

PR #31442 完整报告

sgl-project/sglang

ci: add MLX to coverage report backend display order

合并时间: 2026-07-18 08:32

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/31442>

执行摘要

- 一句话: 修复 CI 覆盖率报告缺少 MLX 后端
- 推荐动作: 该 PR 是常规维护, 技术层面无需精读。作为 CI 基础设施的快速修复, 其设计模式 (使用元组 + 断言强制同步 HWBackend 枚举) 值得在类似场景中参考。

功能与动机

PR #31442 的 body 指出, 自从 #30121 将 MLX 加入 HWBackend 后, 夜间 CI Coverage Overview 作业持续失败。失败根因在于 ci_coverage_report.py 中的断言要求 BACKEND_DISPLAY_ORDER 必须覆盖 HWBackend 的所有成员, 而该元组未及时同步更新。

实现拆解

在 `scripts/ci/utils/ci_coverage_report.py` 的 `BACKEND_DISPLAY_ORDER` 元组末尾添加字符串 "MLX", 并修改注释使其与当前后端列表一致。该元组本是为了控制覆盖率报告表格中各后端列的顺序, 断言同时充当同步检查。变更使断言通过, 覆盖率报告恢复正常渲染。

关键文件:

- `scripts/ci/utils/ci_coverage_report.py` (模块 CI 工具; 类别 infra; 类型 infrastructure; 符号 `BACKEND_DISPLAY_ORDER`): 唯一变更文件。在 `BACKEND_DISPLAY_ORDER` 元组中添加 MLX, 修复断言失败。

关键符号: 未识别

关键源码片段

`scripts/ci/utils/ci_coverage_report.py`

唯一变更文件。在 `BACKEND_DISPLAY_ORDER` 元组中添加 MLX, 修复断言失败。

```
# scripts/ci/utils/ci_coverage_report.py
```

```
# BACKEND_DISPLAY_ORDER 控制覆盖率报告中各后端的列顺序。
# 其内容必须与 HWBackend 枚举完全一致, 否则下方的 assert 会失败。
# 这种设计强迫开发者在新增后端时同步更新此处, 避免遗漏。
BACKEND_DISPLAY_ORDER = (
    "CUDA", "AMD", "NPU", "CPU", "XPU", "MUSA", "MLX" # 新增 MLX
)
```

```
# 运行时断言：若不一致则抛出 AssertionError,
# 从根源上防止新后端被静默忽略。
assert set(BACKEND_DISPLAY_ORDER) == {b.name for b in HWBackend}, \
    "BACKEND_DISPLAY_ORDER is out of sync with HWBackend"
```

评论区精华

无实质性技术讨论。仅有一条来自 [gemini-code-assist\[bot\]](#) 的自动评论，表示没有反馈意见。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：仅修改 CI 工具脚本中的列表末尾，不改变任何核心逻辑或运行时行为。
若 MLX 后端不再存在或名称变更，该元组将再次不同步，但断言会及时暴露问题。
- 影响：直接修复夜间 CI 覆盖率报告生成失败问题。MLX 后端相关的 12 个注册项现在会正确展示在覆盖率表格中。对用户和其他系统无影响。
- 风险标记：暂无

关联脉络

- PR #30121 Add MLX to HWBackend: 本次 PR 直接修复了 #30121 引入 MLX 后未更新 BACKEND_DISPLAY_ORDER 导致的 CI 失败。