

# PR #22586 完整报告

sgl-project/sglang

[Test] Fix flaky `test\_function\_call\_required` by adding `strict=True`

合并时间: 2026-04-11 17:14

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/22586>

## 执行摘要

- 一句话: 修复工具调用测试的随机失败问题, 通过添加 `strict=True` 确保参数模式强制约束。
- 推荐动作: 该 PR 变更简单直接, 无需深入阅读。值得关注的是 `gemini-code-assist[bot]` 的建议——考虑将 `strict=True` 一致地应用于其他相关测试用例, 以全面消除随机失败风险。

## 功能与动机

PR #21593 之后, `tool_choice="required" + strict=false` (默认) 不再通过约束解码强制执行参数模式——只保证产生工具调用, 参数是尽力而为的。小型模型 (如 Llama-3.2-1B) 偶尔在没有模式约束的情况下生成空 `{}` 参数, 导致约 30% 的随机失败。PR #21593 正确更改了协议语义, 但遗漏了更新此测试文件 (只有 `test_tool_choice.py` 更新了 `strict=True`) 。

## 实现拆解

实现非常简单: 在 `test/registered/openai_server/function_call/test_openai_function_calling.py` 文件的 `test_function_call_required` 方法中, 为 `tools` 列表中的每个工具添加 `strict=True` 字段。这确保了约束解码强制执行参数模式, 防止小型模型生成空参数。

关键文件:

- `test/registered/openai_server/function_call/test_openai_function_calling.py` (模块测试 / OpenAI 服务器 / 函数调用): 唯一修改的文件, 修复了 `test_function_call_required` 测试用例的随机失败问题。

关键符号: `test_function_call_required`

## 评论区精华

`gemini-code-assist[bot]` 在 review 中建议将此修复一致地应用于同一文件中使用 `tool_choice='required'` 的其他测试用例 (如 `test_function_call_specific` 和 `test_streaming_multiple_choices_finish_reason`), 以避免类似随机失败。但 PR 作者未回应此建议, PR 已合并。

- 测试用例一致性修复 (testing): PR 作者未回应此建议, PR 已合并, 其他测试用例未修改。

## 风险与影响

- 风险：风险极低：仅修改测试文件，不涉及生产代码。潜在风险是其他使用 `tool_choice='required'` 的测试用例可能仍然存在随机失败，但不会影响系统功能。
- 影响：影响范围仅限于测试稳定性：修复了特定测试用例的随机失败，提高了 CI 可靠性。对用户、系统或团队无直接影响。
- 风险标记：测试覆盖不完整

## 关联脉络

- PR #21593 Fix tool call constrained decoding and parsing for models with native formats: 该 PR 引入了协议语义更改，导致本 PR 修复的测试随机失败问题。本 PR 是 #21593 的后续修复。