

PR #3086 完整报告

radixark/miles

dashboard: pin the anatomy lane to a duplicated index's first leaf

合并时间: 2026-09-04 08:42

原文链接: <http://prhub.com.cn/radixark/miles/pull/3086>

执行摘要

本 PR 是 PR#2881 的后续修复，解决了 Dashboard 轨迹面板因支持多样本叶子而引入的数据映射不一致 bug。通过过滤 `rows` 数组，仅使用第一个叶子（`sample_occurrence` 为 0）的数据构建面板标签映射，确保了用户看到的标签信息与点击后查看的 token 视图内容一致。变更范围极小，仅涉及 `views_rollout.js` 一个文件的一行核心逻辑修改和注释添加。

功能与动机

此变更源于 PR#2881 引入的 TITO v2 多叶子样本支持。如 PR body 所述，一个 `sample_index` 现在可以携带多个摘要行（每个 `sample_occurrence` 对应一行）。原有实现中 `rowsByIndex` 使用 `new Map(rows.map(...))` 构建，当存在多个相同 `sample_index` 的行时，Map 的“最后写入者胜出”特性会导致面板（lifecycle lane）显示最后一个叶子的奖励 / 长度等信息，但点击该行时，`onClickSample` 函数总是打开 `sample_occurrence: 0` 的 token 视图，造成了“标签显示 A，内容却是 B”的用户体验混淆。

实现拆解

1. 变更入口与核心逻辑：修改 `miles/dashboard/static/views_rollout.js` 中 `samplesTab` 函数内传递给 `createAnatomy` 组件的 `rowsByIndex` 参数构建逻辑。
2. 数据过滤与映射：将 `rowsByIndex` 的构建从 `new Map(rows.map(...))` 改为 `new Map(rows.filter((r) => (r.sample_occurrence ?? 0) === 0).map(...))`。此 `filter` 操作确保了 Map 中每个 `sample_index` 对应的行是 `sample_occurrence` 为 0（或未设置，视为 0）的那一行，即该索引下的第一个叶子样本。
3. 意图文档化：在过滤代码前添加了四行注释，清晰地解释了 TITO 索引可能存在多个叶子、生命周期面板应固定展示第一个叶子以保持一致性、以及页面下方的 `samples` 表格仍可访问所有行的设计考量。
4. 验证与配套：通过 `node --check` 语法检查。因前端无测试框架，未增加自动化测试，但变更逻辑简单且自包含，通过代码审查验证。

`miles/dashboard/static/views_rollout.js`

这是本次修复的唯一变更文件。其核心逻辑是构建 `createAnatomy` 面板所需的 `rowsByIndex` 映射表，修复了因数据模型变化（多叶子样本）导致的显示与交互不一致问题。

```
// 从 rows 数组构建行数据的索引映射 (Map)，供 anatomy 面板使用
rowsByIndex: new Map(
  rows
```

```
// 过滤: 仅保留 sample_occurrence 为 0 (或未设置视为 0) 的行。  
// 因为一个 sample_index 可能对应多个叶子 (occurrence > 0) 。  
// 动作执行生命周期面板展示的是所有叶子共享的执行过程,  
// 我们固定展示并打开第一个叶子 (occurrence 0) 以保持一致性。  
// 页面下方的 samples 表格仍可访问所有 (index, occurrence) 的行。  
.filter((r) => (r.sample_occurrence ?? 0) === 0)  
.map((r) => [r.sample_index, r])  
)
```

评论区精华

- [claude\[bot\]](#) 对修复正确性的核心验证: “onClickSample always opens sample_occurrence: 0”, 因此新的过滤逻辑 (保留 occurrence=0 行) 精确地解决了 last-one-wins 带来的标签与点击内容不匹配问题。
- [claude\[bot\]](#) 风险评估: “Level of scrutiny Low. The change is small, self-contained, and directly matches its stated purpose.” 并确认无安全风险。
- [Zhichenzzz](#) 的简洁批准: “LGTM”。

风险与影响

- 风险: 风险极低。变更仅影响 Dashboard 的 Trajectories 面板显示逻辑, 不改变核心数据流。主要风险点在于前端模块缺乏自动化测试覆盖 (risk_flags 中已标出), 未来相关代码变更可能意外破坏此行为。
- 影响: 直接影响使用 Dashboard 观察训练轨迹的用户体验, 修复了标签与内容不一致的 bug, 提升了数据观测的准确性。对后端系统、训练性能、安全性无任何影响。

关联脉络

本 PR 明确是 PR#2881 (“Dashboard: handle multiple TITO leaves per sample index”) 的后续修复 (follow-up)。PR#2881 为 Dashboard 引入了区分多样本叶子的能力, 但遗留了 [anatomy lane](#) 中数据映射的 last-one-wins 问题。本 PR 完整了该功能的实现, 体现了功能开发后细致回归和修复的良好实践。此修复对于维护 Dashboard 作为关键观测工具的数据一致性至关重要。