

PR #1919 完整报告

THUDM/slime

add critic wandb config

合并时间: 2026-05-18 18:55

原文链接: <http://prhub.com.cn/THUDM/slime/pull/1919>

执行摘要

- 一句话: 为 critic 角色添加独立的 W&B 配置记录
- 推荐动作: 此 PR 代码清晰, 改动集中, 无显著风险。推荐合并后同步更新用户文档, 说明 W&B 配置记录的变更。

功能与动机

在 PPO 训练中, actor 和 critic 使用独立的 Megatron 配置, 但之前的 W&B 记录仅将全局 args.__dict__ 作为配置, 导致 critic 的参数被遗漏或混淆。此 PR 旨在实现角色隔离的配置日志, 便于后续对比和调试。

实现拆解

1. 提取配置字典: 新增 _args_to_config_dict 函数, 将 args 对象转为字典 (通过 deepcopy), 解耦配置获取逻辑。
2. 角色特定参数提取: 新增 _get_role_args_for_logging 函数, 当 args.megatron_config_path 存在时, 调用 parse_megatron_role_args 获取指定角色的 args 对象 (如 critic), 否则返回全局 args。
3. 前缀化配置键: 新增 _prefix_config_keys 函数, 将配置字典的所有键添加前缀 (如 "critic/"), 避免与 actor 配置冲突。
4. 主配置函数增强: 修改 _compute_config_for_logging, 在 use_critic=True 时, 提取 critic 参数并添加到输出中。
5. 次级配置函数: 新增 _compute_secondary_config_for_logging, 根据 role 参数返回带前缀或不带前缀的配置, 供 init_wandb_secondary 使用。
6. 初始化调用修改: 修改 init_wandb_secondary 签名, 接受 role 参数, 并在 wandb.init 的 config 中使用新函数; 修改 actor.py 中 init_tracking 调用, 传递 role=role, 使 critic 进程能够记录正确的配置。

关键文件:

- slime/utils/wandb_utils.py (模块 W&B 工具; 类别 source; 类型 core-logic; 符号 _args_to_config_dict, _prefix_config_keys, _get_role_args_for_logging, _compute_secondary_config_for_logging): 核心变更文件: 新增 3 个辅助函数并修改 2 个主要函数, 实现 critic 配置的隔离记录。

- slime/backends/megatron_utils/actor.py (模块 训练引擎; 类别 source; 类型 core-logic)
: 修改一处调用: 将 role 参数传递给 init_tracking, 使 critic 进程能够正确初始化次级 W&B。

关键符号: _compute_config_for_logging, _args_to_config_dict, _prefix_config_keys, _get_role_args_for_logging, _compute_secondary_config_for_logging, init_wandb_secondary

关键源码片段

slime/utils/wandb_utils.py

核心变更文件: 新增 3 个辅助函数并修改 2 个主要函数, 实现 critic 配置的隔离记录。

```
def _compute_config_for_logging(args):
    """为主 W&B 进程构建配置字典, 包含 critic 配置 (如果启用) """
    output = _args_to_config_dict(args)

    whitelist_env_vars = [
        "SLURM_JOB_ID",
    ]
    output["env_vars"] = {k: v for k, v in os.environ.items() if k in whitelist_env_vars}

    if getattr(args, "use_critic", False):
        # 获取 critic 专用的 args 对象, 并添加 critic/ 前缀
        critic_args = _get_role_args_for_logging(args, role="critic")
        output.update(_prefix_config_keys(_args_to_config_dict(critic_args), "critic"))

    return output


def _args_to_config_dict(args):
    """将 args 转成字典, 便于序列化到 W&B"""
    return deepcopy(args.__dict__)


def _prefix_config_keys(config, prefix):
    """给配置字典的每个键添加前缀, 例如 'critic/'"""
    return {f"{prefix}/{key}": value for key, value in config.items()}


def _get_role_args_for_logging(args, role):
    """根据角色获取对应的 args 对象, 用于记录该角色的专用配置"""
    if getattr(args, "megatron_config_path", None) is None:
        return args

    from slime.utils.arguments import parse_megatron_role_args

    return parse_megatron_role_args(args, args.megatron_config_path, role=role)
```

```

def _compute_secondary_config_for_logging(args, role=None):
    """为次级 W&B 进程构建配置，角色为 critic 时添加前缀"""
    config = _args_to_config_dict(args)
    if role == "critic":
        return _prefix_config_keys(config, "critic")
    return config

def init_wandb_secondary(args, role=None):
    """初始化次级 W&B 进程（例如 critic），记录角色特定的配置"""
    wandb_run_id = getattr(args, "wandb_run_id", None)
    if wandb_run_id is None:
        return
    # ... 省略登录和 settings 部分 ...
    init_kwargs = {
        "id": wandb_run_id,
        "entity": args.wandb_team,
        "project": args.wandb_project,
        "config": _compute_secondary_config_for_logging(args, role=role),
        "resume": "allow",
        "reinit": True,
        "settings": wandb.Settings(**settings_kwargs),
    }
    wandb.init(**init_kwargs)

```

评论区精华

无 review 评论或讨论，PR 由作者直接合并。

- 暂无高价值评论线程

风险与影响

- 风险：
 1. 配置键前缀可能影响现有 W&B 面板：actor 配置仍为无前缀键，critic 配置以 critic/ 为前缀，现有 actor 日志不受影响，但用户需要更新面板查询逻辑以包含 critic 键。
 2. parse_megatron_role_args 的导入为延迟导入，若该函数不存在或签名变化可能导致运行时错误。
 3. 新增函数可能增加少量内存开销（deepcopy args），但影响可忽略。- 影响：影响范围限于使用 critic（PPO）训练且启用 W&B 的用户。开启 use_critic 后，W&B 配置中会额外出现 critic/ 开头的键。对于仅使用 actor 的训练，行为不变。- 风险标记：配置键前缀可能影响现有 W&B 面板，延迟导入可能引入运行时错误

关联脉络

- PR #1866 Rename critic config to megatron config: 重构了 critic/actor 配置的解析方式（引入 parse_megatron_role_args），本 PR 依赖该能力来提取角色特定参数。