

PR #2160 完整报告

THUDM/slime

Fix CI

合并时间: 2026-07-01 12:21

原文链接: <http://prhub.com.cn/THUDM/slime/pull/2160>

执行摘要

- 一句话: 修复 CI 测试, 调整梯度比对逻辑与配置
- 推荐动作: 本 PR 是典型的 CI 稳定性修复, 值得关注的是 `_assert_logprob_backward_close` 的设计: 通过 `bfloat16` 对齐来规避不同 kernel 的精度差异, 是一个实用的数值测试技巧。

功能与动机

修复 CI 测试中的数值精度不匹配和配置问题, 确保 GPU 上的 PPO logprob/entropy 反向传播测试能通过, 以及 Qwen3 并行检查测试的稳定性。

实现拆解

1. 新增 `_assert_logprob_backward_close` 函数 (`tests/test_ppo_logprob_entropy_gpu.py`) : 用于统一比较 log-prob 反向梯度。由于 Megatron 融合 CE kernel 会将梯度量化到 `bfloat16`, 而新实现保持 `fp32` 梯度, 因此先将双方梯度转换为 `bfloat16` 再对比, 避免精度差异导致的误报。
2. 重构 `_assert_legacyparity` 中的反向测试 (同上文件) : 将 log-prob 和 entropy 的梯度测试分离。log-prob 梯度使用 `_assert_logprob_backward_close` 在 `bfloat16` 精度下对比; entropy 梯度使用 `ENTROPY_BACKWARD_ATOL` 容差单独测试。同时, 将 `logprob_weights` 和 `entropy_weights` 的构造与 `backward` 调用解耦, 避免混合计算导致精度污染。
3. 添加 SGLang 配置参数 (`tests/test_qwen3_0.6B_parallel_check.py`) : 在 `sglang_args` 中增加 `--sglang-disable-piecewise-cuda-graph`, 禁用分段 CUDA graph 以规避潜在的稳定性问题。

关键文件:

- `tests/test_ppo_logprob_entropy_gpu.py` (模块测试; 类别 test; 类型 test-coverage; 符号 `_assert_logprob_backward_close`, `_assert_legacyparity`) : 主要变更文件, 新增梯度对比辅助函数并重构反向测试逻辑, 提升测试精度和稳定性。
- `tests/test_qwen3_0.6B_parallel_check.py` (模块测试; 类别 test; 类型 test-coverage) : 添加 `--sglang-disable-piecewise-cuda-graph` 参数, 禁用分段 CUDA graph 以提升 CI 稳定性。

关键符号: `_assert_logprob_backward_close`, `_assert_legacyparity`

关键源码片段

tests/test_ppo_logprob_entropy_gpu.py

主要变更文件，新增梯度对比辅助函数并重构反向测试逻辑，提升测试精度和稳定性。

```
# tests/test_ppo_logprob_entropy_gpu.py

def _assert_logprob_backward_close(actual_grad: torch.Tensor, legacy_grad: torch.Tensor) ->
None:
    # Megatron's fused vocab-parallel CE backward quantizes the log-prob
    # gradient to bfloat16 on CUDA. The new implementation keeps fp32 grads, so
    # compare this branch at the legacy kernel's effective precision.
    if actual_grad.is_cuda:
        actual_grad = actual_grad.to(torch.bfloat16)
        legacy_grad = legacy_grad.to(torch.bfloat16)
    torch.testing.assert_close(actual_grad, legacy_grad, rtol=BACKWARD_RTOL, atol=
BACKWARD_ATOL)
```

评论区精华

本 PR 无 review 评论或审核。

- 暂无高价值评论线程

风险与影响

- 风险：风险较低。变更仅影响测试文件，不会影响生产代码。主要风险在于测试重构可能遗漏某些梯度组合场景，但 PARITY_SCENARIOS 已覆盖常见参数组合，且分离后的测试逻辑更清晰可靠。
- 影响：影响范围仅限于 CI 测试管道。修复后，与 PPO logprob/entropy 计算相关的 GPU 测试和 Qwen3 并行检查将更加稳定，减少因数值精度或配置问题导致的 CI 失败。
- 风险标记：仅测试变更

关联脉络

- PR #2152 Optimize memory usage for _VocabParallelLogProbEntropy: 同一测试文件，涉及 PPO logprob/entropy 计算优化。
- PR #2144 perf: fuse PPO logprob entropy computation: 融合 logprob/entropy 计算，本 PR 的测试重构与之直接相关。
- PR #2153 bugfix: 修改了 with_entropy_grad 判断逻辑，本 PR 的测试调整与之配合。
- PR #2158 Remove ctx.set_materialize_grads(False) which may cause issues: 移除梯度 materialize 设置，可能影响测试中的梯度行为。