

PR #1959 完整报告

THUDM/slime

rollout: add forge_load to replay dumped rollouts with slang alive

合并时间: 2026-05-27 15:47

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1959>

执行摘要

- 一句话: 新增 `forge_load` 重放 rollout dump 且保持 `sglang` 运行
- 推荐动作: 建议负责内存测试和 rollout 框架的团队成员精读此 PR, 了解 `forge_load` 的设计细节及如何用于内存测试。值得关注的设计决策: `eval` 没有 fallback 以防止静默数据错配; 不修改 `sample.rollout_id` 以保持分组正确; 与 `--load-debug-rollout-data` 的战略分工。

功能与动机

在进行长上下文训练的内存测试时, 需要运行完整的 `sglang` 推理流程以准确测量 GPU 内存峰值。原有调试模式 `--load-debug-rollout-data` 会设置 `skip_slang=True` 和 `debug_train_only=True`, 从而关闭 `sglang` 等组件, 无法满足内存测试需求。本 PR 通过新增 `--load-forge-rollout-data` 参数和配套的 `rollout` 函数, 实现在不影响 `sglang` 运行的前提下重放已保存的 rollout 数据。这是动机在 PR body 中的明确表述: “Unlike `--load-debug-rollout-data` this does NOT flip `skip_slang/debug_train_only`”。

实现拆解

1. 新增 `slime/rollout/forge_load.py` 核心模块: 实现 `_resolve_path` 和 `generate_rollout` 两个函数。`_resolve_path` 根据 `args.load_forge_rollout_data` 解析路径, 支持文字路径和 `{rollout_id}` 模板路径, 训练时特定文件缺失会回退到 `0.pt`; `eval` 无回退。`generate_rollout` 是符合 `rollout` 函数接口的入口, 根据 `evaluation` 标志加载 `eval` 或 `train` 数据, 返回对应的 `RolloutFnEvalOutput` 或 `RolloutFnTrainOutput`, 并保持 `sample.rollout_id` 不变避免分组错误。
2. 在 `slime/utils/arguments.py` 中添加参数声明: 在 `add_debug_arguments` 函数内新增 `--load-forge-rollout-data` 参数, 类型 `str`, 默认 `None`, 提供详细 `help` 说明与 `--load-debug-rollout-data` 的区别。
3. 集成到 rollout 框架: 用户通过 `--rollout-function-path` `slime.rollout.forge_load.generate_rollout` `--load-forge-rollout-data` 启用。框架自动调用 `generate_rollout` 替代真实生成, 但 `sglang`、路由、权重同步继续运行。

关键文件:

- `slime/rollout/forge_load.py` (模块 加载器; 类别 `source`; 类型 `new-module`; 符号 `_resolve_path`, `generate_rollout`): 核心新增模块, 包含路径解析函数 `_resolve_path` 和 `rollout` 函数 `generate_rollout`, 是实现从 dump 重放 rollout 并保持 `sglang` 运行的核心逻辑

辑。

- slime/utils/arguments.py (模块 参数解析; 类别 source; 类型 configuration) : 修改此文件以添加 --load-forge-rollout-data 参数, 是用户调用该功能的入口。

关键符号: _resolve_path, generate_rollout

关键源码片段

slime/rollout/forge_load.py

核心新增模块, 包含路径解析函数 _resolve_path 和 rollout 函数 generate_rollout, 是实现从 dump 重放 rollout 并保持 sglang 运行的核心逻辑。

```
import logging
import os
from pathlib import Path
import torch
from slime.rollout.base_types import RolloutFnEvalOutput, RolloutFnTrainOutput
from slime.utils.types import Sample
```

```
logger = logging.getLogger(__name__)
```

```
def _resolve_path(args, rollout_id: int, evaluation: bool) -> str | None:
```

```
    '''Resolve the forge rollout data file path.
```

```
    Supports literal path (reused across rollout_ids) and template path
    (with {rollout_id} placeholder). Training path falls back to 0.pt
    if the exact file is missing; evaluation path has no fallback.
    ...
```

```
    tpl = getattr(args, 'load_forge_rollout_data', None)
```

```
    if not tpl:
```

```
        raise RuntimeError('--load-forge-rollout-data must be set')
```

```
    # 文字路径模式 (不含 {rollout_id}) 下, eval 直接返回 None
```

```
    if evaluation and '{rollout_id}' not in tpl:
```

```
        return None
```

```
    rid_str = ('eval_' if evaluation else '') + str(rollout_id)
```

```
    path = tpl.format(rollout_id=rid_str)
```

```
    if os.path.exists(path):
```

```
        return path
```

```
    # 训练路径若不存在, 回退到 0.pt
```

```
    if not evaluation:
```

```
        fallback = tpl.format(rollout_id='0')
```

```
        if os.path.exists(fallback):
```

```
            logger.info('forge_load: %s missing, falling back to %s', path, fallback)
```

```
            return fallback
```

```
    return None
```

```
def generate_rollout(args, rollout_id, data_source, evaluation: bool = False):
```

```
    '''Rollout function that loads pre-generated samples from disk.
```

Called by the training framework via `--rollout-function-path`.

...

```
path = _resolve_path(args, rollout_id, evaluation)
```

```
if evaluation:
```

```
    if path is None:
```

```
        logger.info('forge_load: no eval dump found; returning empty eval result')
```

```
        return RolloutFnEvalOutput(data={})
```

```
    logger.info('forge_load: loading eval samples from %s', path)
```

```
    blob = torch.load(path, weights_only=False)
```

```
    samples = [Sample.from_dict(s) for s in blob['samples']]
```

```
    reward_key = args.eval_reward_key or args.reward_key
```

```
    rewards = [
```

```
        s.reward if (not reward_key or s.reward is None) else s.reward[reward_key]
```

```
        for s in samples
```

```
    ]
```

```
    return RolloutFnEvalOutput(data={
```

```
        'forge_eval': {
```

```
            'rewards': [r if r is not None else 0.0 for r in rewards],
```

```
            'truncated': [s.status == Sample.Status.TRUNCATED for s in samples],
```

```
            'samples': samples,
```

```
        }
```

```
    })
```

```
# 训练路径：必须存在否则报错
```

```
if path is None:
```

```
    raise RuntimeError(
```

```
        f'forge_load: no dump found for rollout_id={rollout_id} '
```

```
        f'--load-forge-rollout-data={args.load_forge_rollout_data!r}'
```

```
    )
```

```
logger.info('forge_load: loading samples from %s', path)
```

```
blob = torch.load(path, weights_only=False)
```

```
samples = [Sample.from_dict(s) for s in blob['samples']]
```

```
# 保持 sample.rollout_id 不变，避免分组崩溃
```

```
return RolloutFnTrainOutput(data={
```

```
    'samples': samples,
```

```
    'policy_clip_eps': args.clip_eps,
```

```
    # ... 其他训练参数
```

```
})
```

评论区精华

本 PR 未产生公开讨论，设计决策在 PR body 和代码注释中已有明确说明。主要设计点包括 literal vs template 路径模式、训练 fallback 仅针对训练路径等。

- 无 review 讨论 (other): 无争议，已合并。

风险与影响

- 风险：

1. 安全隐患: `torch.load` 使用 `weights_only=False`, 允许任意 pickle 序列化, 存在代码执行风险。但该功能面向开发 / 测试环境, 路径由用户显式指定, 风险可控。
2. 数据一致性问题: 训练路径在主文件缺失时回退到 `0.pt`, 可能导致跨 rollout 分组的数据重复或与预期不匹配 (例如预期加载不同批次的数据但实际重复使用了 `0.pt`)。eval 路径无回退以避免静默错误。
3. 性能: 磁盘 I/O 可能增加延迟, 但相比真实生成已显著降低。
4. 兼容性: 与现有 rollout 框架完全正交, 不影响其他功能。
 - 影响: 用户: 提供新的内存测试能力, 可通过该功能在保持 `sglang` 运行的情况下重放 rollout, 准确测量 GPU 内存使用。不影响现有使用 `--load-debug-rollout-data` 或真实生成的用户。系统: 新增一个小模块, 无依赖变更。团队: 需要维护新模块, 注意与 debug rollout 数据的区别, 特别是路径格式和 fallback 行为。
 - 风险标记: `torch.load` 使用 `weights_only=False` 存在 pickle 安全隐患, 训练 fallback 到 `0.pt` 可能导致数据不匹配

关联脉络

- 暂无明显关联 PR