

PR #35419 完整报告

sgl-project/sglang

[Docs] Update contribution guide

合并时间: 2026-08-20 13:31

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/35419>

执行摘要

- 一句话: 更新贡献与评估指南, 对齐最新 CI 与测试流程
- 推荐动作: 值得作为文档维护的样板浏览, 重点看 '测试注册与本地运行' 和 'Mintlify 链接检查' 两节, 它们准确反映了 SGLang CI 的真实行为; 若你在维护内核或自定义库 (DeepGEMM/DeepEP), 新增的工作流章节有直接参考价值。无需精读代码逻辑。

功能与动机

PR body 明确提出要让贡献者文档跟上当前工作流: 'Updates the contribution and model evaluation guides to reflect current SGLang testing, CI rerun permissions, documentation validation, accuracy evaluation, code-style, and kernel-development workflows.' 此前文档中链接检查方式 (lychee)、单元测试执行方式与评估命令均已过时, 容易误导新贡献者并造成无效 PR 往返。

实现拆解

1. 更新 docs/docs/developer_guide/contribution_guide.mdx 的链接检查章节: 将 CI 强制执行的 lychee 描述改为 Mintlify 检查 (mint@4.2.559 的 broken-links --check-anchors --check-redirects), lychee 仅保留为仓库级 README.md 的手动 pre-commit hook。
2. 重写单元测试章节: 明确注册测试必须继承 CustomTestCase、调用 register_*_ci() 注册、提供标准 __main__ 入口, CI 通过 test/run_suite.py 发现注册测试并以 fail-fast 直接执行; 将示例文件名更新为 test_radix_cache_unit.py, 并新增直接 python3 执行单个注册测试文件的本地运行方式, 使其与 CI 行为一致。
3. 补充 CI 重新运行权限、AOT wheel 兼容性要求、自定义内核更新工作流 (含 sgl-deep-gemm、sgl-deep-ep 指引) 以及本地 DeepEP 验证章节, 响应 reviewer Fridge003 的补充建议。
4. 更新 docs/docs/developer_guide/evaluating_new_models.mdx: 将 GSM8K 评估从 python -m sglang.test.few_shot_gsm8k 迁移到 run_eval --eval-name gsm8k, 参数 --num-questions 改为 --num-examples, host 去掉协议前缀; 同步更新 --thinking-mode 支持值列表 (deepseek-v3、qwen-3、glm-45、kimi-k2)。
5. 全程无源码与测试改动, Mintlify preview 部署成功, 最终由 Fridge003 批准合并。

关键文件:

- docs/docs/developer_guide/contribution_guide.mdx (模块 贡献指南; 类别 docs; 类型 documentation) : 主体变更文件 (+97/-43), 重写链接检查、单元测试注册与本地运行、CI rerun 权限、自定义内核更新和 DeepEP 验证等章节, 决定贡献者文档与真实 CI 行为的一致性。
- docs/docs/developer_guide/evaluating_new_models.mdx (模块 评估指南; 类别 docs; 类型 documentation) : 修正 GSM8K 评估命令 (迁移到 run_eval --eval-name gsm8k) 与 --thinking-mode 支持值列表, 避免贡献者运行错误命令。

关键符号: 未识别

关键源码片段

docs/docs/developer_guide/contribution_guide.mdx

主体变更文件 (+97/-43), 重写链接检查、单元测试注册与本地运行、CI rerun 权限、自定义内核更新和 DeepEP 验证等章节, 决定贡献者文档与真实 CI 行为的一致性。

```
# 文档链接与锚点检查: CI 已从 lychee 切换到 Mintlify
# 安装与 CI 相同版本的 Mintlify CLI (固定 4.2.559)
npm install -g mint@4.2.559
# 在 docs 目录下执行与 CI 相同的 broken-links 检查 (含锚点与重定向)
cd docs && mint broken-links --check-anchors --check-redirects

# 单元测试本地运行: 直接执行注册测试文件最接近 CI 行为
# CI 通过 test/run_suite.py 发现 register_*_ci() 注册的测试并以 fail-fast 执行
python3 test/registered/unit/mem_cache/test_radix_cache_unit.py
# 或用 pytest 跑全部 / 单模块
pytest test/registered/unit/ -v
pytest test/registered/unit/mem_cache/ -v
```

docs/docs/developer_guide/evaluating_new_models.mdx

修正 GSM8K 评估命令 (迁移到 run_eval --eval-name gsm8k) 与 --thinking-mode 支持值列表, 避免贡献者运行错误命令。

```
# GSM8K 精度评估: 统一走 run_eval 入口, 用 --eval-name 选择数据集
python -m sglang.test.run_eval \
  --eval-name gsm8k \
  --host 127.0.0.1 \
  --port 30000 \
  --num-examples 200 \
  --num-shots 5

# 推理模型需显式指定 --thinking-mode
# 当前支持 deepseek-v3、qwen-3、glm-45、kimi-k2; 模型已强制 thinking 时可省略
```

评论区精华

Fridge003 在 review 中提出两处补充建议: 一是为自定义库 sgl-deep-gemm (DeepGEMM 的 sgl_deep_gemm 打包) 与 sgl-deep-ep (DeepEP 的 sgl_deep_ep 打包) 补充对应的

README 指引链接；二是希望新增 'How to update kernels in sglang' 章节。作者 mmangkad 回复 'Done, was still doing it when you posted this review', 表示正在同步完善，随后 Fridge003 批准 PR。讨论属于内容完善型反馈，无设计分歧。

- 补充自定义库与内核更新指引 (documentation): 作者回复 'Done, was still doing it when you posted this review', 按建议补充了自定义库与内核更新章节，Fridge003 随后批准。

风险与影响

- 风险：纯文档变更，无回归、性能与安全风险。主要风险是文档与真实 CI 行为的漂移：例如 mint@4.2.559 版本固定、--num-examples 等命令参数若后续变更，需要同步维护文档；新增的 CI 重新运行权限说明需与 CI_PERMISSIONS.json 实际配置保持一致，本 PR 未对该配置做对应验证。另外文档中的测试注册约定 (CustomTestCase、register_*_ci()) 若未来调整，文档也需要同步跟进。
- 影响：影响对象为新贡献者与维护者：文档现在提供与 CI 一致的单测本地执行方式（直接 python3 运行注册测试文件）、Mintlify 链接检查方法、以及 GSM8K 等精度评估的正确命令，能减少无效 PR 往返和 reviewer 的沟通成本；新增的自定义内核更新与 DeepEP 本地验证章节降低了内核开发者的上手门槛。对运行时系统无影响，仓库外部用户基本不受影响。
- 风险标记：文档命令与 CI 行为需持续同步，新增章节缺少自动化校验，无测试与源码配套改动

关联脉络

- PR #35597 [misc] Add a comment style rule to .claude/rules: 同为开发者流程 / 规范类文档调整，与本次贡献指南更新共同体现 SGLang 正在系统化开发者约定与文档体系。