

PR #20543 完整报告

sgl-project/sglang

fix: do not strip whitespace from GLM tool call values

合并时间: 2026-04-10 02:14

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/20543>

执行摘要

- 一句话: 修复 GLM 工具调用参数值中重要空格被错误剥离的问题。
- 推荐动作: 该 PR 值得快速浏览以理解工具调用解析中的空格处理陷阱。关注点: 1. 为什么 `arg_key.strip()` 保留而 `arg_value.strip()` 移除的设计决策。2. 新增测试如何模拟真实场景 (代码缩进)。3. 可扩展思考: 其他检测器是否有类似问题。

功能与动机

修复 Issue #20542 中报告的问题: GLM4 和 GLM47 MoE 检测器的 `_parse_argument_pairs` 方法在解析工具调用参数值时调用了 `arg_value.strip()`, 这会剥离字符串参数值的前导和尾随空格。当参数值包含代码编辑工具调用中的源代码时, 这种剥离会导致语义丢失, 例如代码缩进被移除。Issue 中提供了具体示例: 包含 4 个前导空格的 `old_string` 参数值会被错误地剥离为 `def foo()`。

实现拆解

实现方案分为两部分: 1. 核心修复: 在 `glm4_moe_detector.py` 和 `glm47_moe_detector.py` 的 `_parse_argument_pairs` 方法中移除 `arg_value = arg_value.strip()` 行, 保留 `arg_key.strip()` (键名仍需要清理)。2. 测试验证: 在 `test_function_call_parser.py` 中为两个检测器类分别添加 `test_whitespace_preserved_in_arg_values` 测试方法, 验证包含前导空格的参数值能够正确保留。

关键文件:

- `python/sglang/srt/function_call/glm4_moe_detector.py` (模块 `function_call`): GLM4 MoE 检测器的核心修复文件, 移除 `arg_value.strip()` 调用
- `python/sglang/srt/function_call/glm47_moe_detector.py` (模块 `function_call`): GLM47 MoE 检测器的核心修复文件, 同步移除 `arg_value.strip()` 调用
- `test/registered/unit/function_call/test_function_call_parser.py` (模块 `test`): 新增测试验证空格保留功能, 提供回归防护

关键符号: `_parse_argument_pairs`, `test_whitespace_preserved_in_arg_values`

评论区精华

review 讨论较少但明确: gemini-code-assist[bot] 确认这是 " 正确且必要的修复 ", 指出变更 " 最小化、目标明确且应用一致 "。JustinTong0323 直接批准合并。Issue 评论中 guapisolo 提到这是 "RL token in token out 的重要功能 " 且 " 不侵入整体设计 ", 表明修复的重要性。

- 空格剥离修复的正确性与必要性 (correctness): 修复被批准并合并
- 修复对整体设计的影响 (design): 修复被认可为重要且非侵入性

风险与影响

- 风险: 风险较低但需注意: 1. 回归风险: 移除 strip 可能影响原本依赖空格清理的现有用例, 但 Issue 指出下游 parse_arguments 函数已正确处理原始值 (字符串类型通过 json.dumps 保留空格, 其他类型通过 json.loads 空格不敏感)。2. 测试覆盖: 新增测试仅验证了前导空格场景, 未覆盖尾随空格、混合空格或非字符串类型参数, 但现有测试套件应提供基础保障。3. 兼容性: 修复针对 GLM 特定检测器, 不影响其他工具调用解析器。
- 影响: 影响范围有限但重要: 1. 用户影响: 修复后, 使用 GLM 模型进行代码编辑等工具调用的开发者将获得正确的参数值保留, 避免语义丢失。2. 系统影响: 仅修改两个检测器解析逻辑, 不改变整体架构或性能特征。3. 团队影响: 修复简单明确, 易于理解和维护, 新增测试提供了防护网。
- 风险标记: 潜在回归风险, 测试覆盖有限

关联脉络

- 暂无明显关联 PR