

PR #44011 完整报告

vllm-project/vllm

[CI] Remove redundant test_chat_with_tool_reasoning.py

合并时间: 2026-05-30 03:23

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/44011>

执行摘要

- 一句话: 删除冗余测试并清理 CI ignore 配置
- 推荐动作: 建议合并, 属于常规清理。可关注是否还有其他类似已跳过的测试需要一并清理, 以保持代码库整洁。

功能与动机

PR 描述指出: Remove test_chat_with_tool_reasoning.py because it has redundant coverage to test_completion_with_function_calling.py and has been skipped on CI. Clean up --ignore entries from .buildkite/test_areas/entrypoints.yaml, .

buildkite/test-amd.yaml, and commented-out references in .

buildkite/test_areas/rust_frontend.yaml。主要动机是清理冗余、已跳过的测试, 减少维护负担和 CI 配置复杂度。

实现拆解

1. 删除测试文件: 移除 tests/entrypoints/openai/chat_completion/test_chat_with_tool_reasoning.py, 该文件包含 Qwen/QwQ-32B 模型的流式和非流式工具调用测试, 已被标记跳过且覆盖重复。
2. 清理核心 CI 配置: 在 .buildkite/test_areas/entrypoints.yaml 和 .buildkite/test-amd.yaml 中, 从 pytest 命令的 --ignore 参数中去掉对被删除文件的引用 (保留 test_oop_registration.py 的忽略)。
3. 清理 Rust 前端 CI 注释: 在 .buildkite/test_areas/rust_frontend.yaml 中, 将被删除文件相关的注释行 (# - tests/... 和 # - pytest -v -s ...) 替换为其他先前已注释的文件引用, 保持注释占位风格一致。

关键文件:

- tests/entrypoints/openai/chat_completion/test_chat_with_tool_reasoning.py (模块 工具推理测试; 类别 test; 类型 deletion; 符号 server, client, extract_reasoning_and_calls, test_chat_streaming_of_tool_and_reasoning): 被删除的测试文件, 测试工具调用与推理结合场景, 已被判定冗余且跳过。
- .buildkite/test-amd.yaml (模块 CI 配置; 类别 config; 类型 configuration): 移除两处 --ignore 中对被删除测试文件的引用。

- `.buildkite/test_areas/entrypoints.yaml` (模块 CI 配置; 类别 config; 类型 configuration) : 移除主 CI 配置中对应的 `--ignore` 条目。
- `.buildkite/test_areas/rust_frontend.yaml` (模块 CI 配置; 类别 config; 类型 configuration) : 清理 Rust 前端 CI 中被注释掉的测试文件引用, 替换为其他注释文件。

关键符号: `extract_reasoning_and_calls`, `test_chat_streaming_of_tool_and_reasoning`, `test_chat_full_of_tool_and_reasoning`

关键源码片段

`tests/entrypoints/openai/chat_completion/test_chat_with_tool_reasoning.py`

被删除的测试文件, 测试工具调用与推理结合场景, 已被判定冗余且跳过。

```
# 被删除的测试文件中的一部分: 流式工具调用与推理测试
# 使用 Qwen/QwQ-32B 模型, 启用 deepseek_r1 推理解析器和 hermes 工具调用解析器
# 测试验证流式输出中正确提取推理内容和工具调用参数
```

```
async def test_chat_streaming_of_tool_and_reasoning(client: openai.AsyncOpenAI):
    # 发起流式聊天完成请求
    stream = await client.chat.completions.create(
        model=MODEL_NAME,
        messages=MESSAGES,
        tools=TOOLS,
        temperature=0.0,
        stream=True,
    )
    chunks = []
    async for chunk in stream:
        chunks.append(chunk)

    # 从流式块中提取推理内容和工具调用
    reasoning, arguments, function_names = extract_reasoning_and_calls(chunks)
    # 断言存在推理内容、函数名称和参数
    assert len(reasoning) > 0
    assert len(function_names) > 0 and function_names[0] == FUNC_NAME
    assert len(arguments) > 0 and arguments[0] == FUNC_ARGS
```

评论区精华

该 PR 无 review 评论, 仅有 reviewer yewentao256 的 LGTM 批准。无额外讨论。

- 暂无高价值评论线程

风险与影响

- 风险: 低风险。被删除的测试已在 CI 中跳过 (通过 `--ignore`), 且已有 `test_completion_with_function_calling.py` 提供等效覆盖。需确认是否有其他外部脚本或文档引用该文件路径, 但仓库内未发现。CI 配置的 `--ignore` 移除后, CI 运行不会再尝试跳过不存在的文件, 避免潜在警告。

- 影响：影响范围：仅限测试和 CI 配置，无运行时功能影响。开发者和 CI 系统不再维护冗余测试，CI 日志更简洁，后续维护成本降低。对用户无直接影响。
- 风险标记：低风险，配置清理，无功能影响

关联脉络

- 暂无明显关联 PR