

PR #2490 完整报告

radixark/miles

dashboard: keep sample status/reward chips visible on the Tokens tab

合并时间: 2026-08-13 06:57

原文链接: <http://prhub.com.cn/radixark/miles/pull/2490>

PR 2490 分析: Tokens 标签页保留样本状态与奖励 chips

执行摘要

本 PR 修复了 Rollouts 样本视图中切换至 Tokens 标签后样本状态 (`completed`) 与奖励 (`reward X`) chips 消失的问题。核心做法是把 chips 渲染从 `renderConversation` 中拆分为独立导出函数 `renderSampleChips`, 并在 `renderTokens` 中将其渲染到标签栏上方。改动仅涉及 2 个前端静态文件, 共 16 行新增、7 行删除, 无 API、schema 或 Python 变更。

功能与动机

样本的 status 与 reward 是理解 token 视图的核心上下文。正如 PR body 所述:

The reward is usually the thing you are trying to explain when you open the token view: you go there to see which tokens drove a 0 reward, or where the importance ratio blew up on a sample that scored well.

此前 chips 内嵌在 Conversation 面板中, 切换到 Tokens 标签时整个面板连同 chips 一起被替换, 用户必须来回切换才能确认奖励数值。同样地, `truncated` 与 `completed` 样本的解读方式完全不同, 状态常驻同样重要。

实现拆解

1. 拆分 chips 渲染 (`miles/dashboard/static/conversation.js`) - 新增导出函数 `renderSampleChips(row)`, 返回裸的 `.controls` 行, 包含 status chip 与 (存在时的) reward chip。- `renderConversation(row)` 精简为只渲染面板标题与消息卡片, 不再内嵌 chips。- 保留了 `row.reward !== null && row.reward !== undefined` 的显式判断, 确保真实 0.0 奖励仍能渲染。
2. 接入标签视图 (`miles/dashboard/static/views_tokens.js`) - 导入 `renderSampleChips`。
- 在 `renderTokens` 的 `view.replaceChildren(...)` 中, 将 `renderSampleChips(conversationRow)` 置于 `panels` 之后、`tabs` 之前, 使 chips 位于两个面板之外、标签栏上方, 切换标签时不会被替换。
3. 保持无 sidecar 路径不变 - `conversationRow === null` 的分支不渲染标签栏, 只渲染 token 面板, 行为与改动前一致。
4. 验证方式 - 仓库无 JS 测试 harness (无 `package.json`, Python 测试也不加载静态文件), 因此通过真实 dump 管道生成 dummy dump、`python -m miles.dashboard.serve` 提供服务、Playwright 驱动真实 Chrome 验证 chip 文本与 `#view` 子元素顺序。

miles/dashboard/static/conversation.js

核心改动：将 chips 渲染从 renderConversation 中拆出为独立导出 renderSampleChips，返回裸 .controls 行，使 chips 能在标签栏上方独立渲染。

```
// miles/dashboard/static/conversation.js
import { el, fmtNum } from "../app.js";

// sample outcome (status + reward) from the trajectory sidecar。
// 该函数独立于对话面板，因为样本结果描述的是样本本身，
// 而不是某一次渲染，切换 Tokens 标签时它必须保持可见。
export function renderSampleChips(row) {
  const chips = [el("span", { class: "chip" }, [row.status])];
  // 显式判断 null/undefined，避免把真实的 0.0 奖励当成 falsy 吞掉
  if (row.reward !== null && row.reward !== undefined) {
    chips.push(el("span", { class: "chip" }, [`reward ${fmtNum(row.reward)}`]));
  }
  // 返回裸的 controls 行，由调用方决定放置位置（标签栏上方）
  return el("div", { class: "controls" }, chips);
}

// 对话面板只负责标题与消息卡片，不再内嵌 chips
export function renderConversation(row) {
  return el("div", { class: "panel" }, [
    el("h3", {}, ["Conversation"]),
    ...row.messages.map(messageCard),
  ]);
}
```

miles/dashboard/static/views_tokens.js

变更接入点：将 renderSampleChips 渲染到标签栏上方，使 chips 在两种标签下均可见；同时更新导入。

```
// miles/dashboard/static/views_tokens.js
import { renderConversation, renderSampleChips } from "../conversation.js";

export async function renderTokens(view, meta, route) {
  // ... 前置逻辑：conversationRow 为 null 时直接渲染 token 面板并返回

  const conversationPane = renderConversation(conversationRow);
  const tabs = el("div", { class: "tabs" });
  const body = el("div");
  const select = (name) => {
    tabs.replaceChildren(
      ...["conversation", "tokens"].map((tab) =>
        el("button", { class: tab === name ? "active" : "", onclick: () => select(tab) }, [
          tab[0].toUpperCase() + tab.slice(1),
        ]),
    ),
  );
};
```

```
body.replaceChildren(name === "conversation" ? conversationPane : tokensPane);
if (name === "tokens") startTokens();
};
// chips 位于标签栏上方、两个面板之外,
// 因此切换标签时不会被替换掉, 状态与奖励始终可见
view.replaceChildren(...panels, renderSampleChips(conversationRow), tabs, body);
select("conversation");
}
```

评论区精华

两位 reviewer ([yueming-yuan](#)、[nblintao](#)) 均直接 APPROVED, 无评论内容, 无设计争议。

风险与影响

- 回归风险: conversationRow === null 分支不受影响; 若未来有其他调用方依赖 renderConversation 内部 DOM 结构, 可能受影响, 目前无此类调用。
- 样式风险: chips 使用 .controls 类, 从面板内移到标签栏上方, 视觉位置变化已在 PR 中附前后对比图确认。
- 测试缺口: 仓库无 JS 测试 harness, 本改动没有自动化测试覆盖, 只能依赖手动 /Playwright 验证。这是 dashboard 前端持续存在的技术债。
- 影响范围: 仅 dashboard 前端展示, 无性能、安全或兼容性风险。

关联脉络

本 PR 是 dashboard 观测性持续改进的一环。近期相关改动包括 PR#2476 (token 视图 log-probs 清零修正)、PR#2027 (advisory v2 健康告警优先)、PR#2023 (dp-aware 引擎指标), 表明 dashboard 的样本级分析与 token 级分析正在逐步完善, 本 PR 补齐了样本上下文在 token 视图中的可读性。