particularly for Object-goal Navigation (ObjNav) tasks that require both decision-making and transparent explanation. However, existing work mainly uses LLMs as proxy target indicators, leaving the role of direct action decision-making to other components. This separation causes non-transparent action decisions and extra adaptation requirements. This observation prompts us to reconsider their role: Can LLMs be transformed into the central "brain" of agents, directly outputting action choices and explaining their reasoning? In pursuit of this inquiry, we decouple perception from action reasoning to focus specifically on the feasibility of deploying LLMs as navigation policies. We introduce the Experience-aware Action Cogitator (ExAC) that integrates two kinds of experience, *i.e.*, expert-informed experience and trial & error experience, into prompts. Inspired by David Hume's philosophical principles that knowledge is acquired through reflective experience, these experiences are designed for two critical questions: (i) "What action should be selected as the best option?" and (ii) "What actions have been tried but proven suboptimal?" By analyzing and reflecting on these two types of experience, we show that LLMs can reason navigation actions in unseen environments effectively without costly fine-tuning. Experiments on the widely-adopted iTHOR yield significant improvements in ObjNav performance. These compelling results validate the feasibility of our ExAC. Compared to vanilla LLMs, ExAC nearly doubles both the Success Rate and the Success weighted by Path Length, reaching peak values of 73.93% and 48.35% in unseen scenes, respectively.**

I. INTRODUCTION

Large Language Models (LLMs) exhibit considerable capabilities in complex reasoning and interactive tasks [1]–[3].

979-8-3315-8160-2/26/$31.00 ©2026 IEEE

Their skills in understanding and responding make LLMs naturally suited to tackle various Embodied AI tasks [2], [4], [5]. This suitability is particularly evident in Object-goal Navigation (ObjNav), a task that demands both robust decision-making and clear, transparent explanations. In ObjNav, the goal is to steer an agent to a specific object without relying on explicit path guidance. Compared with LLMs, current reinforcement learning based methods [6]–[8] focus on outputting actions without clear explanations. This lack of transparent reasons is not helpful for collaborating with humans. Consequently, a compelling research question arises: Can LLMs function effectively as agents in ObjNav?

However, most existing approaches utilize LLMs merely as a feature extractor [9] or as proxy target indicators [10], [11] rather than as autonomous agents that directly select actions. LLMs are used to break down complex tasks into simpler sub-goals. For example, they map navigational frontiers and use external modules to output the actual actions. This practice restricts the potential of LLMs to operate as general-purpose Embodied AI agents in the Embodied AI domain. When LLMs are limited to outputting only navigation-related information, they lose the ability to provide other task actions. Thus, a critical research question emerges regarding the practice of LLMs in navigation: *Can LLMs work as independent agents to make action decisions and provide explanations for their decisions in a manner akin to human cognition?*

In this paper, we propose an Experience-aware Action Cogitator (ExAC) to explore the feasibility of utilizing LLMs in directly making and explaining navigation decisions. Drawing inspiration from David Hume's empiricist philosophy, that knowledge is acquired through reflective experience, ExAC integrates two distinct types of experience: expert-informed experience and trial & error experience. These experiences are designed to answer two fundamental questions in navigation: (i) "What action should be selected as the best option?" and (ii) "What actions have been tried but proven suboptimal?"

Specifically, expert-informed experience provides the agent with prior experience and expert demonstrations, guiding it to choose optimal actions. For example, if the target object is on the left, expert-informed experience indicates that the LLM selects the appropriate action, such as `RotateLeft`. Concurrently, trial & error experience exposes the agent to past mistakes, helping it recognize and avoid ineffective actions. By reflecting on these two forms of experience, ExAC equips LLMs to reason through navigation tasks in novel environments without the need for extensive training episodes.

Drawing inspiration from the phenomenon of déjà vu, where elements from past experiences reappear in new contexts, we retrieve relevant experiences that closely resemble the current observations. Note that we convert the observation into text description for retrieval to avoid the effects of image textures and focus on the layout of observed instances. After this retrieval, we acquire navigation experiences that capture both optimal actions and past mistakes. By combining these retrieved experiences with the present observation and the explanation of the previous action, ExAC generates both navigation decisions and their explanations, as shown in Figure 1.

Experiments on the widely-adopted iTHOR [12] environment yield significant improvements in ObjNav performance, validating the feasibility of our ExAC. In unseen scenes, ExAC nearly doubles both the Success Rate (73.93%) and the Success weighted by Path Length (48.35%) compared to vanilla LLMs. To demonstrate the superiority of ExAC, we conduct experiments using various LLMs[1], *i.e.*, LLaMA-3 [1], Qwen2.5 [13], and Mistral [14]. These results demonstrate that ExAC not only enables LLMs to provide transparent action explanations, but also significantly enhances their navigation performance. Overall, our major contributions are summarized as follows:

- We explore the feasibility of using LLMs to directly select actions and provide explanations in ObjNav and introduce the Experience-aware Action Cogitator (ExAC).
- We design two types of navigation experiences: expert-informed experience and trial & error experience. Agents reflect on these experiences to reason optimal actions and improve navigation decisions.

II. RELATED WORKS

A. Object-goal Navigation

Navigation has progressed from full-map pipelines, mapping, localization, and planning [15], [16], to map construction in unseen scenes via SLAM. With deep learning [17], [18], to policies that map visual input directly to actions [19], [20], including VLN, point-goal, and audio-visual variants driven by language, goal cues, and sound [21], [22].

In ObjNav [23], [24], methods typically couple perception with a navigation policy. Unlike memory-based designs [24], [25], we learn from other episodes rather than within-episode history and instantiate the policy with an LLM that, given textualized images, both selects actions and explains decisions, outputting action commands in natural language.

Both our method and RL aim for human-like decision making, but RL requires millions of trial-and-error updates, while we use LLMs to reflect on limited experiences. Unlike PR2L and TWOSOME [26], [27], our approach needs no extra parameters or task-specific training, adapting simply by replacing relevant experiences. In contrast, graph-based planners such as Dijkstra [28] depend on complete maps and fixed edge weights, which our method does not require.

B. Large Language Models in Embodied AI

Recent LLMs such as LLaMA-3 [1], Qwen2.5 [13], and Mistral [14] have broadened embodied AI. Reasoning methods like CoT [29] and ToT [30] improve logical inference, while explanatory approaches [31], [32] integrate interpretability. Two paradigms link LLMs with embodied tasks:

---

[1] Due to restrictions, we are not allowed to use or deploy any products from Deepseek [3], which is disappointing.

one generates executable codes or indirect commands [33], [34], while the other directly trains LLMs for control [35].

In ObjNav, prior work employs LLMs as proxy target indicators with separate sub-goal execution modules [36], [37]. Our approach instead exploits LLM reasoning to directly output navigation actions and explanations without fine-tuning or extra modules, extending to unseen environments [38]. Unlike VLN systems such as NavGPT [39], MapGPT [40], DiscussNav [41], and RAGNav [42], which decompose trajectory-based instructions with proprietary LLMs, our method focuses on ObjNav, where only the target category (e.g., "Microwave") is given, and open-source LLMs directly reason actions from observations while providing human-like explanations.

III. EXPERIENCE-AWARE ACTION COGITATOR

To explore the feasibility of utilizing LLMs in making and explaining navigation decisions, we design the Experience-aware Action Cogitator (ExAC). As demonstrated in Figure 2, the LLM acquires navigation knowledge from historical navigation experiences retrieved based on current observations. With the acquired navigation knowledge and the previous navigation action, the LLM can explain and select effective navigation actions.

A. Task Definition and Setup

As an exploratory effort, we aim to employ LLMs for deciding and explaining navigation actions in ObjNav. We set up the task based on previous works [6], [24], [43], [44]. An agent is placed in a grid-based environment and cannot access any external knowledge regarding the entire environment, including the topological map and 3D meshes. The agent moves through the grid using six actions: `MoveAhead`, `RotateLeft`, `RotateRight`, `LookUp`, `LookDown`, and `Done`, with fixed step sizes and angles.

At the beginning of each episode, the agent is assigned a random target class and starts from a random location in a random room to ensure the uniqueness of each episode. Unlike the previous work, we adopt more than 80 pickupable and moveable instance categories from iTHOR [12]. A successful episode is defined as one in which the agent selects the termination action, `Done`, when the distance between the agent and the target is below a set threshold (*i.e.*, 1.5 meters), within a maximum of 200 steps, and with the target in its field of view. If the termination action is executed at any other time, the agent fails, and the episode ends.

Following prior ObjNav settings [6], [43], agents receive egocentric RGB inputs, which we convert into text for LLMs. Instead of object detection or visual grounding [6], [24], we provide simple spatial descriptions: if the target occupies more than 100 pixels in a 300×300 image, it is considered observed and localized within a 3×3 grid cell. A binary indicator denotes whether it is within a predefined distance. Other objects are also reported with coarse left/center/right labels using the same pixel rule.

B. Navigation Stage Selection

We divide navigation into two phases: searching, where the agent explores unseen areas to locate the target, and approaching, where it moves efficiently toward the target once visible. For example, during searching, the agent should expand its field of view (*e.g.*, choosing `RotateRight` over `RotateLeft`), while in approaching it should act directly toward the target (*e.g.*, `MoveAhead` or `RotateLeft`) instead of persisting with irrelevant rotations. The phase is determined from current observations and recent action explanations.

C. Navigation Experience Collection

To guide action selection and explanation, we introduce two types of navigation experiences—expert-informed and trial & error—following Hume's view that knowledge arises from reflection (Figure 3).

1) Expert-informed Experience: Expert-informed experience provides demonstrations of optimal behavior. First, we design 10 rules linking target locations in a 3×3 grid to navigation actions, with a `Done` action when the target is near (Figure 3a). Second, we collect experiences from optimal paths by running Dijkstra's algorithm [28] in random scenes and recording the agent's observations and corresponding actions (Figure 3b).

2) Trial & Error Experience: To complement expert data, we record suboptimal decisions made when agents act with expert-informed guidance. At each step, we log the chosen action, the optimal action, and the observation (Figure 3c). This reflective feedback, similar to human learning from mistakes, equips LLMs with both successful strategies and knowledge of actions to avoid.

D. Experience-guided Navigation

After obtaining expert-informed experience and trial & error experience, we integrate these experiences into prompts using in-context learning [45], [46] to help LLMs acquire navigation-related knowledge during the navigation episode. When humans encounter new situations, they often experience a form of déjà vu, where the brain makes unexpected connections between new experiences and vaguely familiar memories. This process allows them to draw on past experiences that closely mirror their current context for similar challenges. Motivated by this, we enable the LLM to retrieve the most similar experiences based on the current observation [47]. These similar experiences guide both the selection and explanation of navigation actions.

Considering the differences in the navigation stages, we divide navigation experiences into two stages: searching and approaching. Depending on the selected navigation stage, we retrieve experiences from the corresponding set, as demonstrated in Figure 2. In our implementation, we retrieve ten navigation experiences from expert-informed experience based on the optimal path, and another ten from trial & error experience. When the agent is in the approaching stage, we retrieve four additional experiences from the expert-informed set based on prior knowledge.

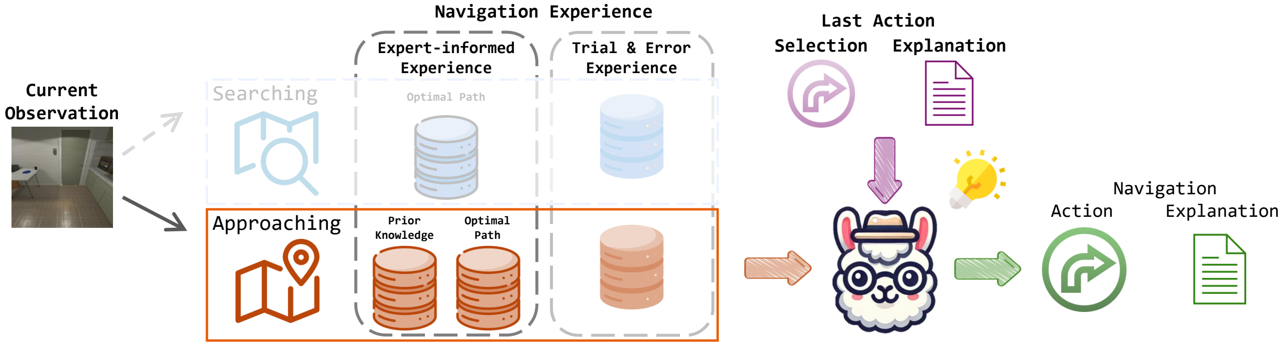
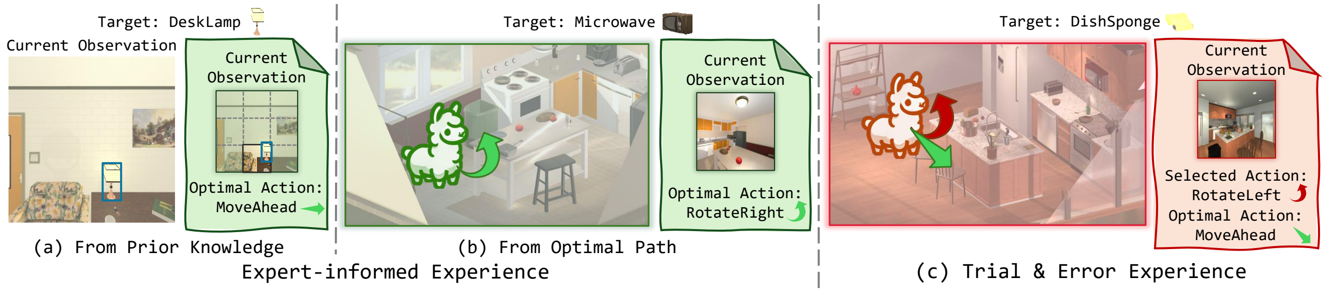


Fig. 2: **Pipeline of Experience-aware Action Cogitator (ExAC).** At each step of navigation, given the observation, the LLM first decides whether the step involves searching for the target or moving closer to it. After that the LLM retrieves the most similar experiences from the historical experiences. The LLM then analyzes the retrieved navigation experiences, the current observation, and the last action and interpretation to provide a transparent selection for the current navigation action.



(a) From Prior Knowledge (b) From Optimal Path

Expert-informed Experience

(c) Trial & Error Experience

Fig. 3: **Illustration of Navigation Experiences.** (a) shows an expert-informed experience derived from prior knowledge. The optimal action `MoveAhead` is chosen because the target, a desk lamp, is directly ahead in the middle-bottom grid of the observation. (b) demonstrates an expert-informed experience from the optimal path, where the action `RotateRight` is selected as the most optimal route to the target, a microwave. (c) presents a trial & error experience where the LLM navigator, only with expert-informed experience, chooses the action `RotateLeft` which deviates from the optimal action `MoveAhead` towards the target, *i.e.*, dish sponge.

Once we obtain the retrieved experiences, we combine the current observation, the retrieved navigation experience, and the explanation of the last navigation action into an experience-enhanced prompt. For expert-informed experiences based on optimal paths and trial & error experiences, the prompt includes the previous navigation action, the observation, and the optimal action. Additionally, when presenting the trial & error experience, the prompt includes previously chosen suboptimal navigation actions. During the approaching phase, the prompt further includes the current observation and the corresponding optimal action to present the retrieved experience based on prior knowledge. The detailed prompt is provided in the supplementary materials. The prompt is then sent to the LLMs to generate the next navigation action along with an explanation.

IV. EXPERIMENTS

To validate the effectiveness and efficiency of ExAC, we conduct experiments on the widely-adopted artificial environment, *i.e.*, iTHOR [12]. These experiments serve as a proof-of-concept that demonstrates the efficacy of ExAC.

A. Protocols and Experimental details

a) Dataset.: The iTHOR [12] environment includes 120 scenes in four types: kitchen, living room, bedroom, and bathroom. Each type of scene consists of 30 distinct rooms arranged with different furniture items and layouts. Following previous work, we utilize the first 20 out of 30 rooms per scene type to construct the navigation experience, while the final five rooms are used for testing. We evaluate ExAC on 1k random episodes, with each episode starting from a random position and a randomly assigned target class.

Evaluation. We evaluate performance using the success rate and Success Weighted by Path Length (SPL). The success rate measures trajectory effectiveness and is defined as $\frac{1}{N}\sum_{n=0}^{N} S_n$, where N is the total number of episodes and S_n is a binary indicator for the success of the n-th episode. SPL assesses model efficiency through the formula $\frac{1}{N}\sum_{n=0}^{N}\frac{L_{opt}}{\max(L_n,L_{opt})}$, where L_n denotes the length of the n-th episode and L_{opt} represents its optimal path length.

Implementation Details. We perform about 16k episodes across 80 different scenes, gathering 200k steps of expert-informed experience based on the optimal path. Additionally, using the LLaMA-3-8B model along with expert-informed experience, we conduct 3k episodes in the same 80 scenes, obtaining 32k steps of trial & error experience. After obtaining these experience, we store the experience in a vector database for retrieval. During testing, we run 50 episodes per scene, totaling 1k episodes. Each agent is run three times, and we report the average and standard deviation. The testing

TABLE I: Comparison of ExAC with various LLMs. We report the average Success rate (%) and SPL in iTHOR [12] as well as their variances in subscripts by repeating experiments three times. $L > 5$ means episodes need more than 5 steps.

LLM	Vanilla LLM				ExAC (Ours)			
	All		$L > 5$		All		$L > 5$	
	SR (%)	SPL	SR (%)	SPL	SR (%)	SPL	SR (%)	SPL
LLaMA-3-8B [1]	16.80 $_{(\pm 0.92)}$	4.32 $_{(\pm 0.33)}$	13.02 $_{(\pm 1.25)}$	3.50 $_{(\pm 0.36)}$	67.77 $_{(\pm 0.80)}$	44.17 $_{(\pm 2.08)}$	50.56 $_{(\pm 0.99)}$	24.10 $_{(\pm 0.69)}$
LLaMA-3.1-8B [1]	35.47 $_{(\pm 1.24)}$	7.31 $_{(\pm 0.25)}$	25.57 $_{(\pm 1.43)}$	5.92 $_{(\pm 0.43)}$	69.65 $_{(\pm 1.95)}$	35.47 $_{(\pm 1.41)}$	54.17 $_{(\pm 1.85)}$	16.99 $_{(\pm 0.11)}$
LLaMA-3.2-3B [1]	7.57 $_{(\pm 0.43)}$	0.85 $_{(\pm 0.05)}$	4.25 $_{(\pm 0.62)}$	0.60 $_{(\pm 0.22)}$	46.53 $_{(\pm 0.32)}$	18.06 $_{(\pm 0.13)}$	30.76 $_{(\pm 0.09)}$	5.80 $_{(\pm 0.23)}$
Mistral-7B [14]	19.92 $_{(\pm 1.95)}$	4.62 $_{(\pm 1.41)}$	16.53 $_{(\pm 1.85)}$	5.01 $_{(\pm 0.11)}$	51.76 $_{(\pm 0.84)}$	23.02 $_{(\pm 1.78)}$	33.53 $_{(\pm 1.02)}$	7.40 $_{(\pm 1.62)}$
Mistral-NeMo-12B [14]	35.56 $_{(\pm 0.67)}$	5.84 $_{(\pm 0.08)}$	22.58 $_{(\pm 0.22)}$	3.11 $_{(\pm 0.10)}$	67.22 $_{(\pm 2.59)}$	41.67 $_{(\pm 2.33)}$	49.35 $_{(\pm 3.05)}$	20.81 $_{(\pm 2.16)}$
Mistral-Small-22B [14]	50.78 $_{(\pm 1.22)}$	15.49 $_{(\pm 1.73)}$	40.64 $_{(\pm 1.75)}$	12.71 $_{(\pm 2.00)}$	71.40 $_{(\pm 0.30)}$	48.22 $_{(\pm 0.59)}$	54.90 $_{(\pm 0.51)}$	26.00 $_{(\pm 0.31)}$
Mixtral-8x7B [14]	56.63 $_{(\pm 3.88)}$	23.31 $_{(\pm 2.84)}$	43.23 $_{(\pm 0.75)}$	17.23 $_{(\pm 0.07)}$	63.45 $_{(\pm 1.45)}$	39.10 $_{(\pm 0.24)}$	42.84 $_{(\pm 1.29)}$	16.88 $_{(\pm 0.42)}$
Qwen2.5-7B [13]	38.34 $_{(\pm 0.56)}$	11.95 $_{(\pm 0.56)}$	34.26 $_{(\pm 1.73)}$	10.25 $_{(\pm 0.13)}$	58.55 $_{(\pm 1.75)}$	31.72 $_{(\pm 1.09)}$	41.54 $_{(\pm 2.24)}$	13.97 $_{(\pm 1.48)}$
Qwen2.5-14B [13]	54.00 $_{(\pm 4.00)}$	22.41 $_{(\pm 1.96)}$	40.43 $_{(\pm 5.88)}$	12.66 $_{(\pm 2.37)}$	70.33 $_{(\pm 3.78)}$	**48.35** $_{(\pm 2.20)}$	52.50 $_{(\pm 6.68)}$	27.36 $_{(\pm 3.68)}$

time ranges from 4 to 8 hours, depending on the LLM used. All experiments are conducted on a workstation equipped with 4 RTX 3090 GPUs, and the LLMs were operated using the Ollama platform.

B. Results and Analysis

1) Baselines: We use the following open-source LLMs to demonstrate the effectiveness of ExAC: **LLaMA-3** [1] includes LLaMA-3-8B, LLaMA-3.1-8B, and LLaMA-3.2-3B. LLaMA-3 is known for its enhanced understanding capabilities. **Mistral** [14] includes Mistral-7B, Mistral-Nemo-12B, Mistral-Small-22B, and Mixtral-8x7B. Notably, Mixtral-8x7B is mixture-of-experts (MoE) models. **Qwen2.5** [13] includes Qwen2.5-7B and Qwen2.5-14B. Qwen2.5 specializes in coding and mathematics tasks. We use the same prompts with the Vanilla LLM as with ExAC, merely removing parts related to the navigation experience.

2) Quantitative Results: As shown in Table I, ExAC consistently improves Success Rate (SR) and Success weighted by Path Length (SPL) across diverse LLMs. For instance, LLaMA-3.1-8B's SR increases from 35.47% to 69.65%, while Mistral-Small-22B rises from 50.78% to 71.40%. Even in harder cases with $L > 5$, substantial gains are observed, such as Qwen2.5-14B's SPL improving from 12.66 to 27.36.

LLaMA-3 models start with lower baseline performance as vanilla agents but benefit greatly from ExAC: SR of LLaMA-3-8B jumps from 16.80% to 67.77%, and from 13.02% to 50.56% for $L > 5$. This improvement suggests that although LLaMA-3 may lack navigation-oriented training data, introducing navigation experiences via ExAC significantly boosts reasoning and decision-making. Beyond weaker baselines, ExAC also strengthens stronger models. For example, Mistral-7B's SR improves from 16.53% to 33.53% for $L > 5$, and Qwen2.5-14B achieves 52.50% SR under the same condition. These results confirm the robustness of ExAC in enhancing both weaker and stronger LLMs across navigation tasks.

3) Case Study: Figure 4 illustrates how ExAC integrates expert experience, trial-and-error feedback, current observations, and recent actions to guide navigation. In the shown case, the LLM infers that the Microwave is absent but unexplored space exists on the right, leading it to choose RotateRight. This highlights the model's ability to summarize context, adapt decisions, and leverage experience for effective navigation.

TABLE II: Ablation study on different components of navigation experiences in iTHOR [12] environment utilizing ExAC with LLaMA-3-8B.

Method	All		$L > 5$	
	SR (%)	SPL	SR (%)	SPL
Vanilla Agent	16.80	4.32	13.02	3.50
ExAC w/o expert-informed experience	41.50	24.65	24.64	10.54
ExAC w/o error-drive trial experience	60.86	38.81	39.90	16.80
ExAC w/o selected action	62.65	37.89	42.23	17.14
ExAC	67.77	44.17	50.56	24.10

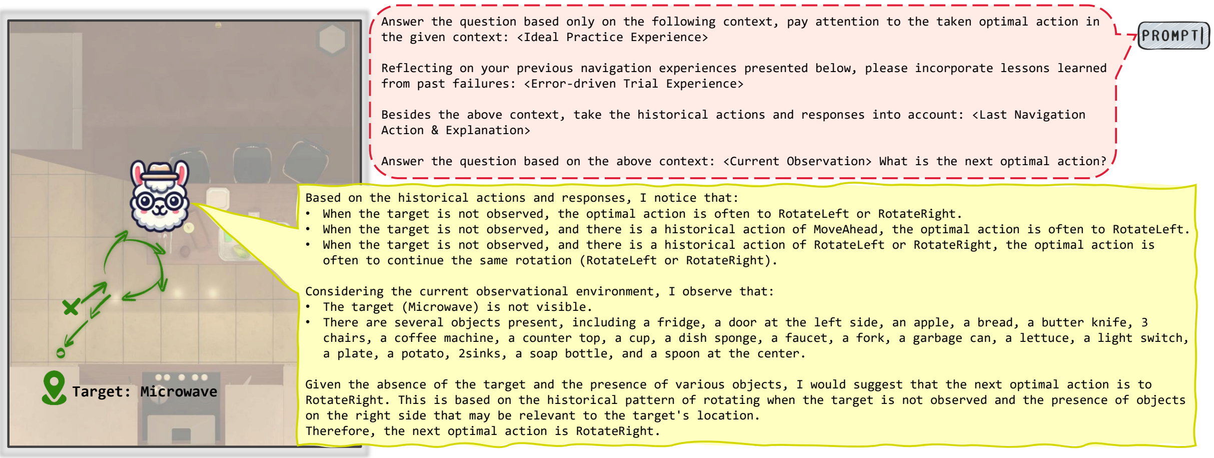
C. Ablation Study

Table II shows the impact of expert-informed and trial-and-error experience on LLaMA-3-8B. Without them, the vanilla agent performs poorly (SR 16.80%, SPL 4.32), while the full ExAC reaches 67.77% SR and 44.17 SPL. Removing expert-informed experience lowers performance most notably (SR 41.50%, SPL 24.65), especially for $L > 5$, while omitting trial-and-error experience also reduces performance (SR 60.86%, SPL 38.81), confirming the importance of both.

Further, excluding selected actions from trial-and-error prompts (ExAC w/o selected action) causes a moderate drop, showing that explicit feedback on suboptimal actions is more valuable than general error exposure. Overall, the ablation demonstrates that both expert-informed and error-driven experiences are crucial for robust LLM navigation in unseen environments.

V. DISCUSSIONS AND INSIGHTS

Comparison with Task-Specific Reinforcement Learning Networks. Our method allows LLMs to perform tasks using related experience without extensive training data required to train task-specific reinforcement learning networks. More critically, the incorporation of LLMs grants agents the capability to interpret their action decisions, facilitating insights



Answer the question based only on the following context, pay attention to the taken optimal action in the given context: <Ideal Practice Experience>

Reflecting on your previous navigation experiences presented below, please incorporate lessons learned from past failures: <Error-driven Trial Experience>

Besides the above context, take the historical actions and responses into account: <Last Navigation Action & Explanation>

Answer the question based on the above context: <Current Observation> What is the next optimal action?

Based on the historical actions and responses, I notice that:
- When the target is not observed, the optimal action is often to RotateLeft or RotateRight.
- When the target is not observed, and there is a historical action of MoveAhead, the optimal action is often to RotateLeft.
- When the target is not observed, and there is a historical action of RotateLeft or RotateRight, the optimal action is often to continue the same rotation (RotateLeft or RotateRight).

Considering the current observational environment, I observe that:
- The target (Microwave) is not visible.
- There are several objects present, including a fridge, a door at the left side, an apple, a bread, a butter knife, 3 chairs, a coffee machine, a counter top, a cup, a dish sponge, a faucet, a fork, a garbage can, a lettuce, a light switch, a plate, a potato, 2sinks, a soap bottle, and a spoon at the center.

Given the absence of the target and the presence of various objects, I would suggest that the next optimal action is to RotateRight. This is based on the historical pattern of rotating when the target is not observed and the presence of objects on the right side that may be relevant to the target's location.
Therefore, the next optimal action is RotateRight.

Fig. 4: **Demonstration of Navigation Explanation.** The left image shows the navigation process: the green cross is the start, the arrow the agent's path, the circle the stop, and the pin the target (Microwave). Our method reaches the target in seven steps. The right image presents the prompt template and explanation for the third-step action, where ExAC combines experience and observation to select the move, as also shown in Figure 1.

into the reasons behind their choices. In future work, we plan to enable these LLM-driven agents to consider unique demands or personalized settings through dialogues or tailored profiles in their task execution.

Comparison with Existing Methods of Using LLMs in Embodied AI Tasks. Unlike existing paradigms using LLMs in Embodied AI tasks, our approach leverages the inherent understanding and reasoning capabilities of LLMs to perform and interpret specific tasks without the need for fine-tuning. By incorporating task-relevant experience directly into prompts, we significantly reduce the resource consumption associated with fine-tuning LLMs and alleviate the need for large-scale datasets specific to tasks. Consequently, it becomes feasible to deploy a universal LLM across diverse tasks rather than dedicating a unique LLM to each. In future work, we will further investigate the impact of experience on the ability of LLMs to execute and interpret specific tasks.

Advantages of Language Input in Navigation Tasks. The use of natural language as an intermediate representation offers multiple benefits in embodied AI. Converting visual observations into textual input abstracts low-level image details while preserving high-level semantic content. This approach facilitates the integration of advanced perception methods, as long as their outputs can be structured to fit a predefined template. Moreover, leveraging in-context learning, language-based input enhances the transparency and interpretability of navigation decisions by allowing the system to clearly explain its actions in human-understandable terms. In addition, this interpretable intermediate variable enables easier monitoring and debugging within complex embodied AI agent systems.

Rationale for Excluding Visual Language Models (VLMs). Although VLMs can process images and text together, we use LLMs instead. This choice allows us to focus on navigation strategy research without the interference of visual perception. By converting images into text, we get a clear and semantic description of the environment. Using textual descriptions also makes it easier to retrieve experiences that are semantically similar to the current state. This is similar to how humans recall abstract ideas rather than detailed pictures when thinking about past experiences. As discussed above, language input brings benefits, such as improved interpretability and simpler integration of advanced perception methods.

Limitations. Our analysis does not yet delve into the detailed examination of individual experiences within our navigation experiences, an aspect we intend to explore in depth. Additionally, our current experiments are validation studies conducted in controlled settings, which may not capture the complexity of real-world navigation. In future work, we will scale our approach to more realistic scenarios. Finally, the nearly endless combinations of actions in navigation make it hard for us to evaluate the explanations quantitatively at this stage, so we can only show them qualitatively. Addressing these issues will help us compare the explanation capabilities of different models more accurately.

VI. CONCLUSION

In this paper, we introduced the Experience-aware Action Cogitator (ExAC) to use Large Language Models (LLMs) as the central decision maker in object-goal navigation tasks. We decouple visual perception from navigation reasoning by converting images to text. We integrate two types of navigation experience, expert-informed experience and trial & error experience, into our in-context learning framework. Inspired by David Hume's idea that knowledge comes from reflective experience, these experiences help answer which action is best and which actions are suboptimal. Experiments on iTHOR show that ExAC significantly improves navigation performance. Our method nearly doubles both the Success Rate and the Success weighted by Path Length compared to vanilla LLMs, reaching 73.93% and 48.35% in unseen scenes. These results validate the feasibility of our approach and pave the way for further research.

ACKNOWLEDGMENTS

This research is funded in part by ARC-Discovery grant (DP220100800 to XY and DP230101753), ARC-DECRA grant (DE230100477 to XY), Nvidia academic grant and the Analytics for the Australian Grains Industry (AAGI) Strategic Partnership with funding allocated by the Grains Research Development Corporation (GRDC UOQ2301-010OPX) and by The University of Queensland (DVCR2201A). We thank A/Prof. Liang Zheng, the area chair and anonymous reviewers for constructive comments.

REFERENCES

[1] A. Dubey, A. Jauhri, A. Pandey, and et al., "The llama 3 herd of models," *CoRR*, vol. abs/2407.21783, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2407.21783

[2] Y. Liu, W. Chen, Y. Bai, G. Li, W. Gao, and L. Lin, "Aligning cyber space with physical world: A comprehensive survey on embodied AI," *CoRR*, vol. abs/2407.06886, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2407.06886

[3] DeepSeek-AI, D. Guo, D. Yang, and et. al., "Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning," *CoRR*, vol. abs/2501.12948, 2025. [Online]. Available: https://doi.org/10.48550/arXiv.2501.12948

[4] Y. Ma, Z. Song, Y. Zhuang, J. Hao, and I. King, "A survey on vision-language-action models for embodied AI," *CoRR*, vol. abs/2405.14093, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2405.14093

[5] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin, W. X. Zhao, Z. Wei, and J. Wen, "A survey on large language model based autonomous agents," *Frontiers Comput. Sci.*, vol. 18, no. 6, p. 186345, 2024. [Online]. Available: https://doi.org/10.1007/s11704-024-40231-1

[6] S. Zhang, X. Song, Y. Bai, W. Li, Y. Chu, and S. Jiang, "Hierarchical object-to-zone graph for object navigation," in *2021 IEEE/CVF International Conference on Computer Vision, ICCV 2021, Montreal, QC, Canada, October 10-17, 2021*. IEEE, 2021, pp. 15 110–15 120. [Online]. Available: https://doi.org/10.1109/ICCV48922.2021.01485

[7] A. Khandelwal, L. Weihs, R. Mottaghi, and A. Kembhavi, "Simple but effective: CLIP embeddings for embodied AI," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2022, New Orleans, LA, USA, June 18-24, 2022*. IEEE, 2022, pp. 14 809–14 818. [Online]. Available: https://doi.org/10.1109/CVPR52688.2022.01441

[8] A. Majumdar, G. Aggarwal, B. Devnani, J. Hoffman, and D. Batra, "ZSON: zero-shot object-goal navigation using multimodal goal embeddings," in *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., 2022. [Online]. Available: http://papers.nips.cc/paper_files/paper/2022/hash/d0b8f0c8f79d3a621af945cafb669f4b-Abstract-Conference.html

[9] D. Shah, B. Osinski, B. Ichter, and S. Levine, "Lm-nav: Robotic navigation with large pre-trained models of language, vision, and action," in *Conference on Robot Learning, CoRL 2022, 14-18 December 2022, Auckland, New Zealand*, ser. Proceedings of Machine Learning Research, K. Liu, D. Kulic, and J. Ichnowski, Eds., vol. 205. PMLR, 2022, pp. 492–504. [Online]. Available: https://proceedings.mlr.press/v205/shah23b.html

[10] B. Yu, H. Kasaei, and M. Cao, "L3MVN: leveraging large language models for visual target navigation," in *IROS*, 2023, pp. 3554–3560. [Online]. Available: https://doi.org/10.1109/IROS55552.2023.10342512

[11] W. Cai, S. Huang, G. Cheng, Y. Long, P. Gao, C. Sun, and H. Dong, "Bridging zero-shot object navigation and foundation models through pixel-guided navigation skill," in *IEEE International Conference on Robotics and Automation, ICRA 2024, Yokohama, Japan, May 13-17, 2024*. IEEE, 2024, pp. 5228–5234. [Online]. Available: https://doi.org/10.1109/ICRA57147.2024.10610499

[12] E. Kolve, R. Mottaghi, D. Gordon, Y. Zhu, A. Gupta, and A. Farhadi, "AI2-THOR: an interactive 3d environment for visual AI," *CoRR*, vol. abs/1712.05474, 2017. [Online]. Available: http://arxiv.org/abs/1712.05474

[13] Q. Team, "Qwen2.5: A party of foundation models," September 2024. [Online]. Available: https://qwenlm.github.io/blog/qwen2.5/

[14] A. Q. Jiang, A. Sablayrolles, A. Mensch, C. Bamford, D. S. Chaplot, D. de Las Casas, F. Bressand, G. Lengyel, G. Lample, L. Saulnier, L. R. Lavaud, M. Lachaux, P. Stock, T. L. Scao, T. Lavril, T. Wang, T. Lacroix, and W. E. Sayed, "Mistral 7b," *CoRR*, vol. abs/2310.06825, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2310.06825

[15] G. Oriolo, M. Vendittelli, and G. Ulivi, "On-line map building and navigation for autonomous mobile robots," in *Proceedings of the 1995 International Conference on Robotics and Automation, Nagoya, Aichi, Japan, May 21-27, 1995*. IEEE Computer Society, 1995, pp. 2900–2906. [Online]. Available: https://doi.org/10.1109/ROBOT.1995.525695

[16] J. Borenstein and Y. Koren, "Real-time obstacle avoidance for fast mobile robots in cluttered environments," in *Proceedings of the 1990 IEEE International Conference on Robotics and Automation, Cincinnati, Ohio, USA, May 13-18, 1990*. IEEE, 1990, pp. 572–577. [Online]. Available: https://doi.org/10.1109/ROBOT.1990.126042

[17] X. Yu and F. Porikli, "Ultra-resolving face images by discriminative generative networks," in *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V*, ser. Lecture Notes in Computer Science, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds., vol. 9909. Springer, 2016, pp. 318–333. [Online]. Available: https://doi.org/10.1007/978-3-319-46454-1_20

[18] M. Jia, L. Shen, X. Shen, L. Liao, M. Chen, X. He, Z. Chen, and J. Li, "MNER-QG: an end-to-end MRC framework for multimodal named entity recognition with query grounding," in *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, B. Williams, Y. Chen, and J. Neville, Eds. AAAI Press, 2023, pp. 8032–8040. [Online]. Available: https://doi.org/10.1609/aaai.v37i7.25971

[19] P. Mirowski, R. Pascanu, F. Viola, H. Soyer, A. Ballard, A. Banino, M. Denil, R. Goroshin, L. Sifre, K. Kavukcuoglu, D. Kumaran, and R. Hadsell, "Learning to navigate in complex environments," in *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017. [Online]. Available: https://openreview.net/forum?id=SJMGPrcle

[20] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *2017 IEEE International Conference on Robotics and Automation, ICRA 2017, Singapore, Singapore, May 29 - June 3, 2017*. IEEE, 2017, pp. 3357–3364. [Online]. Available: https://doi.org/10.1109/ICRA.2017.7989381

[21] G. Sepulveda, J. C. Niebles, and A. Soto, "A deep learning based behavioral approach to indoor autonomous navigation," in *2018 IEEE International Conference on Robotics and Automation, ICRA 2018, Brisbane, Australia, May 21-25, 2018*. IEEE, 2018, pp. 4646–4653. [Online]. Available: https://doi.org/10.1109/ICRA.2018.8460646

[22] K. Chen, J. P. de Vicente, G. Sepulveda, F. Xia, A. Soto, M. Vázquez, and S. Savarese, "A behavioral approach to visual navigation with graph localization networks," in *Robotics: Science and Systems XV, University of Freiburg, Freiburg im Breisgau, Germany, June 22-26, 2019*, A. Bicchi, H. Kress-Gazit, and S. Hutchinson, Eds., 2019. [Online]. Available: https://doi.org/10.15607/RSS.2019.XV.010

[23] A. Mousavian, A. Toshev, M. Fiser, J. Kosecká, A. Wahid, and J. Davidson, "Visual representations for semantic target driven navigation," in *International Conference on Robotics and Automation, ICRA 2019, Montreal, QC, Canada, May 20-24, 2019*. IEEE, 2019, pp. 8846–8852. [Online]. Available: https://doi.org/10.1109/ICRA.2019.8793493

[24] H. Du, X. Yu, and L. Zheng, "Learning object relation graph and tentative policy for visual navigation," in *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part VII*, ser. Lecture Notes in Computer Science, A. Vedaldi, H. Bischof, T. Brox, and J. Frahm, Eds.,

vol. 12352. Springer, 2020, pp. 19–34. [Online]. Available: https://doi.org/10.1007/978-3-030-58571-6_2

[25] L. Mezghani, S. Sukhbaatar, T. Lavril, O. Maksymets, D. Batra, P. Bojanowski, and K. Alahari, "Memory-augmented reinforcement learning for image-goal navigation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2022, Kyoto, Japan, October 23-27, 2022*. IEEE, 2022, pp. 3316–3323. [Online]. Available: https://doi.org/10.1109/IROS47612.2022.9981090

[26] W. Chen, O. Mees, A. Kumar, and S. Levine, "Vision-language models provide promptable representations for reinforcement learning," *CoRR*, vol. abs/2402.02651, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2402.02651

[27] W. Tan, W. Zhang, S. Liu, L. Zheng, X. Wang, and B. An, "True knowledge comes from practice: Aligning llms with embodied environments via reinforcement learning," *CoRR*, vol. abs/2401.14151, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2401.14151

[28] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[29] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. H. Chi, Q. V. Le, and D. Zhou, "Chain-of-thought prompting elicits reasoning in large language models," in *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, Eds., 2022. [Online]. Available: http://papers.nips.cc/paper_files/paper/2022/hash/9d5609613524ecf4f15af0f7b31abca4-Abstract-Conference.html

[30] S. Yao, D. Yu, J. Zhao, I. Shafran, T. Griffiths, Y. Cao, and K. Narasimhan, "Tree of thoughts: Deliberate problem solving with large language models," in *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*, A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, Eds., 2023. [Online]. Available: http://papers.nips.cc/paper_files/paper/2023/hash/271db9922b8d1f4dd7aaef84ed5ac703-Abstract-Conference.html

[31] N. F. Rajani, B. McCann, C. Xiong, and R. Socher, "Explain yourself! leveraging language models for commonsense reasoning," in *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28-August 2, 2019, Volume 1: Long Papers*, A. Korhonen, D. R. Traum, and L. Màrquez, Eds. Association for Computational Linguistics, 2019, pp. 4932–4942. [Online]. Available: https://doi.org/10.18653/v1/p19-1487

[32] H. Zhao, H. Chen, F. Yang, N. Liu, H. Deng, H. Cai, S. Wang, D. Yin, and M. Du, "Explainability for large language models: A survey," *ACM Trans. Intell. Syst. Technol.*, vol. 15, no. 2, pp. 20:1–20:38, 2024. [Online]. Available: https://doi.org/10.1145/3639372

[33] W. Huang, C. Wang, R. Zhang, Y. Li, J. Wu, and L. Fei-Fei, "Voxposer: Composable 3d value maps for robotic manipulation with language models," in *Conference on Robot Learning, CoRL 2023, 6-9 November 2023, Atlanta, GA, USA*, ser. Proceedings of Machine Learning Research, J. Tan, M. Toussaint, and K. Darvish, Eds., vol. 229. PMLR, 2023, pp. 540–562. [Online]. Available: https://proceedings.mlr.press/v229/huang23b.html

[34] J. Lin, H. Gao, X. Feng, R. Xu, C. Wang, M. Zhang, L. Guo, and S. Xu, "The development of llms for embodied navigation," *CoRR*, vol. abs/2311.00530, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2311.00530

[35] M. Ahn, D. Dwibedi, C. Finn, M. G. Arenas, K. Gopalakrishnan, K. Hausman, B. Ichter, A. Irpan, N. J. Joshi, R. Julian, S. Kirmani, I. Leal, T. E. Lee, S. Levine, Y. Lu, S. Maddineni, K. Rao, D. Sadigh, P. Sanketi, P. Sermanet, Q. Vuong, S. Welker, F. Xia, T. Xiao, P. Xu, S. Xu, and Z. Xu, "Autort: Embodied foundation models for large scale orchestration of robotic agents," *CoRR*, vol. abs/2401.12963, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2401.12963

[36] D. Driess, F. Xia, M. S. M. Sajjadi, C. Lynch, A. Chowdhery, B. Ichter, A. Wahid, J. Tompson, Q. Vuong, T. Yu, W. Huang, Y. Chebotar, P. Sermanet, D. Duckworth, S. Levine, V. Vanhoucke, K. Hausman, M. Toussaint, K. Greff, A. Zeng, I. Mordatch, and P. Florence, "Palm-e: An embodied multimodal language model," in *International Conference on Machine Learning, ICML 2023, 23-29 July 2023, Honolulu, Hawaii, USA*, ser. Proceedings of Machine Learning Research, A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett, Eds., vol. 202. PMLR, 2023, pp. 8469–8488. [Online]. Available: https://proceedings.mlr.press/v202/driess23a.html

[37] S. Huang, Z. Jiang, H. Dong, Y. Qiao, P. Gao, and H. Li, "Instruct2act: Mapping multi-modality instructions to robotic actions with large language model," *CoRR*, vol. abs/2305.11176, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2305.11176

[38] P. Chen, P. Bu, J. Song, Y. Gao, and B. Zheng, "Can vlms play action role-playing games? take black myth wukong as a study case," *CoRR*, vol. abs/2409.12889, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2409.12889

[39] G. Zhou, Y. Hong, and Q. Wu, "Navgpt: Explicit reasoning in vision-and-language navigation with large language models," in *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, M. J. Wooldridge, J. G. Dy, and S. Natarajan, Eds. AAAI Press, 2024, pp. 7641–7649. [Online]. Available: https://doi.org/10.1609/aaai.v38i7.28597

[40] J. Chen, B. Lin, R. Xu, Z. Chai, X. Liang, and K. K. Wong, "Mapgpt: Map-guided prompting with adaptive path planning for vision-and-language navigation," in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand, August 11-16, 2024*, L. Ku, A. Martins, and V. Srikumar, Eds. Association for Computational Linguistics, 2024, pp. 9796–9810. [Online]. Available: https://doi.org/10.18653/v1/2024.acl-long.529

[41] Y. Long, X. Li, W. Cai, and H. Dong, "Discuss before moving: Visual language navigation via multi-expert discussions," in *IEEE International Conference on Robotics and Automation, ICRA 2024, Yokohama, Japan, May 13-17, 2024*. IEEE, 2024, pp. 17 380–17 387. [Online]. Available: https://doi.org/10.1109/ICRA57147.2024.10611565

[42] X. Bao, Z. Lv, and B. Wu, "Enhancing large language models with rag for visual language navigation in continuous environments," *Electronics*, vol. 14, no. 5, 2025. [Online]. Available: https://www.mdpi.com/2079-9292/14/5/909

[43] M. Wortsman, K. Ehsani, M. Rastegari, A. Farhadi, and R. Mottaghi, "Learning to learn how to learn: Self-adaptive visual navigation using meta-learning," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, 2019, pp. 6750–6759. [Online]. Available: http://openaccess.thecvf.com/content_CVPR_2019/html/Wortsman_Learning_to_Learn_How_to_Learn_Self-Adaptive_Visual_Navigation_Using_CVPR_2019_paper.html

[44] H. Du, L. Li, Z. Huang, and X. Yu, "Object-goal visual navigation via effective exploration of relations among historical navigation states," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17-24, 2023*. IEEE, 2023, pp. 2563–2573. [Online]. Available: https://doi.org/10.1109/CVPR52729.2023.00252

[45] S. Min, X. Lyu, A. Holtzman, M. Artetxe, M. Lewis, H. Hajishirzi, and L. Zettlemoyer, "Rethinking the role of demonstrations: What makes in-context learning work?" in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, Y. Goldberg, Z. Kozareva, and Y. Zhang, Eds. Association for Computational Linguistics, 2022, pp. 11 048–11 064. [Online]. Available: https://doi.org/10.18653/v1/2022.emnlp-main.759

[46] E. Akyürek, D. Schuurmans, J. Andreas, T. Ma, and D. Zhou, "What learning algorithm is in-context learning? investigations with linear models," in *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023. [Online]. Available: https://openreview.net/forum?id=0g0X4H8yN4I

[47] P. S. H. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive NLP tasks," in *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., 2020. [Online]. Available: https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html