

PR #29376 完整报告

sgl-project/sglang

Fix test_type_based_dispatcher after TokenizedGenerateReqInput field changes

合并时间: 2026-06-30 13:22

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/29376>

执行摘要

- 一句话: 修复测试适配 struct 字段变更
- 推荐动作: 该 PR 属于常规测试修复, 值得精读程度低。但可关注 PR #29214 的结构体变更, 若使用相关 IPC 类需注意参数变化。

功能与动机

`test/registered/utils/test_type_based_dispatcher.py` 在 AMD CI 中因 `TokenizedGenerateReqInput.__init__()` 缺少 `input_embeds` 和 `token_type_ids` 抛出 `TypeError`。根因是 PR #29214 重构了 IPC 结构体定义, 但未同步更新测试构造器。

实现拆解

1. 添加 `input_embeds=None` 和 `token_type_ids=None`: 在两个 `TokenizedGenerateReqInput(...)` 构造调用中增加缺失的参数, 以匹配新的必需字段签名。
2. 重命名 `image_inputs` 为 `mm_inputs`: 在两个 `TokenizedEmbeddingReqInput(...)` 调用中将参数名同步更新为新名称。
3. 批量构造: 对 `BatchTokenizedGenerateReqInput` 内的嵌套 `TokenizedGenerateReqInput` 同样应用上述修改。

关键文件:

- `test/registered/utils/test_type_based_dispatcher.py` (模块测试; 类别 test; 类型 test-coverage): 唯一变更文件, 修正了因 IPC 结构体字段重排和重命名导致的测试失败。

关键符号: 未识别

关键源码片段

`test/registered/utils/test_type_based_dispatcher.py`

唯一变更文件, 修正了因 IPC 结构体字段重排和重命名导致的测试失败。

```
# 修改后的测试构造调用片段
# TokenizedGenerateReqInput: 补充新必需的 input_embeds 和 token_type_ids 字段
test_requests.append(
    TokenizedGenerateReqInput(
        input_text="",
```

```

        input_ids=[1, 2],
        input_embeds=None, # 新增：从可选变为必需，默认为 None
        mm_inputs=dict(),
        token_type_ids=None, # 新增：从可选变为必需，默认为 None
        sampling_params=SamplingParams(),
        return_logprob=False,
        logprob_start_len=0,
        top_logprobs_num=0,
        token_ids_logprob=[1, 2],
        stream=False,
    )
)

# TokenizedEmbeddingReqInput: 将 image_inputs 重命名为 mm_inputs
test_requests.append(
    TokenizedEmbeddingReqInput(
        input_text="",
        input_ids=[1, 2],
        mm_inputs=dict(), # 原 image_inputs=dict()
        token_type_ids=[1, 2],
        sampling_params=SamplingParams(),
    )
)

```

评论区精华

无 review 评论或争议。HaiShaw 直接批准并触发 CI。

- 暂无高价值评论线程

风险与影响

- 风险：无风险。仅修正测试文件中的参数匹配问题，不涉及生产代码。
- 影响：影响范围极小，仅修复单个测试文件，使其通过 CI。对系统稳定性、性能、安全无影响。
- 风险标记：暂无

关联脉络

- PR #29214 [Cleanup] IPC struct renames, better typing, and SenderWrapper removal: 本 PR 的直接原因：该 PR 重构了 TokenizedGenerateReqInput 和 TokenizedEmbeddingReqInput 的字段，导致测试构造失败。