

PR #1982 完整报告

THUDM/slime

[agent] Fix dropping overlong sample

合并时间: 2026-05-29 21:03

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1982>

执行摘要

- 一句话: 修复 merge_turn_segments 丢弃超长 segment 的 bug
- 推荐动作: 该 PR 是一次轻量的 bugfix, 逻辑清晰, 改动集中。建议快速合并, 无需深入精读。

功能与动机

在 coding_agent_rl 场景中, multi-turn 的 prompt+response 总长度可能超过 max_context_tokens, 但这些 segment 是有效的训练数据, 不应被丢弃。PR 标题和 body 明确指出要修复“dropping overlong sample”的问题。

实现拆解

1. 修改 merge_turn_segments 函数 (slime/agent/trajectory.py) : 移除 max_context_tokens 参数和相关的长度判断逻辑, 只保留 response 非空检查。
2. 更新调用者: 在 slime/agent/adapters/anthropic.py 和 slime/agent/adapters/openai.py 的 finish_session 方法中, 移除传递 max_context_tokens 参数, 直接调用 merge_turn_segments(s.segments)。
3. 更新测试 (tests/test_agent_trajectory.py) : 新增 test_merge_turn_segments_keeps_oversized_segments 测试用例, 验证 oversize 的 segment 不会被丢弃; 同时导入 TurnSegment 和 merge_turn_segments。
4. 更新文档 (examples/coding_agent_rl/README.md) : 更新 --rollout-max-context-len 的说明, 澄清该参数仅作用于生成时的截断, 不再影响轨迹合并 / 导出时的 segment 丢弃。

关键文件:

- slime/agent/trajectory.py (模块 轨迹合并; 类别 source; 类型 core-logic; 符号 merge_turn_segments) : 核心修复, 修改了 merge_turn_segments 函数, 移除了 overlong 过滤逻辑。
- tests/test_agent_trajectory.py (模块 测试; 类别 test; 类型 test-coverage; 符号 test_merge_turn_segments_keeps_oversized_segments) : 新增测试用例验证 oversize segment 不会被丢弃。
- slime/agent/adapters/anthropic.py (模块 Anthropic 适配器; 类别 source; 类型 core-logic; 符号 finish_session) : 适配器调用处移除 max_context_tokens 参数。

- `slime/agent/adapters/openai.py` (模块 OpenAI 适配器; 类别 `source`; 类型 `core-logic`; 符号 `finish_session`) : 适配器调用处移除 `max_context_tokens` 参数。
- `examples/coding_agent_rl/README.md` (模块 文档; 类别 `docs`; 类型 `documentation`) : 文档说明更新, 澄清参数作用范围。

关键符号: `merge_turn_segments`, `finish_session`

评论区精华

该 PR 没有 review 评论和 discussion, 因此没有可提炼的讨论内容。

- 暂无高价值评论线程

风险与影响

- 风险: 风险较低。移除 `max_context_tokens` 过滤后, 所有非空 `segment` 都会被保留, 可能增加训练数据量, 但不会引发错误。需要确保训练流程对 `segment` 长度没有隐式依赖。此外, 如果其他未发现的调用点依赖了旧行为, 可能需要额外清理。
- 影响: 影响范围包括 `agent` 模块的 `merge_turn_segments` 核心函数及其所有调用者 (Anthropic 和 OpenAI adapter), 以及测试和文档。对用户的影响是: 之前因过长而被丢弃的训练样本现在会被保留, 可能提升模型训练效果。对系统的影响是略微增加训练数据量, 但无性能副作用。
- 风险标记: 核心路径变更

关联脉络

- PR #1963 Fix trajectory merging logic: 同样涉及轨迹合并逻辑的修复, 与本次改动相关。