

PR #33518 完整报告

sgl-project/sglang

feat(api): add sglex_spec

合并时间: 2026-08-19 11:54

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/33518>

执行摘要

- 一句话:
- 推荐动作:

功能与动机

实现拆解

1. 核心源码路径 - `python/sglang/srt/entrypoints/openai/utils.py` (core-logic) : 源码主路径; 涉及符号 `spec_tokens_details_from_meta_info`, `process_spec_tokens_details_from_ret`; 包含 控制流调整、配置键调整、符号定义调整; +48/-0; 关键符号: `spec_tokens_details_from_meta_info`, `process_spec_tokens_details_from_ret` - `python/sglang/srt/entrypoints/openai/serving_chat.py` (core-logic) : 源码主路径; 包含 控制流调整、配置键调整; +47/-5 - `python/sglang/srt/entrypoints/openai/serving_completions.py` (core-logic) : 源码主路径; 包含 控制流调整、配置键调整; +47/-5
2. 测试与验证 - `test/registered/unit/entrypoints/openai/test_serving_completions.py` (test-coverage) : 测试配套; 涉及符号 `_spec_result`, `test_parallel_sampling_returns_spec_details_per_choice`, `test_streaming_parallel_sampling_orders_spec_details_by_choice`, `mock_generate`; 包含 测试覆盖调整、控制流调整、配置键调整; +121/-0; 关键符号: `_spec_result`, `test_parallel_sampling_returns_spec_details_per_choice`, `test_streaming_parallel_sampling_orders_spec_details_by_choice`, `mock_generate` - `test/registered/unit/entrypoints/openai/test_serving_chat.py` (test-coverage) : 测试配套; 涉及符号 `_spec_result`, `test_parallel_sampling_returns_spec_details_per_choice`, `test_streaming_parallel_sampling_orders_spec_details_by_choice`, `mock_generate`; 包含 测试覆盖调整、控制流调整、配置键调整; +78/-0; 关键符号: `_spec_result`, `test_parallel_sampling_returns_spec_details_per_choice`, `test_streaming_parallel_sampling_orders_spec_details_by_choice`, `mock_generate`

关键文件:

- `python/sglang/srt/entrypoints/openai/utils.py` (模块 `sglang/srt`; 类别 `source`; 类型 `core-logic`; 符号 `spec_tokens_details_from_meta_info`, `process_spec_tokens_details_from_ret`) : 源码主路径; 涉及符号

spec_tokens_details_from_meta_info, process_spec_tokens_details_from_ret; 包含 控制流调整、配置键调整、符号定义调整; +48/-0

- test/registered/unit/entrypoints/openai/test_serving_completions.py (模块 serving/completions; 类别 test; 类型 test-coverage; 符号 _spec_result, test_parallel_sampling_returns_spec_details_per_choice, test_streaming_parallel_sampling_orders_spec_details_by_choice, mock_generate) : 测试配套; 涉及符号 _spec_result, test_parallel_sampling_returns_spec_details_per_choice, test_streaming_parallel_sampling_orders_spec_details_by_choice, mock_generate; 包含 测试覆盖调整、控制流调整、配置键调整; +121/-0
- test/registered/unit/entrypoints/openai/test_serving_chat.py (模块 serving/chat; 类别 test; 类型 test-coverage; 符号 _spec_result, test_parallel_sampling_returns_spec_details_per_choice, test_streaming_parallel_sampling_orders_spec_details_by_choice, mock_generate) : 测试配套; 涉及符号 _spec_result, test_parallel_sampling_returns_spec_details_per_choice, test_streaming_parallel_sampling_orders_spec_details_by_choice, mock_generate; 包含 测试覆盖调整、控制流调整、配置键调整; +78/-0
- python/sglang/srt/entrypoints/openai/serving_chat.py (模块 sglang/srt; 类别 source; 类型 core-logic) : 源码主路径; 包含 控制流调整、配置键调整; +47/-5
- python/sglang/srt/entrypoints/openai/serving_completions.py (模块 sglang/srt; 类别 source; 类型 core-logic) : 源码主路径; 包含 控制流调整、配置键调整; +47/-5

关键符号: spec_tokens_details_from_meta_info, process_spec_tokens_details_from_ret, _spec_result, test_parallel_sampling_returns_spec_details_per_choice, test_streaming_parallel_sampling_orders_spec_details_by_choice, mock_generate, run_stream, mock_single_generate, SpecTokensDetails

评论区精华

- 暂无高价值评论线程

风险与影响

- 风险:
- 影响:
 - 风险标记: 暂无

关联脉络

- 暂无明显关联 PR