

MoonBit 开源大赛 · 项目申报书

国产基础软件生态开源大赛 (MGPIC 2026)

项目名称: moon-diff

参赛赛道: 基础软件生态 · 库开发 (MoonBit 标准库及生态工具)

负责人: yuzhiblue (个人参赛)

开发语言: MoonBit

开源协议: Apache License 2.0

仓库地址: <https://github.com/yuzhiblue/moon-diff>

一、项目简介

moon-diff 是一个面向 MoonBit 语言的文本 diff 与 patch

库。它计算两段序列之间的差异, 并提供统一 diff (diff -u 风格) 的渲染与回放, 以及 Git 风格 diff、二进制文件差异、语义 JSON Patch、三路合并与多文件树 diff 能力。项目以纯 MoonBit、零外部依赖实现, 旨在为 MoonBit 生态补上版本控制与文本比对领域的基础构件。

二、技术亮点

- 五种差异算法: 经典 LCS 动态规划 (diff)、Myers $O(ND)$ 最小编辑脚本 (myers_diff)、Patience (patience_diff)、Histogram (histogram_diff), 以及线性空间 Hirschberg 算法 (diff_linear, 仅 $O(|a|+|b|)$ 内存)。
- 标准统一 diff 与 Git 风格: to_unified 输出 GNU diff -u 风格补丁; git_diff_text 输出 Git diff --git / index <sha> 头部, binary_diff 输出二进制文件差异; 均可通过 apply_unified / apply_unified_fuzzy 还原, 兼容 Git / patch 工具链。
- 补丁应用与逆转: apply_unified 正向还原, apply_unified_reverse (patch -R) 反向还原, apply_unified_fuzzy 容错 offset / fuzz。
- 语义 JSON diff 与 RFC 6902 应用: json_diff_text 生成 JSON Patch (对象顺序无关), apply_json_patch / apply_json_patch_text 回放, 闭环比对闭环。
- 三路合并与多文件树 diff: merge3 实现 diff3 风格合并 (含冲突标记与 ours/theirs 解析); diff_trees / render_tree_patch / apply_tree_patch 生成并消费 Git 风格多文件补丁 (含重命名检测)。
- Unicode 友好的文本比对: tokenize_unicode 按 CJK 字符 / 词 / 标点切分, 使中文、日文、韩文逐字比对; ratio 基于 LCS 计算相似度 (类似 Python difflib.ratio), 可用于查重与排序。
- 零依赖、纯 MoonBit 实现: 基于泛型 Array[T] (T: Eq / Show), 可对任意序列 (行、词元、AST 节点) 做差异计算。
- 工程完备: 自带 GitHub Actions CI (moon build + moon test), 54 项测试覆盖五算法重建一致性、最小编辑距离交叉验证、JSON / tree 补丁往返、Unicode tokenizer、ratio 边界、前后缀公共段剪枝、git diff --stat 摘要、忽略空格/大小写比较等; 已发布到 MoonCakes (v0.2.2, mooncakes.io/package/yuzhiblue/moon-diff)。

三、应用场景

- 代码版本控制与差异展示 (类 git diff 的基础能力)
- 文本 / 文档比对工具、在线协作文档的差异高亮

- 补丁 (patch) 系统的生成与应用
- 任意结构化数据 (含 JSON) 的变更检测与同步
- 多文件项目的批量差异与合并 (tree diff / 3-way merge)

四、当前进度与后续计划

核心功能与扩展能力已全部完成: diff / myers_diff / to_unified / apply_unified / diff_lines / 五种算法 / 三路合并 / 语义 JSON diff 与 apply / 多文件 tree diff / Git 风格与二进制 diff / Unicode 友好 tokenizer / 相似度 ratio, 均已实现、测试并发布到 MoonCakes。GitHub Actions CI 在每次 push/PR 实跑 moon build + moon test 且全绿。后续计划: 基于 moonbitlang/x/fs 的 CLI 文件 I/O (当前为参数 / 标准输入方式)。

五、参赛承诺

本作品为作者原创, 以 Apache-2.0

协议开源, 可公开访问、可构建、可测试, 满足大赛验收要求, 并接受评审与社区反馈。

申报日期: 2026-07-09 | 负责人: yuzhiblue | 仓库: github.com/yuzhiblue/moon-diff