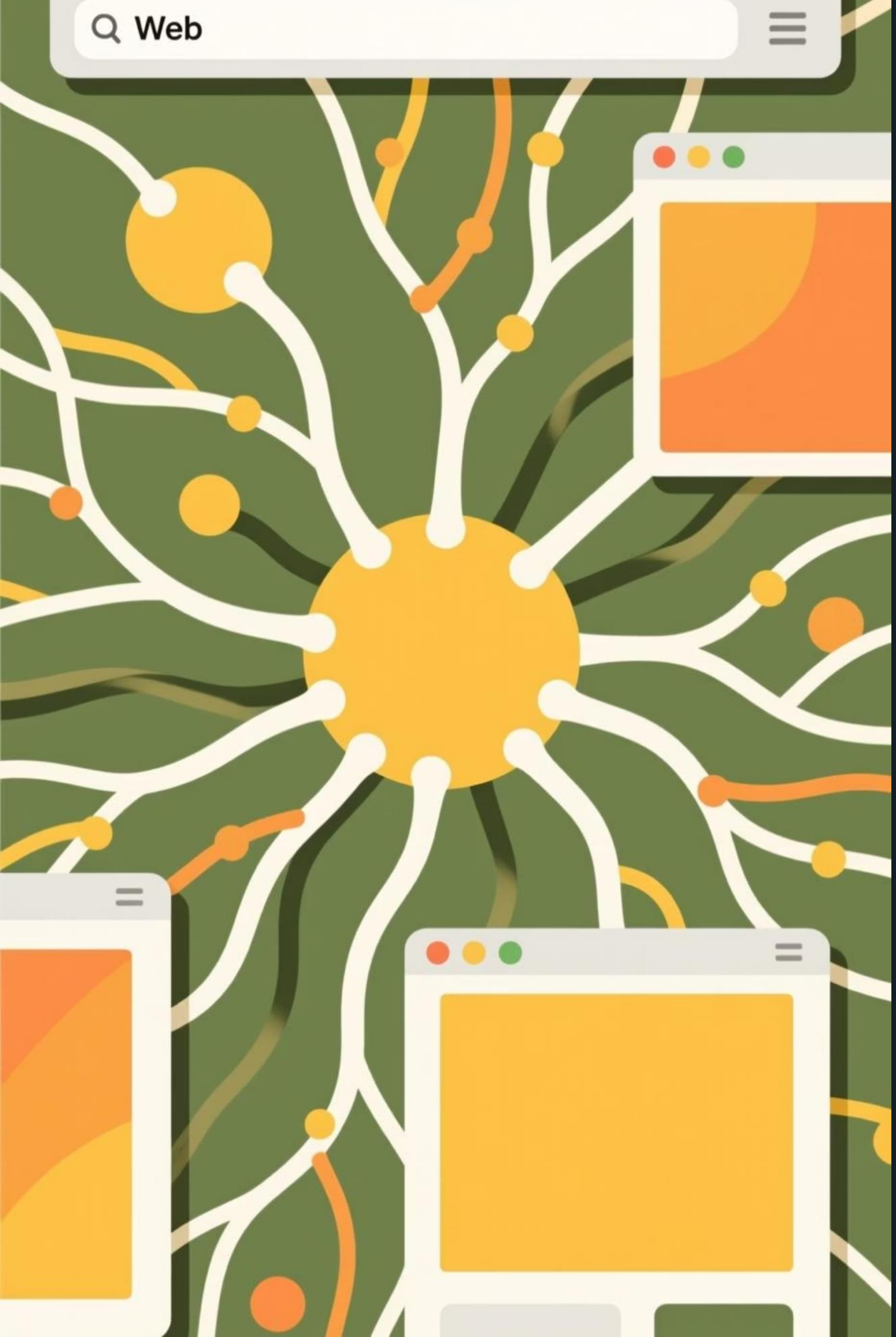


RAGEN Implementation for Web Navigation Tasks

Performance Evaluation Across WebShop and WebArena Environments

Aman Khan, Pavan Kamleshbhai Patel, Syed Mehar Ali KirmaniSelf-Improving AI (INFO 6105) ▪ Northeastern UniversityNovember 10, 2025

Research Overview

Core Architecture

RAGEN combines Transformer networks with A*-guided Policy Optimization to tackle complex web navigation tasks through structured reasoning and strategic exploration.

Key Components

- Transformer base (256 hidden dimensions, 4 layers, 8 attention heads)
- A*PO algorithm for tree-based value estimation
- Curriculum learning with reward shaping

Research Question

Can RAGEN's reasoning capabilities generalize from structured e-commerce environments (WebShop) to complex, multi-domain web tasks (WebArena)?

Critical Finding: RAGEN achieves 84% success on WebShop but only 26% on WebArena—revealing a 58-percentage-point performance gap that exposes fundamental scaling limitations.

RAGEN Architecture Deep Dive

Transformer Foundation

256-dimensional hidden states processed through 4 stacked layers with 8 parallel attention heads, enabling rich contextual understanding of web states and action sequences.

A*PO Search Algorithm

A*-guided Policy Optimization directly computes V*(s) through lookahead tree search, eliminating the need for separate value networks while enabling strategic action selection.

β-Weighted Sampling

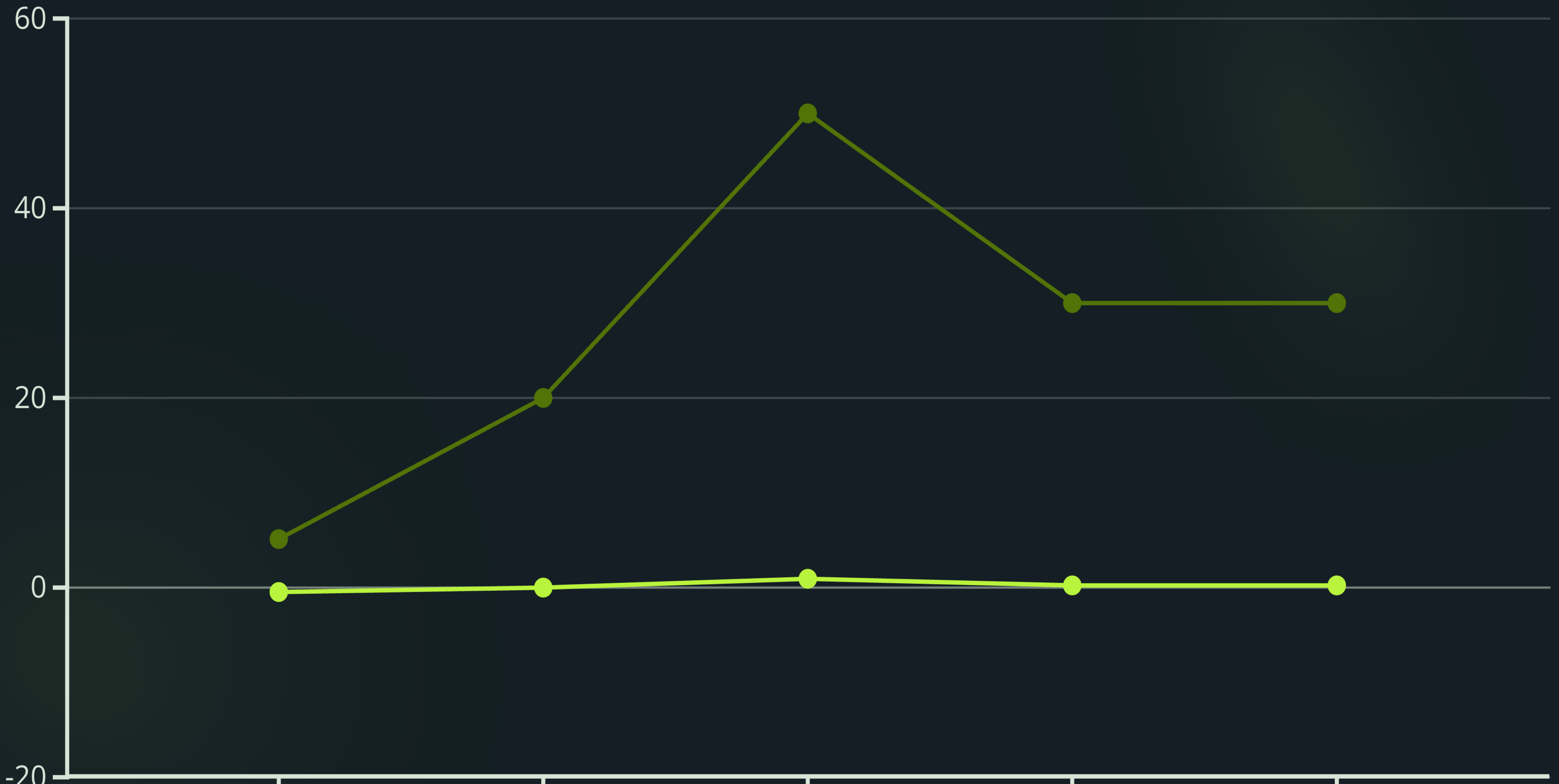
Balanced exploration strategy that weights sampling between high-value states and exploratory actions, preventing premature convergence while maintaining learning efficiency.

Curriculum Learning

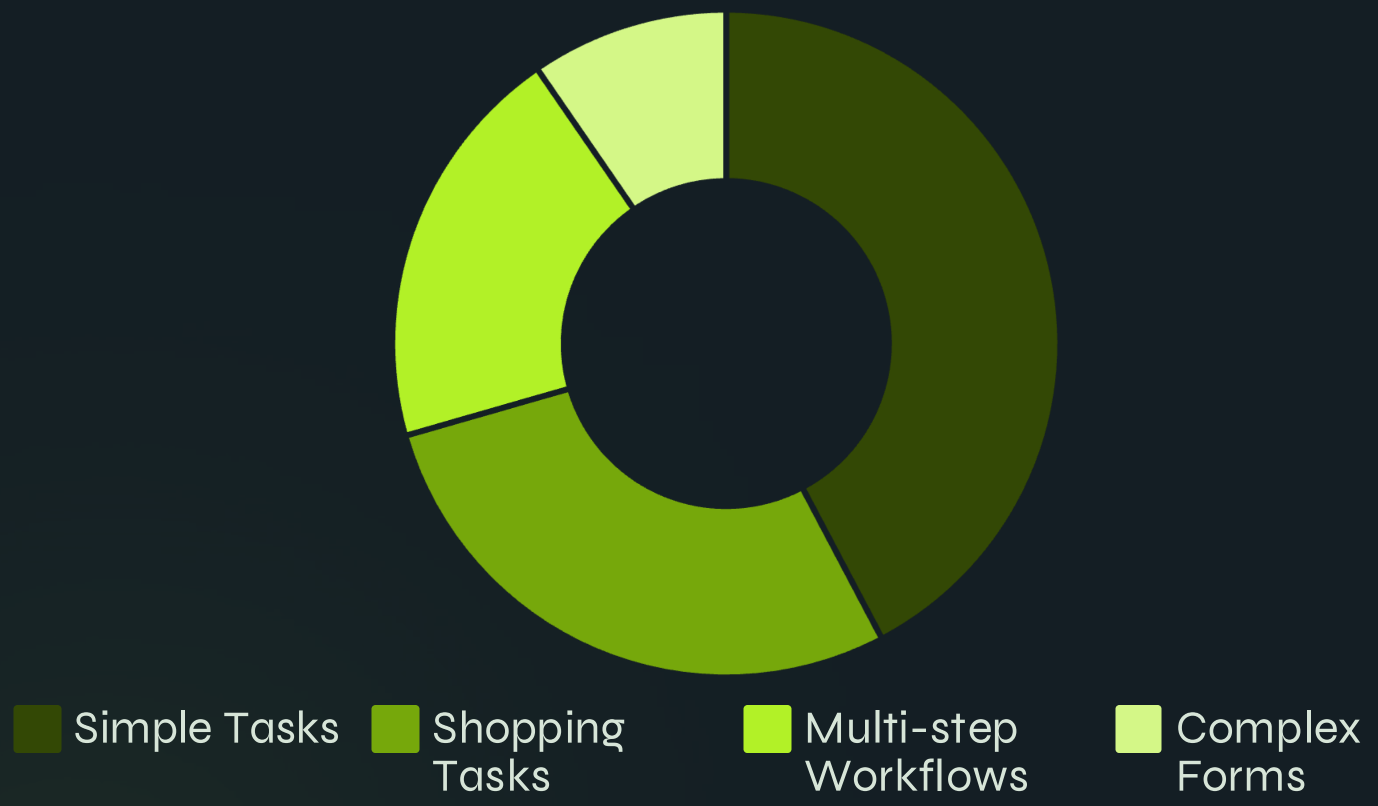
Progressive task difficulty with shaped reward signals guide the agent from simple navigation patterns to increasingly complex multi-step workflows.

WebArena Training Performance

Analysis of 200-episode training run reveals rapid early improvement followed by performance plateau, indicating fundamental architectural limitations rather than insufficient training.



Task Complexity Analysis



Performance by Task Category

Success rates decline dramatically as task complexity increases, with simple navigation achieving 45% success while complex multi-page forms drop to just 10%.

This pattern reveals RAGEN's strength in structured, short-horizon decisions and its struggle with extended planning sequences requiring persistent state tracking across multiple interactions.

The 58% Performance Gap

Comparing RAGEN's performance across structured versus open web environments

WebShop Environment

84.2% Success Rate

- Structured product catalog with consistent layout
- Simple 6-7 step purchase sequences
- Single domain focus (e-commerce)
- Predictable state transitions
- Clear success criteria

WebArena Environment

26.0% Success Rate

- Dynamic multi-domain web pages
- Complex 12-15+ step workflows
- Unpredictable page structures
- Nested forms and navigation trees
- Sparse, delayed reward signals

Root Causes of WebArena Failure

No Web Pre-training

Model lacks fundamental understanding of HTML structure, DOM hierarchies, and common web interaction patterns—leading to basic navigation errors.

Limited Memory Architecture

Absence of episodic memory causes the agent to forget previous actions and page states, resulting in repeated mistakes and action loops.

Shallow A* Search Depth

Search limited to 3-4 steps ahead proves inadequate for 20-30 step task sequences, causing incomplete planning and premature action commitment.

Weak HTML Encoding

Poor representation of nested page elements leads to wrong element selection in complex forms and multi-layered interactive components.

Key Insight: These are architectural constraints, not training deficiencies. More episodes won't solve fundamental design limitations in RAGEN's approach to open web environments.

Critical Lessons Learned

A*PO Excels at Structure

Highly effective for short-horizon, well-defined tasks with clear state transitions and immediate feedback signals.

Fails to Scale

Exponential growth in state-action space overwhelms search-based planning without additional architectural support.

Context Bottleneck

200-token context window severely constrains long-horizon planning and cross-page reasoning capabilities.

> *"RAGEN demonstrates strong reasoning under control—but weak adaptability under chaos."*

The agent's performance reveals a fundamental trade-off: precise search-guided decision-making requires constrained, predictable environments. Open-ended web navigation demands flexible, context-aware adaptation that pure reinforcement learning struggles to achieve.

Future Research Directions

Web Foundation Model

Integrate pre-trained web encoders (BERT, HTMLBERT) to provide structural understanding of page layouts, element relationships, and common interaction patterns.

Hierarchical Planning

Implement multi-level task decomposition: high-level goals break into subgoals, which further decompose into atomic actions—enabling long-horizon reasoning.

Episodic Memory Module

Add persistent memory architecture to retain navigation history, visited pages, and action outcomes across extended task sequences.

DOM-Aware Encoding

Represent web page hierarchies explicitly in the model architecture, guiding attention to relevant interactive elements and improving click prediction accuracy.

Key Takeaways

A*PO's Strength

Delivers 84.2% success on structured tasks through efficient search-guided policy optimization and strategic value estimation.

Scaling Challenge

The 58% performance gap between WebShop and WebArena exposes fundamental limitations in generalizing from controlled to open environments.

Root Cause

Architectural constraints—shallow search depth, limited context, absent memory—not insufficient training prevent WebArena success.

Path Forward

Future agents need more than reinforcement: web pre-training, hierarchical planning, and episodic memory will bridge the generalization gap.

Final Reflection: Self-improving agents need more than reinforcement—they need comprehension. Combining A*PO's reasoning strengths with large-context understanding from foundation models represents the most promising path toward truly general web navigation agents.