dsh-memory

给 DeepSeek Harness 装上跨会话项目记忆



厦门大学《学科实践三》期末工程
原生 Cordis 插件研究与开发
单人组 · AI 协同开发 · MIT 开源

陈悦祺 · 学号 22920242203194

鲸鱼讲解员 · 项目原创非官方形象

v0.5.1 / v0.2.2 · 历史评测 v0.4.1

01 需要跨会话维护的项目知识

包管理器、发布规则与长期偏好会随项目演进



换会话，信息断开

上次约定“依赖统一用 pnpm”，新会话却看不到；
只能重新说明，或让助手在缺少依据时拒答。

写入文件，需要维护

AGENTS.md 适合稳定规则；如果 npm 已改为 pnpm，
文件没有同步，新会话仍会看到旧约定。

让变化自动沉淀

从对话提取值得保留的信息，更新项目记忆；
新会话只带入相关记录，不重放整段聊天。

工程问题：自动更新能带来多少收益，需要付出多少额外调用成本？

02 工程范围与设计取舍



宿主基线：本机 DeepSeek Harness 0.2.0-rc.2

复用宿主能力

通过原生接口使用模型、工具和存储，不另建记忆服务；
实现更集中，但需要适配 dsh 的接口与生命周期。

轮末尝试保存

轮末钩子允许等待抽取完成，退出事件只做兜底；
减少进程先退出造成的遗漏，也会增加轮末等待。

围绕一条主线验证

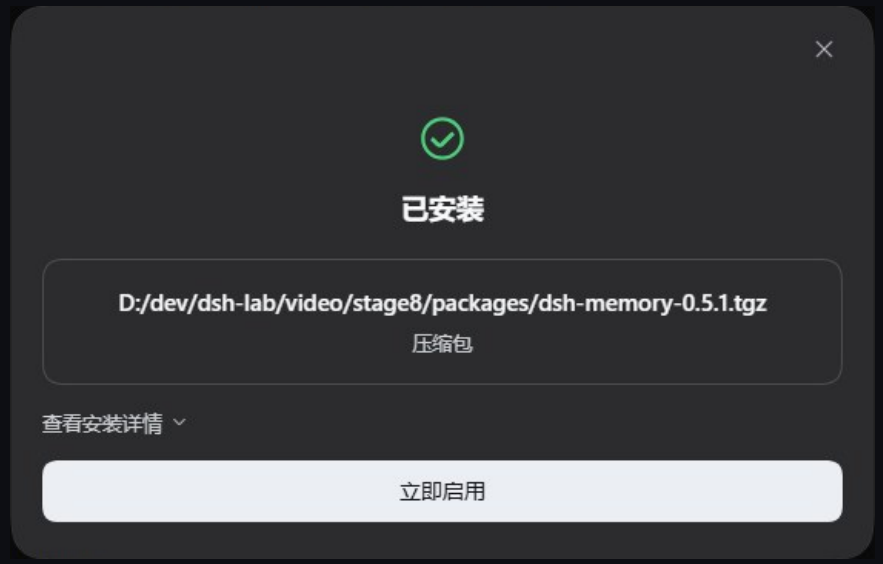
重点解决“保存—更新—跨会话使用”，并测量成本；
dsh-commit 仅作为生成提交说明的配套插件。

本次对照只比较无插件、冻结 AGENTS.md 与 dsh-memory，不对其他产品作能力排名。

03 真实使用：安装后，让新会话接上项目约定

v0.5.1 官方 dsh 网页录屏截图；裁切原始画面，未改写界面输出

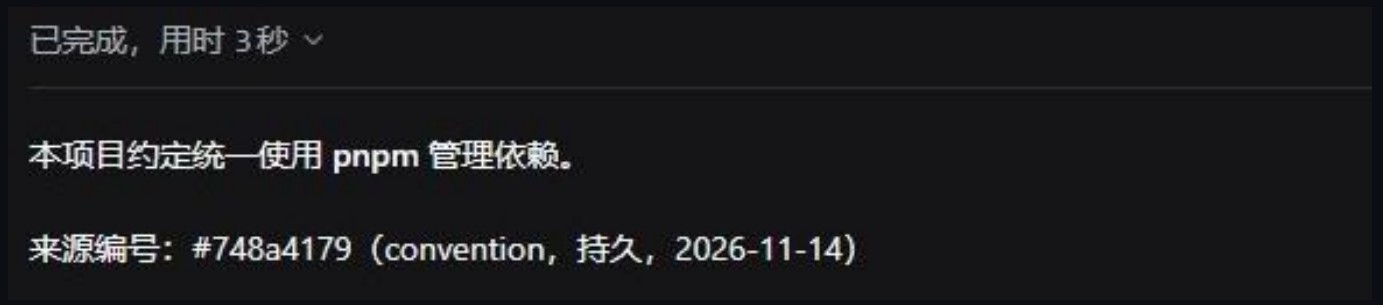
① 添加插件 → 安装成功 → 启用



本次演示安装本地已构建的 tgz 包。

② 保存约定 → 换会话提问

上一会话保存“统一使用 pnpm”；
新会话只询问规则，不再提供答案，也不调用工具。



回答带回 #748a4179，可定位到已保存的条目。

完整 Demo 约 138 秒：还包括设置长期记忆、查看和取消。

04 插件架构：抽取、治理、再注入



记忆保存在模型外部；不训练模型，也不修改模型权重

① 从对话抽取

轮末调用宿主模型，把事实、决策、约定与偏好整理为结构化候选；不直接保存整段聊天。

② 治理后落盘

合并重复项、处理新旧冲突与审批状态，通过官方存储域按工作区保存，关闭会话后仍可读取。

③ 按问题注入

新会话按工作区、相关性和容量筛选记录，放入系统上下文并附来源编号，再由模型回答。

管理入口：查看 / 删除 / 手动提取 / 批准 / 设置或取消持久保留

模型负责理解语义，代码负责存储与可见性边界；抽取失败不推进处理游标。

05 可靠性修复与验收

v0.5.1：已有会话补挂注入；卸载关闭存储域，支持重新启用



新环境也能安装

prepare 在安装源码包时生成运行入口；
用干净 Git 安装复核，避免开发机已有构建掩盖问题。

热安装也能生效

安装前已创建的会话不会再触发“新建”事件；
v0.5.1 为现有 Agent 补挂注入，并去重、清理卸载资源。

更新不能丢掉规则

npm 改 pnpm 时，先写入新事实，再让旧事实退役；
若需要审批，则等整批替代项获准后再切换。

58 + 7 个插件回归测试通过；评测器另有 19 个离线自检。

另有 180/365 天模拟、旧数据兼容与真实网页验证；测试数不等于真实任务准确率。

06 记忆更新与可管理性



内容类别与保留策略分开：未再提及，不代表约定失效

普通记忆

事实 14 天、约定/决策 30 天、
偏好 45 天：有效分数减半，
不是删除一半文字。

长期约定：用户指定

memory_pin 设置持久保留，
豁免时间衰减；
仍可更新、替代或删除。

取消持久：恢复规则

保留内容，不重置计时；
按原时间计算，可能立即
停止自动注入。

实际设置

已完成：#748a4179 已设为持久保留（pinned=true），其他条目未改动。

实际取消

已取消该约定的长期保留：`#748a4179` 现在是一条普通记忆，不再是持久记忆，会按原来的时间继续衰减。

这条记忆内容保留下来，没有删除，仍然是：本项目依赖统一使用 pnpm 管理。

注意：取消持久保留后，它可能不再被自动注入到对话里。

持久仍受相关性和容量限制；14/30/45 天是尚未长期校准的工程参数。

07 实验设计

20 个固定合成主案例，每例每组重复 3 次；另有 4 个独立治理案例



组别	新上下文可见资料
无插件	统一回答指令 + 问题
冻结静态 AGENTS.md	正确的初始项目说明，之后不更新
dsh-memory	真实插件抽取、存储、更新后的注入段

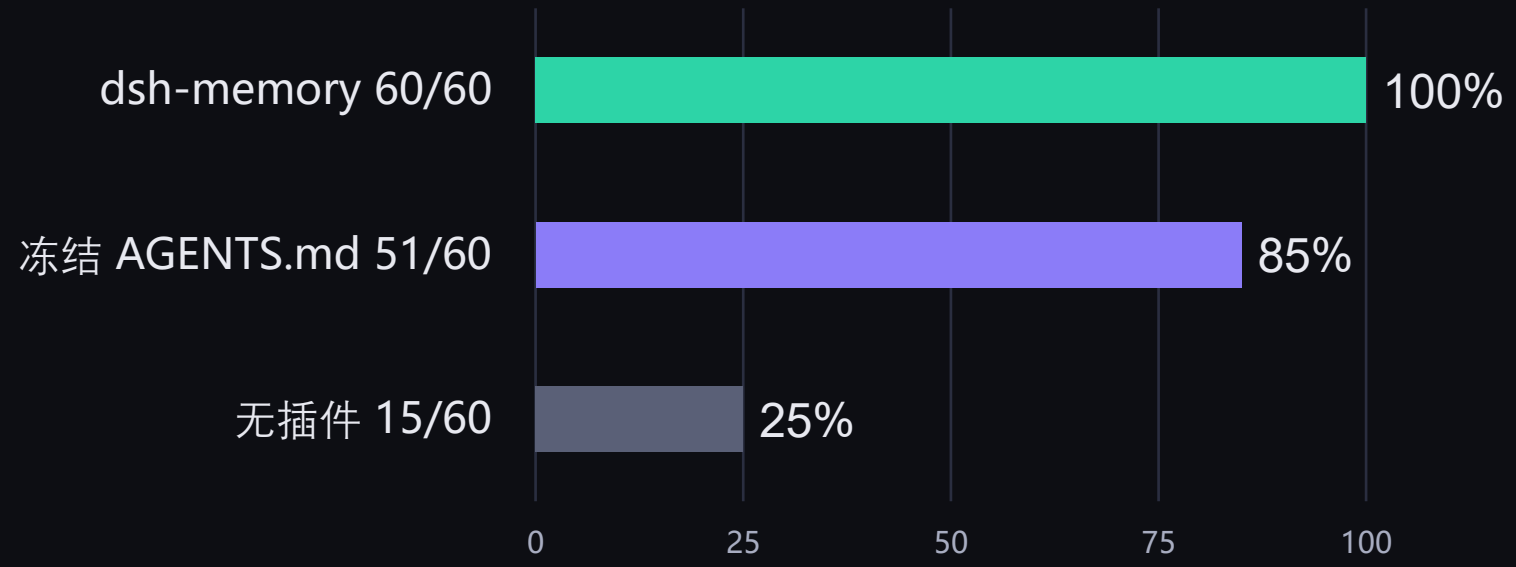
三组使用同一模型、问题与评分；学习后创建新上下文，避免旧聊天泄漏。

192 单元、306 次请求全部完成。答案表不进入模型输入。

deepseek-flash / temperature=0 / 固定 seed；真实插件后端，非完整编程任务评测。

08 结果：自动更新减少旧约定残留

仅限这 20 个固定合成案例；60 次作答不是 60 个独立案例



+15

个百分点

相对冻结静态文件
多答对 9 次

三个重复轮次一致：5/20、17/20、20/20。

差异全部来自：npm 改 pnpm、Node 18 改 22、周三发布改周五。

恢复旧发布规则时，静态组也答对；无插件的 45 次 UNKNOWN 属于缺信息拒答。

09 成本：历史抽取需要额外调用

token 数量来自实际 provider usage；治理诊断单独计费量，不混入主对照



主对照组	调用数	合计 tokens
无插件	60	9,237
冻结静态 AGENTS.md	60	10,455
dsh-memory	159（含 99 抽取）	59,949

插件主实验 token 用量约为静态组的 5.73 倍。

全部 306 次请求共 89,057 tokens，含独立治理诊断的 9,416 tokens。

未折算货币费用或人工维护时间；实验频次不能代替默认配置下的长期成本。

10 结果适用范围

报告同时保留收益、成本与尚未验证的边界



没有证明胜过人工维护

静态基线在学习后冻结；持续维护的 AGENTS.md 也可能答对，维护时间与综合收益还没有测量。

没有证明普遍 100%

只有单模型、20 个合成短案例；没有独立留出集，不能直接推断真实开发效率或统计显著性。

新版功能与旧评测分开

正式评测来自 v0.4.1；v0.5.1 增加时间、存储和网页操作验证，不能用旧分数证明持久记忆的长期效果。

AI 逐条审阅 96 条新增候选，未发现无历史支持事实；未做独立人工标注。

未证明长期准确率或时间参数最优；真实代码任务与复杂攻击仍待验证。

11 单人组与 AI 协同开发

课程工程采用 AI 开发；作者负责范围决策与提交答辩



作者

选题与范围决策、配置凭证、运行真实评测；负责最终提交与答辩。

AI

源码研究与实现、回归与安装验证、评测脚本、结果复核、材料生成。

配套工程

dsh-commit 读取暂存区生成中文提交信息；工具本身不执行提交。

交付：插件源码、展示网站、PPT/PDF、评测报告与介绍视频。

项目网站提供公开源码、真实操作视频与评测证据，便于复现和核验。