

PR #7836 完整报告

verl-project/verl

[recipe] fix: update GLM-5.2 GRPO training example on Ascend NPUs

合并时间: 2026-09-14 09:31

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7836>

执行摘要

- 一句话: 修复 GLM-5.2 Ascend 示例脚本: 关闭 checkpoint 严格校验并更新依赖 refs
- 推荐动作: 值得快速浏览, 但无需精读。该 PR 是典型的 recipe 维护类变更, 核心价值在于两点: 一是学会在示例脚本中通过 `checkpoint.strict=False` 处理权重冗余 / 缺失场景; 二是注意依赖 ref 与镜像栈的版本对齐。可作为 Ascend 示例脚本维护的参考。

功能与动机

PR body 中说明: 'MTP layers are unused in both training and rollout and are omitted from exported Hugging Face weights', 因此必须关闭 checkpoint 的严格模式, 否则权重加载会因缺少 MTP 层而失败; 同时修正 Megatron-LM 与 MegatronAdaptor 的依赖 commit, 使其与当前 Ascend 镜像 /CI 使用的版本保持一致。该 PR 是此前 PR#7788 新增 GLM-5.2 示例脚本后的跟进修复。

实现拆解

1. 关闭 checkpoint 严格校验: 在 `examples/ascend_extras/grpo_trainer/run_glm5_2_megatron.sh` 的 `ACTOR_ARGS` 数组中新增 `actor_rollout_ref.actor.checkpoint.strict=False`, 并添加注释说明原因 (MTP 层在训练 /rollout 中未使用, 且已从导出权重中移除)。这避免了 HF 权重加载时因缺失 MTP 层而抛出的 `strict` 模式异常。
2. 更新依赖 refs: 将脚本头部依赖表中的 Megatron-LM 分支从 `core_r0.18.0` 改为 `core_v0.18.0`, commit 从 `dc0ee41...` 更新为 `ba7b5eb...`; MegatronAdaptor 的 commit 从 `746fe27...` 更新为 `eb0e504...`。这些更新使脚本与 Ascend 镜像实际安装的依赖版本对齐, 避免因 commit 不匹配导致的编译或运行问题。
3. 无测试 / 配置配套改动: 该 PR 仅修改示例脚本, 没有新增测试或修改正式配置; 风险集中在脚本本身的可复现性, 不涉及核心训练代码。

关键文件:

- `examples/ascend_extras/grpo_trainer/run_glm5_2_megatron.sh` (模块 示例脚本; 类别 other; 类型 core-logic): 唯一变更文件, 包含两处修正: 新增 `checkpoint.strict=False` 配置和更新 Megatron-LM/MegatronAdaptor 依赖 commit。是 GLM-5.2 在 Ascend 上训练的关键入口脚本。

关键符号: 未识别

关键源码片段

examples/ascend_extras/grpo_trainer/run_glm5_2_megatron.sh

唯一变更文件，包含两处修正：新增 checkpoint.strict=False 配置和更新

Megatron-LM/MegatronAdaptor 依赖 commit。是 GLM-5.2 在 Ascend 上训练的关键入口脚本。

```
# 关键配置段（来自 run_glm5_2_megatron.sh）
# 关闭 checkpoint 严格校验：
# MTP 层在训练 /rollout 中均未使用，且已从导出的 Hugging Face 权重剔除，
# 若保持 strict=True，加载权重时会因缺少 MTP 层而报错。
ACTOR_ARGS=(
    actor_rollout_ref.actor.use_torch_compile=False
    actor_rollout_ref.actor.use_kl_loss=${use_kl_loss}
    actor_rollout_ref.actor.kl_loss_coef=${kl_loss_coef}
    actor_rollout_ref.actor.checkpoint.strict=False # 允许权重中缺少未使用的 MTP 层
    actor_rollout_ref.actor.clip_ratio_low=${clip_ratio_low}
    actor_rollout_ref.actor.clip_ratio_high=${clip_ratio_high}
    actor_rollout_ref.actor.clip_ratio_c=10.0
)

# 依赖版本表（更新部分）：
# NVIDIA/Megatron-LM core_v0.18.0 ba7b5ebce12af60627a80985792a1449ce45f46c
# Ascend/MegatronAdaptor core_r0.18.0 eb0e5043438dd753099432bc3b3d8e5916ed1844
```

评论区精华

本 PR 无公开 review 评论或讨论线程，唯一 reviewer (wucong25) 直接 APPROVED。讨论信息不足，无法提炼更多设计权衡。

- 暂无高价值评论线程

风险与影响

- 风险：
 1. 依赖 ref 固化风险：脚本将依赖锁定到特定 commit，若上游仓库历史被改写或新 commit 存在问题，脚本可能无法复现；但这是示例脚本的常见做法，风险可控。
 2. checkpoint.strict=False 的副作用：关闭严格校验后，若 checkpoint 中确实缺少某些关键层，训练可能在后续启动时才报错，问题定位成本更高。不过本例中 MTP 层确实未使用，因此影响有限。
 3. 仅示例脚本变更：不直接影响 verl 核心库，回归面极小。- 影响：影响范围仅限于 Ascend 平台上的 GLM-5.2 GRPO 示例脚本（examples/ascend_extras/grpo_trainer/run_glm5_2_megatron.sh）。对使用该脚本跑 GLM-5.2 训练的用户有直接收益：权重加载不再因 MTP 层缺失而失败，依赖版本与镜像一致，提高脚本可用性。对其它模块无影响。- 风险标记：依赖 commit 固化，strict 校验放宽的潜在隐藏问题，仅示例脚本影响

关联脉络

- PR #7788 [recipe] feat: add GLM-5.2 GRPO training example on Ascend NPUs: 本 PR 是对 PR#7788 新增的 GLM-5.2 示例脚本的跟进修复，修改同一文件。
- PR #7779 [vllm, hardware] fix: patch vllm-ascend for GLM-5.2 RL training on Ascend NPUs: 同为 GLM-5.2 在 Ascend 上的兼容性修复，属于同一条功能线。
- PR #7604 [megatron, ci] chore: upgrade Ascend stack to Megatron 0.18.0: Megatron-LM 和 MegatronAdaptor 依赖 refs 的更新与 Ascend 镜像栈的版本对齐背景相关。