

PR #1987 完整报告

THUDM/slime

[ci] don't compare ref_logprob and logprob when R3 is on

合并时间: 2026-05-30 09:56

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1987>

执行摘要

- 一句话: R3 开启时跳过 CI 中的 logprob 对比检查
- 推荐动作: 可快速合入, 无精读必要。但注释中对 R3 路由差异的解释值得关注。

功能与动机

R3 模式下 actor forward 和 reference forward 的路由逻辑不同, log_probs 与 ref_log_probs 原本就不应完全一致, 但 CI 中的断言 `< 1e-8` 会误判为失败。PR 标题和改动注释明确说明了这一背景。

实现拆解

1. 在 `slime/backends/megatron_utils/data.py` 的 `log_rollout_data` 函数中, CI 检查前增加了 `not getattr(args, 'use_rollout_routing_replay', False)` 条件。
2. 添加了注释解释 R3 模式下 actor 与 reference forward 路由差异导致 logprob 不一致是预期行为。
3. 仅 4 行新增, 逻辑清晰, 无其他文件变更。

关键文件:

- `slime/backends/megatron_utils/data.py` (模块 数据层; 类别 source; 类型 core-logic; 符号 `log_rollout_data`): 核心改动文件, 在 CI 检查中添加了 R3 模式下的条件跳过逻辑。

关键符号: `log_rollout_data`

关键源码片段

`slime/backends/megatron_utils/data.py`

核心改动文件, 在 CI 检查中添加了 R3 模式下的条件跳过逻辑。

```
def log_rollout_data(...):
    # ... 前面的代码 ...
    reduced_log_dict = gather_log_data("rollout", args, rollout_id, log_dict)
    if args.ci_test and reduced_log_dict is not None:
        # R3 replays rollout routing for the actor forward, while the
        # reference forward intentionally falls through to normal routing.
        # Their log-probs are not expected to match bit-for-bit in CI.
        if (
```

```
rollout_id == 0
and not getattr(args, "use_rollout_routing_replay", False) # <-- 新增条件
and "rollout/log_probs" in reduced_log_dict
and "rollout/ref_log_probs" in reduced_log_dict
):
    # 仅当非 R3 模式时才执行数值对比
    assert abs(reduced_log_dict["rollout/log_probs"] - reduced_log_dict["rollout/ref_log_
    probs"]) < 1e-8
    if "rollout/log_probs" in reduced_log_dict:
        assert -0.5 < reduced_log_dict["rollout/log_probs"] < 0
    if "rollout/entropy" in reduced_log_dict:
        assert 0 < reduced_log_dict["rollout/entropy"] < 0.5
# ... 后续代码 ...
```

评论区精华

该 PR 无 review 评论，属于直接合并的简单修复。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：变更仅新增一个条件判断，且默认值为 False，不影响既有逻辑。若未来 R3 配置名称变化，需同步更新 getattr 的参数。
- 影响：影响范围仅限于 CI 环境：当启用 R3（use_rollout_routing_replay=True）时，不会再因 logprob 数值差异而报错。用户无感知。
- 风险标记：暂无

关联脉络

- PR #1965 Don't use sample.index as default rollout_id: 同为 CI 相关修复，涉及 rollout_id 使用。