

# PR #25012 完整报告

sgl-project/sglang

[Spec] Drop Rule 5 (`\_len` / `\_lens` ban) from speculative naming rule

合并时间: 2026-05-12 05:35

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/25012>

## 执行摘要

本 PR 移除了投机解码命名规范中关于禁止使用 `_len` / `_lens` 的 Rule 5, 并相应调整后续规则编号。纯文档变更, 无源码修改, 不影响运行时行为。

## 功能与动机

原规则 Rule 5 明确禁止在内部代码中使用 `length` / `lens`, 但实践中 `_len` / `_lens` 在投机解码标量名中很常见 (如 `accept_len`、`ema_accept_len`), 且与 PyTorch 生态中 Triton kernel 的惯用法一致 (`seq_lens`、`cu_seqlens_q`)。团队认为该限制过于严格, 决定放宽以更好地对齐社区实践。

## 实现拆解

- 删除 Rule 5 及其例外说明: 原规则占 10 行, 包括禁令和 Triton kernel 签名的特例说明, 全部移除。
- 重新编号后续规则: 原 Rule 6 (冗余 `_token_id` 后缀) 变为 Rule 5; 原 Rule 7 (单复数规则) 变为 Rule 6。
- 新增 out-of-scope 说明: 在文档末尾添加一条说明, 明确 `_len` / `_lens` 是可接受的命名形式, 同时强调 `num_X` 仍是首选 (Rule 4), 且 Rule 1 (去掉 `-ed`) 仍然要求使用 `accept_length` 而非 `accepted_length`。

变更文件 `.claude/rules/speculative-naming.md` 末尾新增内容如下:

```
- `_len` / `_lens` names: `num_X` is preferred for counts (Rule 4), but `_len` / `_lens` names are acceptable. Triton kernel params in particular often use `_lens` / `_len` to align with the PyTorch ecosystem (`seq_lens`, `cu_seqlens_q`). Rule 1 still requires the `-ed`-less form (`accept_length` OK, `accepted_length` not).
```

## 评论区精华

无实质性讨论 (仅有自动化消息)。

## 风险与影响

- 风险: 无。纯文档变更, 不涉及任何代码。
- 影响: 开发者编写投机解码相关代码时, 可以更自由地使用 `_len` / `_lens`, 无需担心违反规范。同时, `num_X` 作为首选形式仍被推荐, 保持了命名的一致性。

## 关联脉络

- 同属投机解码模块的清理工作，与 PR #25010（移除死参数和过时注释）形成小型重构系列。
- 与 PR #24965（修复 ngram 指标命名中的 num\_accepted\_drafts\_per\_req\_cpu）呼应，体现了对命名规范的持续完善。