SOVEREIGN: What is the impact of varying reflection memory length on the success rate and average reward of LLM agents us

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

This paper develops LongNav-R1, an end-to-end multi-turn reinforcement learning (RL) framework designed to optimize Visual-Language-Action (VLA) models for long-horizon navigation. Unlike existing single-turn paradigm, LongNav-R1 reformulates the navigation decision process as a continuous multi-turn conversation between the VLA policy and the embodied environment. This multi-turn RL framework offers two distinct advantages: i) it enables the agent to reason about the causal effects of historical interactions and sequential future outcomes; and ii) it allows the model to learn directly from on

1 Introduction

Analysis of: LongNav-R1: Horizon-Adaptive Multi-Turn RL for Long-Horizon VLA Navigation. Research goal: What is the impact of varying reflection memory length on the success rate and average reward of LLM agents using Reflexion on the WebShop benchmark compared to fixed-length memory baselines?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

11 papers retrieved. 3 claims extracted, 3 verified. Tribunal: 7.8/10 → APPROVE (revision_round=0). Policy: AUTO_APPROVE.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
LongNav-R1 is an end-to-end framework that reformulates navigation as a multi-turn Reinforcement Learning process	✓	0.28
The model's generalizability and robustness are validated by its zero-shot performance in long-horizon real-world naviga	✓	0.26
LongNav-R1 outperforms state-of-the-art methods in long-horizon VLA navigation	✓	0.26

References

- http://arxiv.org/abs/2207.01206v4
- http://arxiv.org/abs/2602.12351v1
- http://arxiv.org/abs/2110.07276v3