SOVEREIGN: Does Reflexion's verbal reinforcement learning generalize to multimodal agents on the ALFRED benchmark for emb

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

Large language models (LLMs) have been increasingly used to interact with external environments (e.g., games, compilers, APIs) as goal-driven agents. However, it remains challenging for these language agents to quickly and efficiently learn from trial-and-error as traditional reinforcement learning methods require extensive training samples and expensive model fine-tuning. We propose Reflexion, a novel framework to reinforce language agents not by updating weights, but instead through linguistic feedback. Concretely, Reflexion agents verbally reflect on task feedback signals, then maintain the

1 Introduction

Analysis of: Reflexion: Language Agents with Verbal Reinforcement Learning. Research goal: Does Reflexion's verbal reinforcement learning generalize to multimodal agents on the ALFRED benchmark for embodied task completion, measured by success rate and subgoal accuracy?.

2 Methodology

Multi-query arXiv search (1 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

3 papers retrieved. 6 claims extracted, 0 verified. Tribunal: 1.7/10 $\rightarrow$ REJECT (revision_round=0). Policy: ESCALATE_TO_OWNER.

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
Reflexion improves performance over strong baselines by 22% in AlfWorld.	×	0.02
Reflexion improves performance over strong baselines by 20% in HotPotQA.	×	0.02
Reflexion improves performance over strong baselines by 11% on HumanEval.	×	0.04
ReAct + Reflexion completes 130 out of 134 tasks in AlfWorld.	×	0.02
ReAct + Reflexion learns to solve additional tasks by learning in 12 consecutive trials in AlfWorld.	×	0.05
In the ReAct-only approach, performance increase halts between trials 6 and 7 in AlfWorld.	×	0.02

References

- http://arxiv.org/abs/2602.12375v1
- http://arxiv.org/abs/2110.15191v1
- http://arxiv.org/abs/2303.11366v4