PR #6601 完整报告

verl-project/verl

[misc] fix: revert run_qwen3_8b_fsdp_npu.sh

合并时间：2026-06-05 14:33

原文链接：http://prhub.com.cn/verl-project/verl/pull/6601

执行摘要

- 一句话：回退 Qwen3-8B FSDP NPU 训练脚本到旧版结构
- 推荐动作：此 PR 为结构清理性质，值得快速了解以理解 NPU 脚本的维护策略变更。未来迭代可考虑将 MODEL_PATH 改为可配置环境变量，或在通用脚本中更优雅地处理平台差异。

功能与动机

PR 标题和 body 明确说明 "revert run_qwen3_8b_fsdp_npu.sh"，旨在将之前可能合并的 NPU 专用脚本恢复为独立文件，确保 GPU 和 NPU 用户各自的脚本不受干扰。推测是为了解决 NPU 配置与通用脚本耦合导致的问题。

实现拆解

1. 恢复 NPU 专用脚本：在 *examples/ascend_extras/ppo_trainer/* 下新增 run_qwen3_8b_fsdp.sh（实际内容与旧版 run_qwen3_8b_fsdp_npu.sh 一致），该脚本硬编码了 NPU 相关的环境变量和参数（如 ulysses_sequence_parallel_size=2、enable_chunked_prefill=True 等）。
2. 移除通用脚本中的 NPU 配置：从 *examples/ppo_trainer/run_qwen3_8b_fsdp.sh* 中删除了根据 DEVICE 变量设置 NPU 环境变量的 *case* 语句块（约 15 行），使通用脚本不再污染 NPU 特定配置。
3. 更新文档中的脚本引用：在 docs/ascend_tutorial/model_support/model_and_algorithm_support.md 中，将 Qwen3-8B PPO 示例的链接从通用脚本指向新的 NPU 专用脚本。

关键文件：

- *examples/ascend_extras/ppo_trainer/run_qwen3_8b_fsdp.sh*（模块 NPU 示例；类别 other；类型 *core-logic*）：新增的 NPU 专用脚本，包含了完整的 PPO 训练配置和 NPU 特定参数（如 ulysses_sequence_parallel_size、enable_chunked_prefill 等）。
- *examples/ppo_trainer/run_qwen3_8b_fsdp.sh*（模块 PPO 示例；类别 other；类型 *core-logic*）：从通用脚本中移除了 NPU 特定环境变量设置，使其保持平台无关性。
- docs/ascend_tutorial/model_support/model_and_algorithm_support.md（模块 Ascend 文档；类别 docs；类型 documentation）：更新 Qwen3-8B PPO 示例的文档链接，指向新的 NPU 专用脚本，避免用户误用。

关键符号：未识别

关键源码片段

examples/ascend_extras/ppo_trainer/run_qwen3_8b_fsdp.sh

新增的 NPU 专用脚本，包含了完整的 PPO 训练配置和 NPU 特定参数（如 ulysses_sequence_parallel_size、enable_chunked_prefill 等）。

```
# NPU 专用 PPO 训练脚本 for Qwen3-8B (FSDP)
set -x

MODEL_PATH="Qwen/Qwen3-8B" # TODO: 可改为 ${MODEL_PATH:-"Qwen/Qwen3-8B"} 支持环境变量覆盖

python3 -m verl.trainer.main_ppo \
    algorithm.adv_estimator=gae \
    data.train_files=$HOME/data/gsm8k/train.parquet \
    data.val_files=$HOME/data/gsm8k/test.parquet \
    data.train_batch_size=32 \
    data.max_prompt_length=2000 \
    data.max_response_length=2000 \
    data.shuffle=False \
    actor_rollout_ref.model.path="${MODEL_PATH}" \
    actor_rollout_ref.model.use_remove_padding=True \
    actor_rollout_ref.model.enable_gradient_checkpointing=True \
    actor_rollout_ref.actor.optim.lr=1e-6 \
    actor_rollout_ref.actor.ppo_mini_batch_size=32 \
    actor_rollout_ref.actor.ppo_micro_batch_size_per_gpu=1 \
    actor_rollout_ref.actor.fsdp_config.param_offload=True \
    actor_rollout_ref.actor.fsdp_config.optimizer_offload=True \
    actor_rollout_ref.actor.use_kl_loss=False \
    actor_rollout_ref.actor.ulysses_sequence_parallel_size=2 \  # NPU 典型设置
    actor_rollout_ref.actor.use_dynamic_bsz=True \
    actor_rollout_ref.actor.use_torch_compile=False \
    actor_rollout_ref.rollout.log_prob_micro_batch_size_per_gpu=1 \
    actor_rollout_ref.rollout.tensor_model_parallel_size=1 \
    actor_rollout_ref.rollout.name=vllm \
    actor_rollout_ref.rollout.gpu_memory_utilization=0.9 \
    actor_rollout_ref.rollout.max_num_batched_tokens=4000 \
    actor_rollout_ref.rollout.max_num_seqs=64 \
    actor_rollout_ref.rollout.checkpoint_engine.update_weights_bucket_megabytes=4096 \
    actor_rollout_ref.rollout.log_prob_use_dynamic_bsz=True \
    actor_rollout_ref.rollout.enable_chunked_prefill=True \  # NPU 推荐
    actor_rollout_ref.rollout.enforce_eager=False \
    actor_rollout_ref.rollout.calculate_log_probs=True \
    critic.optim.lr=1e-5 \
    critic.model.use_remove_padding=True \
    critic.model.path="${MODEL_PATH}" \
    critic.model.enable_gradient_checkpointing=True \
    critic.ppo_micro_batch_size_per_gpu=1 \
    critic.ulysses_sequence_parallel_size=2 \
```

```
    critic.fsdp.param_offload=True \
    critic.fsdp.optimizer_offload=True \
    critic.use_dynamic_bsz=True \
    trainer.critic_warmup=0 \
    trainer.logger=console \
    trainer.project_name='verl_example_ppo_gsm8k' \
    trainer.experiment_name='qwen3_8b_fsdp' \
    trainer.n_gpus_per_node=8 \
    trainer.nnodes=1 \
    trainer.save_freq=-1 \
    trainer.test_freq=-1 \
    trainer.val_before_train=False \
    trainer.max_actor_ckpt_to_keep=1 \
    trainer.max_critic_ckpt_to_keep=1 \
    trainer.total_training_steps=15 2>&1 | tee ppo_qwen3-8b_fsdp_npu.log
```

评论区精华

机器人审查员 gemini-code-assist[bot] 对 MODEL_PATH 的硬编码提出高优先级建议，推荐改为支持环境变量覆盖（如 MODEL_PATH=${MODEL_PATH:-"Qwen/Qwen3-8B"}），以提升灵活性。但该建议未被采纳，最终 reviewer wucong25 直接批准了 PR。可能因为 PR 的 revert 性质不鼓励引入额外改动。

- MODEL_PATH 硬编码 vs 环境变量覆盖 (design): 未采纳，PR 被直接批准。可能是因为 revert 性质避免附加改动。

风险与影响

- 风险：低风险。本质是脚本组织方式回退，不涉及核心库代码。主要风险是用户如果习惯使用之前的通用脚本路径 examples/ppo_trainer/run_qwen3_8b_fsdp.sh 运行 NPU 训练，需要切换到新路径；文档已更新以降低混淆。
- 影响：影响范围局限于 NPU 用户和 Ascend 文档。GPU 用户不受影响。对于 NPU 用户，专用脚本提供了更明确的配置，但硬编码的 MODEL_PATH 仍可能给需要自定义路径的用户带来不便。
- 风险标记：硬编码路径限制灵活性

关联脉络

- 暂无明显关联 PR