

PR #6825 完整报告

verl-project/verl

[doc] chore: add verl extension guide

合并时间: 2026-06-23 21:30

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6825>

执行摘要

- 一句话: 添加 verl 扩展指南文档, 介绍自定义组件的方法
- 推荐动作: 值得所有希望扩展 verl 的研究人员和开发者阅读, 尤其是 Agent Loop 和 Replay Buffer 的自定义示例。建议在后续 PR 中及时更新文档以匹配实际代码接口。

功能与动机

PR body 中仅说明 'As title.', 即添加 verl 扩展指南。该文档旨在为用户提供扩展 verl 的完整指导, 降低自定义组件的门槛。

实现拆解

1. 创建文档文件 (docs/extend_guide.rst): 从零编写 273 行的扩展指南, 分为 RL 研究者、Engine 用户两大类读者, 覆盖 reward 函数定制、tool 定义、自定义 Agent Loop (如 MyAgentLoop.run)、自定义 Replay Buffer (UserCustomReplayBuffer.sample)、自定义模型引擎 (MyAgentLoopManager.create / generate_sequences / _load_vexact_replica) 等场景, 每部分包含配置示例和代码骨架。
2. 注册到索引 (docs/index.rst): 在 Programming guide 分类下添加一行 extend_guide 条目, 使新文档出现在 Sphinx 文档目录中。

关键文件:

- docs/extend_guide.rst (模块 扩展指南; 类别 docs; 类型 documentation; 符号 MyAgentLoop, run, UserCustomReplayBuffer, sample): 新增的扩展指南主体, 详细说明如何自定义 reward、tool、agent loop、replay buffer、model engine 等组件。
- docs/index.rst (模块 文档索引; 类别 docs; 类型 documentation): 注册新文档到索引, 使扩展指南出现在文档目录中。

关键符号: MyAgentLoop, run, UserCustomReplayBuffer, sample, MyAgentLoopManager, create, generate_sequences, _load_vexact_replica

评论区精华

Gemini Code Assist 的 review 指出了两个文档格式问题:

- :doc: 指令中链接文本与目标路径之间缺少空格 (第 21 行), 可能导致 Sphinx 构建警告。

- 配置代码块开头有多余的 + 字符（第 180 行），可能是从 git diff 复制导致，会使用户复制配置时出错。这两个建议均未被采纳（PR 已合并时未修复）。
- :doc: 指令格式错误 (documentation): 建议修复，但 PR 合并前未采纳。
- 配置中多余的 '+' 字符 (documentation): 建议删除该字符。

风险与影响

- 风险：纯文档变更，无运行时风险。但格式错误（如 :doc: 缺少空格）可能影响 Sphinx 渲染，降低阅读体验；配置示例中的多余字符可能误导用户。
- 影响：对用户：提供了一站式扩展指南，显著降低自定义组件的学习成本，尤其是 Agent Loop 和 Replay Buffer 部分。对系统：无影响。对团队：需与代码演进保持同步，防止指南过时。
- 风险标记：文档格式错误可能误导用户

关联脉络

- 暂无明显关联 PR