

PR #2577 完整报告

radixark/miles

fix: make the Verifiers E2E self-contained

合并时间: 2026-08-17 02:33

原文链接: <http://prhub.com.cn/radixark/miles/pull/2577>

执行摘要

- 一句话: 修复 Verifiers E2E 运行环境, 纳入 short CI 范围
- 推荐动作: 值得精读 `prepare()` 的环境隔离思路 (`system-site-packages venv + uv tool dir` 路径解析), 并关注 CI scope 从专用 label 回退到 short 的决策过程。建议合并后观察 short CI 中该 E2E 的时长与稳定性, 若超时再考虑恢复专用 scope。

功能与动机

release CI run 31862426780 的 `stage-c-2-gpu-h200` shard 1 中, `test_qwen3_0.6B_verifiers.py` 在 `prepare()` 阶段因 `error: externally-managed-environment` 失败——测试尝试向系统 Python `/usr/bin/python3` 安装可选依赖。此外, `prime env install` 指向的解释器与 Ray workers 不一致, `MilesSGLangTransport.post` 缺少 `Renderer` 所需的 token 标签、`uv tool install` 的 `prime` 可执行文件不在继承的 `PATH` 中, 共同导致 Verifiers E2E 无法在冻结的 release 镜像中自足运行。

实现拆解

1. 隔离依赖安装: 在 `tests/e2e/long/test_qwen3_0.6B_verifiers.py` 中引入 `VERIFIERS_VENV` 与 `VERIFIERS_SITE_PACKAGES`, 通过 `uv venv --clear --python {sys.executable} --system-site-packages` 创建基于系统解释器且带 `system-site-packages` 的虚拟环境, 将 `requirements.txt` 装进该 `venv`, 避免向受管理的 `/usr/bin/python3` 写入。同时将 `prime env install` 改为 `"$(uv tool dir --bin)/prime"` 调用, 解决 `uv tool install` 后可执行文件不在 `PATH` 的问题。
2. 打通 Ray worker 依赖: `execute()` 里把 `VERIFIERS_SITE_PACKAGES` 与 `CODE_GOLF_DIR` 前置到 `PYTHONPATH`, 确保 Ray 各 worker 能同时见到 `venv` 中的可选包和本地克隆的 CodeGolf taskset; 并新增 `--attention-backend flash` 与发布的 `launcher` 保持一致, 避免推理后端不一致。
3. 修正 transport 输出格式: 在 `examples/experimental/verifiers/verifiers_rollout.py` 的 `MilesSGLangTransport.post` 中, 构造 `logprobs.content` 时对每个 token 增加 `"token": f"token_id:{token_id}"`, 并用 `zip(..., strict=True)` 保证 token id 与 logprob 位置绑定, 满足当前 `Renderer` 的解析要求。
4. CI scope 收敛: 最初提交为 Verifiers 增加了专用 `run-ci-verifiers` 选择器并排除其隐式 scope, 最终提交将其简化为直接挂在 `short` 标签下, 移除策略改动, 减少维护面。

5. 测试配套: tests/fast/rollout/test_verifiers_rollout.py 的

test_transport_translates_renderer_request_to_sglang 从仅断言 token_ids 改为完整比对 choices[0], 覆盖新 token 标签输出。

关键文件:

- examples/experimental/verifiers/verifiers_rollout.py (模块 实验示例; 类别 source; 类型 core-logic; 符号 MilesSGLangTransport) : 修复 MilesSGLangTransport.post 未输出 Renderer 所需的 token 标签, 导致 rollout 数据解析失败。
- tests/e2e/long/test_qwen3_0.6B_verifiers.py (模块 E2E 测试; 类别 test; 类型 test-coverage; 符号 prepare, execute) : E2E 测试脚本重构 prepare/execute, 解决系统 Python 污染和 Ray worker 找不到 CodeGolf 环境的问题, 并改 short scope。
- tests/fast/rollout/test_verifiers_rollout.py (模块 单测覆盖; 类别 test; 类型 test-coverage; 符号 test_transport_translates_renderer_request_to_sglang) : 更新传输测试断言以匹配新的 token 标签格式。

关键符号: MilesSGLangTransport.post, prepare, execute,
test_transport_translates_renderer_request_to_sglang

关键源码片段

examples/experimental/verifiers/verifiers_rollout.py

修复 MilesSGLangTransport.post 未输出 Renderer 所需的 token 标签, 导致 rollout 数据解析失败。

```
# 从 SGLang 响应 meta_info 中取 token 级 logprob 列表, 每项形如 [logprob, token_id]
token_logprobs = list(meta_info.get("output_token_logprobs") or [])
completion_ids = [int(item[1]) for item in token_logprobs]
completion_logprobs = [float(item[0]) for item in token_logprobs]

# 组装 OpenAI 兼容的 choices 结构
response_body = {
    "request_id": output.get("request_id") or f"vf-{uuid.uuid4().hex}",
    "choices": [
        {
            "token_ids": completion_ids,
            "logprobs": {
                "content": [
                    # 当前 Renderer 依赖 "token" 标签, 格式为 token_id:<id>, 必须与 completion_ids
                    位置一一对应
                    {"token": f"token_id:{token_id}", "logprob": logprob}
                    for token_id, logprob in zip(completion_ids, completion_logprobs, strict=True)
                ]
            },
            "finish_reason": _finish_reason(output),
        }
    ],
}
```

tests/e2e/long/test_qwen3_0.6B_verifiers.py

E2E 测试脚本重构 prepare/execute，解决系统 Python 污染和 Ray worker 找不到 CodeGolf 环境的问题，并改 short scope。

```
def prepare():
    # 准备目录并下载模型 checkpoint
    U.exec_command_cpu(f"mkdir -p {MODEL_DIR} {RUN_DIR}")
    U.exec_command_cpu(f"hf download Qwen/{MODEL_NAME} --local-dir {MODEL_DIR}/{MODEL_NAME}")

    # 为可选依赖创建 system-site-packages venv，避免写入受管系统 Python
    U.exec_command_cpu(f"uv venv --clear --python {sys.executable} --system-site-packages {VERIFIERS_VENV}")
    U.exec_command_cpu(
        f"{VERIFIERS_VENV}/bin/python -m pip install "
        f"-r {U.repo_base_dir}/examples/experimental/verifiers/requirements.txt"
    )

    # 安装 prime 工具，并通过 uv tool dir --bin 解析其真实路径，因为 tool 目录未必在 runner 的 PATH 上
    U.exec_command_cpu("uv tool install 'prime==0.6.19'")
    if not VERIFIERS_DIR.exists():
        U.exec_command_cpu(
            f"git clone --depth 1 --branch v0.2.0 "
            f"https://github.com/PrimeIntellect-ai/verifiers.git {VERIFIERS_DIR}"
        )
    shutil.copytree(
        VERIFIERS_DIR / "environments" / "code_golf_v1",
        CODE_GOLF_DIR,
        dirs_exist_ok=True,
    )
    U.exec_command_cpu(f'cd {RUN_DIR} && "$(uv tool dir --bin)/prime" --plain env install code-golf-v1')

    # 将 HF checkpoint 转换为 Megatron 格式
    U.convert_checkpoint(
        model_name=MODEL_NAME,
        megatron_model_type=MODEL_TYPE,
        num_gpus_per_node=NUM_GPUS,
        dir_dst=str(MODEL_DIR),
        hf_checkpoint=str(MODEL_DIR / MODEL_NAME),
        megatron_path=str(MEGATRON_PATH),
    )
```

评论区精华

本 PR 没有产生实质 review 讨论；[nblintao](#) 直接批准。提交历史体现了 CI scope 的决策演进：第 2 个提交为 Verifiers 增加了专用 [run-ci-verifiers](#) 选择器并排除隐式 scope，最终提交又

将其回退为直接使用 **short** 标签，理由是 E2E 已估算 900 秒、属于现有 short 范围，避免引入额外策略维护。

- 暂无高价值评论线程

风险与影响

- 风险：主要风险集中在测试环境配置：uv venv 与 uv tool dir 依赖 CI runner 的 uv 版本，若 uv 升级可能影响路径解析；PYTHONPATH 前置 venv site-packages 可能遮蔽镜像内固定的核心依赖（如 torch），尽管使用 --system-site-packages 缓解，但顺序问题仍需观察；--attention-backend flash 与发布 launcher 的对齐依赖镜像内置后端支持。另外 scope 从 long 改为 short 使该 E2E 在 PR CI 中更频繁执行，若 est_time 900 秒被低估可能拉长队列。
- 影响：影响范围集中在 Verifiers 示例的 E2E 验证链路：修复后该测试可在冻结 release 镜像中自足运行，CI 从 long 移到 short 后会在每次 PR CI 中执行，提升回归发现速度。对生产功能无直接影响；对使用 MilesSGLangTransport 的其他调用方无破坏（token 标签是新增字段）。团队维护者获得更可靠的环境隔离范式和 CI scope 取舍参考。
- 风险标记：环境隔离依赖 uv, CI scope 调整，Renderer 格式耦合

关联脉络

- PR #2522 Make the class-based rollout the default and convert legacy path to env var gated: 该 PR 将 class-based rollout 设为默认并让旧路径受环境变量门控；本 PR 的 E2E 明确设置 MILES_USE_LEGACY_ROLLOUT_V1=1，并要求 verifiers_rollout 在旧路径下可用，二者共同维护 Verifiers 示例的兼容性。
- PR #2544 Do not kill the run when one sample's collect_samples loses its connection: 该 PR 涉及 rollout 生成链路（agentic_tool_call）的断连容忍，与本 PR 同属 rollout 运行时稳定性改进，且都修改了 tests/fast/rollout 下的测试。