

PR #2477 完整报告

radixark/miles

fix(docs): correct agentic rollout guidance

合并时间: 2026-08-13 15:32

原文链接: <http://prhub.com.cn/radixark/miles/pull/2477>

执行摘要

- 一句话: 修正 agentic rollout 文档, 整合 TITO 会话与生成端点指南
- 推荐动作: 值得精读。对使用 agentic rollout 与 TITO 的工程用户, 这是当前最权威的配置与契约说明; 对文档维护者, 该 PR 展示了“症状 - 根因 - 修复 - 验证”的文档修正方法和页面整合时保留 redirect 的迁移策略。值得关注的设计决策: 文档围绕“token 所有权”这一单一事实组织 agentic 路径, 用一张对比表呈现 generate 函数的两种兼容形式, 并把 CLI help 修正与文档重构放进同一个 PR 以保证契约一致性。

功能与动机

PR body 以症状 - 复现 - 根因 - 修复四段式列出动机: (1) `rollout-endpoints.md` 将 chat-turn 的 token 化职责分配给 SGLang, 并要求 agent 提供 prompt token IDs (`git show 35c701ba:docs/user-guide/rollout-endpoints.md` 可复现), 这与 session-server 实际持有 token 的 TITO 架构不符; (2) `agentic-chat-template.md` 遗漏 v2 分支语义, 并将 TITO 从整体 agentic 工作流中割裂; (3) `--use-session-server` 的帮助文本要求同时设置 `--chat-template-path`, 即使模型自带模板 (`git show 35c701ba:miles/utils/arguments.py` 可复现该过时契约); (4) custom-generate 文档把两种受支持的函数形式割裂描述, 未解释 `load_generate_function` 的适配关系。

实现拆解

1. 归档旧页面: 整体删除 `docs/user-guide/rollout-endpoints.md` (270 行) 与 `docs/user-guide/agentic-chat-template.md` (81 行)。前者将 chat-turn 的 token 化职责分配给 SGLang 并要求 agent 提供 prompt token IDs, 属于 pre-session-server 时代的陈旧描述; 后者把 TITO 与整体 agentic 工作流割裂且未覆盖 v2 分支语义。两个旧页面的内容按主题拆分进入两个新页面, 消除同一主题两处维护造成的矛盾。
2. 重建 `agentic-rollout.md` (+227 行): 作为 agentic 路径的唯一权威指南, 整合了如下内容: `run_agent` 钩子契约 (`base_url` 已含 `/sessions/<id>`、`prompt` 恒为 messages 列表、`request_kwargs` 完成 `max_new_tokens` 到 `max_tokens` 的映射)、可选的 `abort` teardown 钩子、token 所有权与 TITO 生命周期 (首轮渲染模板、后续复用 checkpoint 只 token 化后缀)、v1 (线性追加) 与 v2 (append-only 树、`finish_reason=length` 路径不可扩展) 的语义对比、`--tito-model` 家族选择表 (qwen3/qwen35/qwennext/glm47/nemotron3/kimi25/minimax_m25/deepseekv32 等映射)、`--session-message-matcher` 的 replay matching 说明、TITO 暂不支持 VLM 的警告, 以及新模型家族的注册与验证流程 (

verify_chat_template.py 与 verify_session_tito_tokenizer.py 两道检查)。

3. 重建 generate-endpoint.md (+111 行)：聚焦底层 stateless /generate 路径。核心是一张对比厘清两种 generate 函数形式：新形式 `async def generate(input: GenerateFnInput) -> GenerateFnOutput` 与旧形式 `async def generate(args, sample, sampling_params) -> Sample`，并说明 `load_generate_function` (`miles/rollout/inference_rollout/compatibility.py`) 在加载时自动适配旧形式；同时澄清 `MILES_EXPERIMENTAL_ROLLOUT_REFACTOR` 默认关闭、需显式设置，且两种形式在开关 / 关上都能工作。
4. 修正 CLI 契约与导航： `miles/utils/arguments.py` 的 `add_session_arguments` 中 `--use-session-server` 帮助文本移除“必须同时设置 `--chat-template-path`”的过时要求，改为“具名 `--tito-model` 解析其注册模板，`default` 使用 checkpoint 原生或显式模板”； `docs/docs.json` 将 User Guide 导航替换为 `generate-endpoint` 与 `agentic-rollout`，并新增两条永久 redirect (`rollout-endpoints` 到 `generate-endpoint`、`agentic-chat-template` 到 `agentic-rollout`)； `customization.md` 新增 `--custom-agent-function-path` 钩子条目及 `run_agent` 契约，并把 `--session-message-matcher` 交叉链接指向新页面； `cli-reference.md` 新增“Agentic sessions” flag 表格 (`use-session-server`、`tito-model`、`max-seq-len`、`session-server-port`、`session-sample-picker-path`、`session-sample-postprocessor-path` 等)。
5. 配套链接与验证：同步 `index.md`、`deepseek-v3-2.md`、`nemo-gym.md`、`debug.md`、`environments.md`、`README.md` 中的交叉链接。PR body 报告 165 个 agentic/session 测试与 110 个 generate 测试通过、`py_compile` 通过、JSON/ 导航 / 链接 / 过期 slug/Markdown 检查与 `git diff --check` 通过。本次没有新增测试文件，变更以文档与展示字符串为主。

关键文件：

- `docs/user-guide/agentic-rollout.md` (模块 用户指南；类别 docs；类型 documentation；符号 `run_agent`, `abort`)：本 PR 的核心产出：将 OpenAI 消息契约、TITO token 所有权、v1/v2 历史语义、模型家族选择表、replay matching、VLM 限制与验证流程整合为 agentic 路径的唯一权威指南，取代旧的 `agentic-chat-template.md`。
- `miles/utils/arguments.py` (模块 参数解析；类别 source；类型 core-logic；符号 `add_session_arguments`)：全 PR 唯一的源码变更：修正 `--use-session-server` 的 help 文本，移除过时的 `--chat-template-path` 强制要求，使 CLI 契约与具名 `--tito-model` 的注册模板行为保持一致。
- `docs/user-guide/generate-endpoint.md` (模块 用户指南；类别 docs；类型 documentation；符号 `generate`, `_add_arguments`)：重建后的底层 /generate 指南，核心贡献是一张对比厘清新旧两种 generate 函数形式以及 `load_generate_function` 的自动适配关系，消除旧文档把两种形式割裂描述的误解。
- `docs/user-guide/rollout-endpoints.md` (模块 用户指南；类别 docs；类型 deletion；符号 `generate`, `_add_arguments`, `run_agent`, `abort`)：被删除的 270 行旧页面，其对 token 化职责的描述停留在 pre-session-server 时代 (要求 agent 提供 prompt token IDs)，是本次修正的主要症状来源。

- docs/docs.json (模块 文档导航; 类别 config; 类型 configuration) : 文档站导航与重定向配置: 将 User Guide 导航替换为新页面, 并新增两条永久 redirect 保证旧 slug 不断链。
- docs/user-guide/agent-chat-template.md (模块 用户指南; 类别 docs; 类型 deletion) : 被 agent-chat-rollout.md 取代的旧页面 (81 行), 其内容将 TITO 与整体 agent 工作流分离且未覆盖 v2 语义; 删除后通过 redirect 保留旧 slug。
- docs/user-guide/customization.md (模块 用户指南; 类别 docs; 类型 documentation; 符号 run_agent) : 补全 Python 钩子总表: 新增 --custom-agent-function-path 行及其 run_agent 契约说明, 并把 --session-message-matcher 交叉链接指向新页面, 是钩子文档与指南之间的桥接。

关键符号: run_agent, abort, generate, _add_arguments, add_session_arguments

关键源码片段

docs/user-guide/agent-chat-rollout.md

本 PR 的核心产出: 将 OpenAI 消息契约、TITO token 所有权、v1/v2 历史语义、模型家族选择表、replay matching、VLM 限制与验证流程整合为 agent 路径的唯一权威指南, 取代旧的 agent-chat-template.md。

docs/user-guide/agent-chat-rollout.md 中定义的 agent 钩子契约:

```
# 旧文档要求 agent 自行提供 prompt token IDs, 并把 token 化职责归给 SGLang;
# 新契约明确 token 所有权在 session server: base_url 已带会话路径,
# prompt 始终是 messages 列表, Miles 负责把各轮输出对齐到累计 TITO 序列。
async def run_agent(
    base_url: str, # 已包含 /sessions/<id>, 不要手动拼接 session 路径
    prompt, # 输入 sample 的 OpenAI messages 列表
    request_kwargs: dict, # 已映射好 sampling 参数 (max_new_tokens -> max_tokens)
    metadata: dict, # sample 元数据、session 标识与 max_seq_len
    **kwargs,
) -> dict | None: # 返回 dict 会合并进 sample metadata, None 表示无附加
    payload = {"model": "default", "messages": prompt, **request_kwargs}
    await post(f"{base_url}/v1/chat/completions", payload)
    return None
```

文档同时警告: 不要传 --apply-chat-template, 不要设置 TITO 控制字段 (如 logprob_start_len=0, 会破坏前缀缓存); v2 模式下 wrapper 返回 list[Sample], 并显式拒绝 --group-rm、--partial-rollout、--recompute-logprobs-via-prefill。

miles/utils/arguments.py

全 PR 唯一的源码变更: 修正 --use-session-server 的 help 文本, 移除过时的 --chat-template-path 强制要求, 使 CLI 契约与具名 --tito-model 的注册模板行为保持一致。

```
def add_session_arguments(parser):
    # 旧文案要求必须同时设置 --chat-template-path,
    # 但具名 --tito-model (如 qwen3、glm47) 已注册固定模板,
    # checkpoint 也可能原生自带模板, 该要求是过时契约, 会误导用户。
    parser.add_argument(
```

```

"--use-session-server",
nargs="?",
const=True,
default=False,
help="Start a standalone session server for TITO/session support. "
"Requires --hf-checkpoint. A named --tito-model resolves its registered template; "
"--tito-model=default uses the checkpoint-native or explicit --chat-template-path template.
"

"Bare flag (or 'v1') selects the append-only linear v1 server; "
"--use-session-server v2' selects the tree-serving v2 "
"(multi-lineage trajectories, always-branch).",
)
# 相邻的 --tito-model 与 --session-message-matcher 参数保持不变,
# 前者控制前缀 token 复用, 后者控制重放消息匹配策略 (#2240)。

```

评论区精华

该 PR 没有任何 review 评论或讨论线程，两位 reviewer 均直接批准：Zhichenzzz 提交了无文字的 APPROVED，Shi-Dong 给出 **LGTM**。虽然没有显式交锋，但 18 个 commit 记录了一次值得注意的文档组织迭代：replay-matching 小节先被移到 `--tito-model` 选择之后（commit [e1989be2](#)），随后又被移到 TITO 章节末尾（commit [9c2132b4](#)），体现作者对“模型选择 -> 会话语义 -> 重放匹配”阅读顺序的持续调整；另外 commit [41cdb5c](#) 与 [113b54e](#) 记录了 CI 侧 Hugging Face Xet 下载限流的临时 workaround 与移除，最终未带入本 PR 的变更文件。

- 暂无高价值评论线程

风险与影响

- 风险：

1. 重定向依赖：rollout-endpoints.md 与 agentic-chat-template.md 被删除后，站外书签与旧链接全部依赖 docs/docs.json 中新增的两条 redirect，若文档部署流程未同步该配置，旧链接将 404。
2. 深层锚点兼容：页面合并后部分锚点发生变化，例如旧页 `agentic-chat-template#choose-replay-matching` 的标题层级与位置均有调整，虽然新页面保留了同名小节，但无法保证旧页面的全部深层锚点（如 `#pick-your---tito-model`）仍然有效。
3. 文档 - 代码漂移：`--tito-model` 家族映射表与 `tito_tokenizer.py` 的注册、`--use-session-server` 的 help 文本均没有代码测试保护，后续实现变更可能再次造成文档与 CLI 契约脱节——本 PR 恰恰是为了修复这类漂移而来。
4. 运行时影响：唯一源码变更为 `arguments.py` 中 3 行 help 文案，属于展示层修改，不影响参数解析逻辑本身，回归风险极低。- 影响：对用户：agentic rollout 配置者是最直接受益方——旧文档会让用户误以为需要自行完成 token 化并向 SGLang 提供 prompt token IDs，且模型自带模板时仍被迫传 `--chat-template-path`；新文档明确了“消息交换 + Token 所有权归 session server”的边界，并给出 v1/v2 语义、模型家族映射与 replay matching 的选择依据。对系统：除 help 文本外无任何运行时行为变化，无数据、性能或安全影响。对团队：文档结构从“端点视角 + 分散的 chat-template 页”收敛为“generate-endpoint（下层 token 级）与 agentic-rollout（上层消息级）”的清晰二分，

减少了两处重复维护与相互矛盾的来源，并为后续新增模型家族的验证流程（见 issue#712）提供了操作入口。 - 风险标记：删除页面依赖重定向，文档 - 代码同步风险, CLI 帮助文本无测试覆盖，深层锚点兼容风险

关联脉络

- PR #2240 feat(session): add configurable replay matching: 本 PR 在 PR body 与 commit 210520c 中明确提到要整合 #2240 引入的 `--session-message-matcher` 文档：将 'Choose replay matching' 段落移植进重命名后的 `agentic-rollout.md`，并保留旧 slug 重定向。
- PR #2368 fix(rollout): group session v2 leaf samples: `agentic_tool_call` 的 v2 多叶子返回 `list[Sample]` 语义在 `agentic-rollout.md` 中有专门描述，两个 PR 共同构成 session v2 `agentic` 路径的行为与文档闭环。
- PR #2369 fix(rollout): normalize rewards per rollout: session v2 变扇出统计偏差的修复与 `agentic-rollout.md` 中 v2 多叶子样本、批量 reward 的说明相呼应，属于同一功能线。