

# SOVEREIGN: How does the multi-turn RL framework in LongNav-R1 compare to fixed-length memory baselines on the VLN-CE bench

SOVEREIGN Research Kernel

Autonomous draft — Owner review required before publication

May 29, 2026

## Abstract

This paper surveys the field of reinforcement learning from a computer-science perspective. It is written to be accessible to researchers familiar with machine learning. Both the historical basis of the field and a broad selection of current work are summarized. Reinforcement learning is the problem faced by an agent that learns behavior through trial-and-error interactions with a dynamic environment. The work described here has a resemblance to work in psychology, but differs considerably in the details and in the use of the word “reinforcement.” The paper discusses central issues of reinfo

## 1 Introduction

Analysis of: Reinforcement Learning: A Survey. Research goal: How does the multi-turn RL framework in LongNav-R1 compare to fixed-length memory baselines on the VLN-CE benchmark in terms of success rate and SPL for navigation tasks with varying horizon lengths?.

## 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

9 papers retrieved. 3 claims extracted, 3 verified. Tribunal: 7.3/10 → 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
Reinforcement learning is the problem faced by an agent that learns behavior through trial-and-error interactions with a	✓	0.38
The work described in this paper has a resemblance to work in psychology, but differs considerably in the details and in	✓	0.33
The paper discusses central issues of reinforcement learning, including trading off exploration and exploitation, establ	✓	0.60

## References

- <https://doi.org/10.1186/s13059-016-0974-4>
- <https://doi.org/10.1613/jair.301>
- <https://doi.org/10.1051/0004-6361/201629272>