

Horizon-Adaptive Mechanisms in LongNav-R1 Boost Out-of-Distribution Navigation Success

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: To what extent does the horizon-adaptive mechanism in LongNav-R1 improve success rates on out-of-distribution navigation instructions compared to standard fine-tuned VLA models. Language models (LMs) possess a strong capability to comprehend natural language, making them effective in translating human instructions into detailed plans for simple robot tasks. Nevertheless, it remains a significant challenge to handle long-horizon tasks, especially in. 15 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: LaMMA-P: Generalizable Multi-Agent Long-Horizon Task Allocation and Planning with LM-Driven PDDL Planner. Research question: To what extent does the horizon-adaptive mechanism in LongNav-R1 improve success rates on out-of-distribution navigation instructions compared to standard fine-tuned VLA models?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.5/10.

3 Results

16 papers retrieved. 15 claims extracted; 0 independently verified. Quality review score: 3.5/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The addition of pre-defined domains (D), Problem Generator (G), Precondition Identifier (P), and PDDL Validator (V) lead	×	0.04
Excluding all major components (P, V, G, and D) results in significantly lower performance compared to the full method.	×	0.02
The absence of pre-defined robot domains significantly decreases the LLM’s ability to identify appropriate sub-tasks for	×	0.05
Adding pre-defined domains (D) substantially increases the Success Rate (SR) for all categories of tasks.	×	0.05
The addition of the Problem Generator (G) increases Efficiency (Eff) for compound tasks and Robot Utilization (RU) for c	×	0.03
The PDDL Validator (V) ensures that all generated plans are executable before task execution, reducing task failures.	×	0.04
LaMMA-P using GPT-4o achieves the best performance when all components (D, G, P, V) are included.	×	0.04
The addition of the Precondition Identifier (P) simplifies complex preconditions, enabling LLMs to generate action seque	×	0.03
LaMMA-P significantly improves task success rates and robot utilization efficiency over existing methods.	×	0.09
LaMMA-P outperforms the strongest baseline SMART-LLM on the MAT-THOR benchmark in long-horizon tasks.	×	0.12
LaMMA-P assumes fully observable, static environments.	×	0.04
The ‘Prepare the Plate with Egg’ subtask requires the skills: GoToObject, PickupObject, and PutObject.	×	0.01
The ‘Microwave the Plate’ subtask requires the skills: GoToObject, OpenObject, PutObject, SwitchOn, SwitchOff, and Close	×	0.01
The ‘GoToObject’ action has preconditions including ‘(at-location ?egg ?location)’, ‘(at ?robot ?location)’, and ‘(not (×	0.02
The ‘CloseObject’ action has effects ‘(object-close ?robot ?microwave)’ and ‘(not (inaction ?robot))’.	×	0.03

References

- <https://arxiv.org/abs/2409.20560>
- <https://arxiv.org/abs/2509.02324>
- <http://arxiv.org/abs/2602.12351v1>