1. SKYRL-AGENT: EFFICIENT RL TRAINING FOR MULTI-TURN LLM AGENT

Accession number: 20250542710
Authors: Cao, Shiyi (1); Li, Dacheng (1); Zhao, Fangzhou (1); Yuan, Shuo (1); Hegde, Sumanth R. (2); Chen, Connor (1); Ruan, Charlie (1); Griggs, Tyler (1); Liu, Shu (1); Tang, Eric (2); Liaw, Richard (2); Moritz, Philipp (2); Zaharia, Matei (1); Gonzalez, Joseph E. (1); Stoica, Ion (1)
Author affiliation: (1) UC Berkeley, United States; (2) Anyscale, United States
Source title: arXiv
Abbreviated source title: arXiv
Issue date: November 20, 2025
Publication year: 2025
Language: English
E-ISSN: 23318422
Document type: Preprint (PP)
Repository: arXiv
Abstract: We introduce SKYRL-AGENT, a framework for efficient, multi-turn, long-horizon agent training and evaluation. It provides efficient asynchronous dispatching, lightweight tool integration, and flexible backend interoperability, enabling seamless use with existing RL frameworks such as SkyRL-train, VeRL, and Tinker. Using SKYRL-AGENT, we train SA-SWE-32B, a SWE agent trained from Qwen3-32B (24.4% Pass@1) purely with RL. We introduce two key components: an optimized asynchronous pipeline dispatcher that achieves a 1.55× speedup over naive asynchronous batching, and a tool-enhanced training recipe leveraging an AST-based search tool to facilitate code navigation, boost rollout Pass@K, and improve training efficiency. Together, these optimizations enable SA-SWE-32B to reach 39.4% Pass@1 on SWE-Bench Verified1 with more than 2× cost reduction than prior models reaching similar performance. Despite being trained solely on SWE tasks, SA-SWE-32B generalizes effectively to other agentic tasks, including Terminal-Bench, BrowseComp-Plus, and WebArena. We further demonstrate SKYRL-AGENT's extensibility through case studies on deep research, computer use, and memory agents, each trained using a different training backend. Copyright © 2025, The Authors. All rights reserved.
Number of references: 0
Main heading: Copyrights
Controlled terms: Agents - Cost reduction - Interoperability - Multi agent systems - Personnel training
Uncontrolled terms: Asynchronous pipeline - Code navigation - Costs reduction - Enhanced training - Multi-turn - Optimisations - RL framework - Search tools - Tool integration - Training efficiency
Classification code: 803 Chemical Agents and Basic Industrial Chemicals - 902.3 Legal Aspects - 911.2 Industrial Economics - 912.3 Personnel - 1101 Artificial Intelligence - 1106.9 Computer Software
Numerical data indexing: Percentage 2.44E+01%, Percentage 3.94E+01%
DOI: 10.48550/arXiv.2511.16108
Compendex references: YES
Preprint ID: 2511.16108v1
Preprint source website: https://arxiv.org
Preprint ID type: ARXIV
Database: Compendex
Data Provider: Engineering Village