

roboRSI：任务图辅助的执行反馈与双臂策略框架改进

roboRSI: TaskGraph-assisted Execution Feedback for Bimanual Harness Improvement

技术报告 v0.4 · 2026 年 10 月 20 日

作者与机构待定。主要实证包括已归档的五任务架构研究、历史重放恢复筛选、一次新的 TaskGraph 接入诊断与模型复盘，以及一个 RealBot 真机插笔回合。

摘要

双臂操作中的错误可能发生在视觉判断、技能调用和实际物理状态之间。为了让这些执行经验进入后续策略框架改进，我们在线分层执行与回合后轨迹分析连接起来：在线阶段由 Astra 通过代码技能或 Simate-beta 操作机器人，原生环境独立判定结果；回合结束后，TaskGraph 使用本地 SDK 把轨迹整理为阶段、进度、故障环和恢复候选，再向复盘与候选生成入口提供有来源的摘要。

该反馈路径已经接通。我们对既有 75 份动作重放侧录完成本地筛选，得到十八个恢复候选、六个尝试未恢复、四十七个未观察到故障和四个待复核回合。十八个恢复候选中，十个回合最终失败，一个没有图环，显示局部恢复需要与整合成功和图环分别记录。另一个真实笔筒诊断回合执行 940 次动作，以 Score 25/100、任务失败结束；一次真实 Astra 复盘消费 TaskGraph 摘要及十八帧三视角采样，输出四项通用改进提案。它们目前是文本假设，尚无可执行子代的配对选择与继承结果。

真机方面，我们在 RealBot 上为 Astra 增加了一层任务工具：胸部 RGB-D 物体检测，以及取笔、侧夹笔筒、递交、插入和放回等双臂技能。技能在开始时按实测关节与当前检测重新规划，展开为普通动作，仍逐步经过安全层。10 月 20 日的一次运行历时约 41 分钟、47 次模型调用，左手把五支笔中的四支插入右手持握的笔筒，第五支落在筒外后由操作员停止。运行无遥操作、动作全部自动批准，但操作员在中途向模型发送了三条文字提示。这是单回合演示，不构成真机成功率。

此前完成的五任务、三架构研究共 75 个原生回合，Astra+H₀、混合组和独立 Simate 的成功率分别为 40%、40% 和 20%，平均 Score 为 47.47、56.13 和 21.60。该研究提供执行与工具分工的基础证据，不用于估计新 TaskGraph 接入的收益。本文报告两类历史数据和新诊断各自的范围，并据此提出下一阶段的分类核验、候选比较和独立测试协议。

目录

1. 研究问题	2. 系统定义与证据范围
3. 分层执行方法	4. TaskGraph 回合后反馈机制
5. 恢复筛选与真实复盘结果	6. 策略框架改进的实现与协议
7. 五任务架构实验设置	8. 实验结果与解释
9. 执行失败揭示的接口问题	10. 跨机型与真机工作
11. 局限与下一阶段研究	附录 A. 补充研究快照
附录 B. 复核方式与数据口径	参考与证据索引

1. 研究问题

放笔入筒要求一只手稳定容器，另一只手依次抓取、搬运和插入物体，再完成放筒与回位。任务中的错误经常具有过程性：夹爪请求已经闭合，物体却未可靠附着；一次撤回遇到停滞，后续清障运动可能恢复了手臂运动，却没有恢复笔的放置。最终成功标记只说明终局，难以指出哪些局部修正值得保留。

roboRSI 研究两种相连的反馈。在线反馈让模型根据新的观测调整当前行动；回合后反馈组织完整执行记录，为下一版框架提供改动假设。TaskGraph 在后者中承担结构化分析：把动作与测量记录转换为阶段和进度，在实际时间顺序中定位故障、修正与恢复，并将来源绑定到模型复盘。

本报告围绕三个已经具备材料的问题展开：分层执行如何连接高层模型与控制器；任务图怎样筛出可复核的恢复候选；一次真实执行的日志、任务图和图像如何形成改进提案。候选在独立评测中是否有效，是后续需要回答的问题。

1.1 方法来源与研究范围

PhysicalRSI 提出可编辑 harness 与具身反馈驱动的候选选择，并区分回合内状态与跨版本技能。我们沿用这一组织方式，将本地双臂执行记录和 TaskGraph 分析接到复盘与候选入口。[R1, R2]

参考项目的研究材料还区分了候选被验证、搜索后策略提高，以及反馈搜索优于随机搜索。这些结论依赖不同的对照。本文因此将执行能力、接入完成、候选提案与最终收益分别报告。[R3, R4]

roboRSI 指本地 RoboDojo、TaskGraph 和 RealBot 研究。历史代码保留 `physical_rsi` 名称。外部 PhysicalRSI 榜单、其他同名项目及其他训练路线的结果不进入本报告的实验分母。

2. 系统定义与证据范围

2.1 在线执行与跨回合更新

设任务目标为 g ，公开观测为 o_t ，回合内记忆为 m_t ，基础模型为 F ，策略框架为 H_k ，技能和工具集合为 S 、 T 。在线执行为：

```
(a_t, m_(t+1)) = H_k(o_t, g, m_t; F, S, T)
τ = {(o_t, a_t, measured_feedback_t)}_(t=0..T)
y = V(environment, τ)
```

这里的 H_k 包含调用模型的时机、上下文组织、技能程序和执行路由。a_t 最终需要转换为机器人接口接受的动作。V 是原生环境的任务判定，与模型文本中的完成声明分开。一次回合结束时，研究系统同时保存成功标记、过程分和结束原因，以区分任务失败、预算耗尽及基础设施异常。

跨回合更新的目标形式为：

```
C_k = Propose(H_k, Analyze(D_dev))
H_(k+1) = Select({H_k} ∪ C_k, D_sel)
Report(H_frozen, D_test)
```

D_dev、D_sel 和 D_test 分别表示开发、选择与最终测试数据。Analyze 现在包含日志整理、TaskGraph 摘要与视觉复盘。入口已经接通，是否形成可继承子代仍取决于完整开发与选择过程。父代保留在比较池中；没有满足条件的候选时，下一轮继续使用父代。

这两个时间尺度对应不同的变化。回合内，模型根据新观测调整计划和动作，但正在执行的框架版本保持固定。回合之间，研究流程才允许修改并冻结新版本。历史 75 回合中的四个源码版本属于有人工参与的工程迭代，不能用轮次编号直接代替自主继承代数。

2.2 需要分别记录的状态

对象	记录内容	生命周期与用途
回合状态 m_t	当前阶段、关联手臂、模型记忆与待执行动作	随回合 reset 清空，支持当前任务执行
交互记忆	观测、模型决策、工具回执与图像锚点	当前回合内检索，用于重新判断进度
框架 H_k	技能、路由、上下文配置和控制依赖	冻结到版本，修改后需重新评测
实验证据	布局、源码、动作、录像、原生结果和费用	跨回合保存，支持结果核对与后续开发
TaskGraph 派生记录	阶段、进度、故障环与恢复候选	离线分析后进入复盘和候选入口，保留数据用途约束

“记住返回到接近位姿的规则”与“保存本回合测得的位姿数值”属于不同对象。前者可以成为技能程序的一部分，后者必须随回合重建。同样，一段成功录像可以作为后续研究依据，但其中的场景坐标不能自动变成下一布局的控制目标。

2.3 四种产物不能混算

产物	数量与来源	当前可支持的结论
历史架构研究	75 个原生回合，25 组同布局对照	多版本架构表现与执行机制分析
恢复筛选	上述历史经验的 75 份重放侧录	本地筛选运行与候选分布；没有新增 75 次机器人实验
接入诊断	1 个新的原生笔筒 smoke，940 次动作	原生采集、SDK 处理与真实复盘消费
候选建议	复盘中的 4 项通用文本提案	产生了改动假设，尚无可执行子代、选择或继承收益

历史筛选没有独立人工标签，三项分类指标保持 null。新的诊断回合标记为 smoke，`eligible_for_improvement=false`；它可以验证复盘输入，但不作为研究继承依据。后续停止或未启动的回合不加入完成分母。

3. 分层执行方法

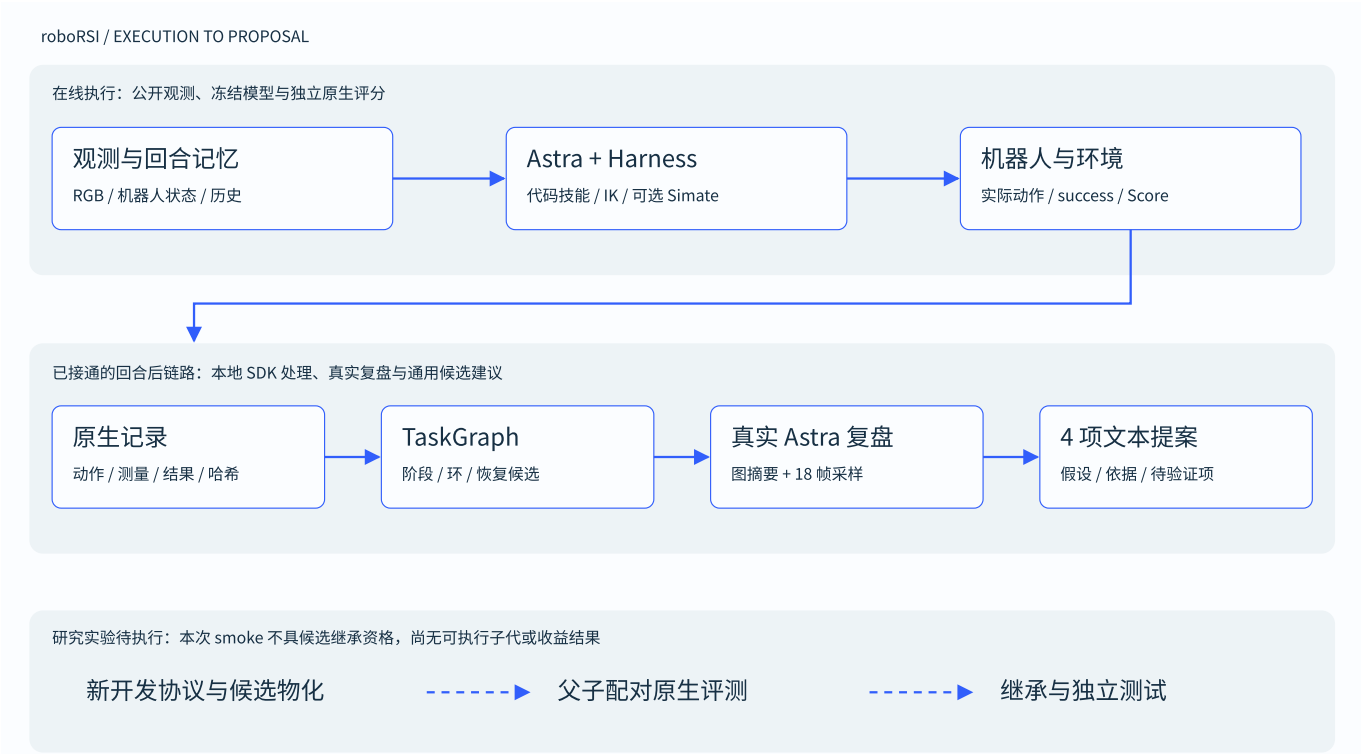


图 1：在线执行结束后，TaskGraph 在 CPU 上组织轨迹并向复盘和候选入口提供摘要。实线部分已有实现与消费记录；配对选择、继承和独立测试仍需新的研究实验。

3.1 公开观测与模型上下文

75 回合研究使用 RoboDojo 中的双 ARX X5，每臂六个关节和一个夹爪命令，共 14 维动作。头部和左右腕部提供三路 RGB。Astra 的常规视觉输入包含当前完整图像、腕部裁剪及有限历史图像；文本部分包含任务指令、机器人关节、LINK6 位姿、夹爪命令与实测手指状态、公开相机标定和近期执行反馈。[L1]

历史图像用于恢复执行上下文。例如，当前画面里没有笔时，模型可以查询较早时刻的观测，检查它此前被移动到了哪里。历史采样与模型决策频率分别管理：该研究的 Astra 图像历史约每八次原生动作采样，缓存最近十二个条目，再按上下文预算选择输入帧；inspect_history 允许补取指定时刻的图像。Simate 的历史输入是另一套序列，不能与 Astra 的上下文窗口混算。

物体真值、私有布局文件、深度和奖励内部变量不进入该研究的在线策略。几何工具只根据公开标定、模型指定的像素对应或假设平面计算射线、交点和投影。因此，一次几何查询是否有用，既取决于标定是否正确，也取决于模型选择的像素是否对应目标。数值计算正确不能消除识别错误。

3.2 结构化决策与固定技能 H_0

Astra 输出结构化工具决策。查询类工具提供历史或几何信息；运动类请求进入技能与控制层。 H_0 将抓取等请求展开为多个阶段，使接近后的新观测有机会影响接触与闭爪。这样，模型不必在一次请求中决定整个任务的所有关节动作。

技能状态与物理动作的提交需要保持一致。若动作因参数、姿态或可达性检查被拒绝，状态不应先推进到“已经执行”的下一阶段。否则模型下一次读到的阶段标签会与实际机器人位置不符，错误可能沿着后续技能继续传播。

H_0 中的技能是程序化的执行结构。lift、push 等入口通过末端运动原语实现，没有各自训练独立神经网络。受限 AST 解释器约束程序操作、循环和状态大小，并限制导入、文件与网络访问。它约束技能代码能够做什么；物理执行是否达到任务要求仍由环境判断。[L1]

3.3 VLA 委托与返回高层

混合组在 H_0 工具集之外增加 Simate-beta。Astra 可以给出自然语言子目标，由 Simate 根据三路当前 RGB、机器人状态和头部历史生成动作。该研究使用冻结的 FP32 权重 Simate-beta-RoboDojo-11424ad0-fp32，网络预测五十步动作，适配器执行前十步后重新观察。头部历史保留二十五帧，采样间隔为二十次动作；单次高层委托最多执行一百次原生动作。[L1, L9]

动作块执行期间仍检查反馈和剩余预算。委托结束或发生中断后，高层读取新的观测与执行摘要，再决定继续委托、修正子目标或调用代码技能。这里的层级关系依赖实际路由：混合组具备 VLA 工具，并不要求每个回合都使用它。

独立 Simate 组使用相同的权重和动作适配，输入原始任务指令，不接受 Astra 在线子目标、反思或终止判断。它也没有在线 Astra 请求。混合组相对独立 VLA 的变化因此同时包含高层规划、额外工具与执行路由，二者的差异不能被拆解为单一组件的收益。

3.4 动作坐标与执行反馈

历史研究把运动位置解释为 LINK6 手腕原点，位置单位为米，四元数采用 wxyz。TCP 偏移在明确的坐标转换处处理。控制器把目标位姿转换为关节目标，并限制单步关节变化；该研究使用的上限为 0.08 rad，VLA 输出也经过共同动作约束。

记录中区分三类量：模型或技能请求的动作、经过转换与限制后实际派发的命令，以及执行后的实测状态。请求被接受、关节到位、夹爪稳定、物体被可靠持有，需要不同的观测支持。尤其在接触遮挡下，末端到位和夹爪停止运动并不足以判断笔是否抓住或插入。

双臂任务还需要保持另一只手的已有状态。后期版本加入可选的手臂保持约束：以委托开始时的观测为参考，检查指定手臂的位移、旋转和夹爪请求变化，超出条件时取消后续委托并交回高层。它可以限制一类破坏支撑关系的动作，但其是否减少失败、是否增加误中断，需要配对实验测量。

3.5 时间、调用额度与回合结束

仿真原生观测按 25 Hz 更新，Astra 推理期间仿真不自动推进。API 等待三十秒不等于机器人经历了三十秒物理运动。控制稳定性使用仿真时间，系统成本使用墙钟；视频播放时间又需要由帧号和帧率核对。

Astra 使用 `reasoning.effort=high`，单次输出上限为 6144 tokens。历史研究的回合调用上限为 `min(60, ceil(native_action_limit / 10))`，查询与修复请求同样占额度。上限约束可以防止无限查询，也会使重复观察与动作执行争夺预算，因此需要保留每次请求的用途，而不能仅看总调用数。

任务成功、动作预算耗尽、模型请求耗尽和服务异常分别记为结束原因。模型输出 `done` 是一项决策，其含义需要结合原生状态解释。后文的案例表明，过早接受或机械截断这一声明，都可能改变任务结果。

4. TaskGraph 回合后反馈机制

4.1 接入位置与数据流

TaskGraph 已接入回合结束后的处理、复盘和候选生成入口。在线策略仍使用公开观测、回合记忆与执行反馈；TaskGraph 在回合完成后读取已有轨迹，通过本地 SDK 生成阶段图、环特征与恢复候选。模型权重保持冻结，任务图不进入在线动作选择，也不作为动作奖励。[L11]

数据流从原生结果开始。模拟器保存终局和采集记录，回合后处理构造带来源哈希的请求，本地 SDK 产生节点与边，随后生成恢复记录和有界摘要。`visual_review()`、代码候选生成器与 Genome 候选生成器消费通过校验的摘要。候选是否继承仍由原有父子配对原生评测决定。

实现中有两个后处理入口。`sim_worker` 在结束策略与环境、但尚未关闭 Isaac 应用时，启动 CPU 子进程运行 `watchdog`；父进程 `after_episode()` 再核对并补做反馈。实跑发现 `app.close()` 会直接终止 Python，放在它后面的处理不会执行。这一问题已修正，最终 CPU 路径在保留的真实轨迹上重新处理并通过核对，应用退出时序另有行为测试。修正后没有为此重新生成一个 GPU 回合，因此应区分真实轨迹处理验证与新版本完整重跑。[L11, L12]

4.2 输入绑定与语义状态

请求记录回合与来源 ID、源码身份、实际可用的指令、动作、状态、回执、原生结果及媒体引用。缺失的历史字段保持空值。消费前检查采集完整性、动作编号、样本索引、实测时间、压缩数据及原生结果的绑定，避免将旧图或其他回合的摘要接到当前提案上。

任务插件将测量点映射为 `State(phase, progress, adverse, terminal)`，再把相同的连续状态压缩为片段。节点身份包含任务、插件、配置与采集约定。同一回合重复访问某节点，会增加访问次数，但不会增加独立回合支持数。

阶段是解释执行记录的一种抽象。笔筒插件根据物体与末端几何推断抓取、搬运和插入等阶段，含经验阈值和离线平滑；这些标签不等同于接触、物体附着或原生成功。当前新 campaign 的原生采集先覆盖笔筒，历史重放可以复用已有五任务适配器，二者的支持范围不同。[L3, L4]

4.3 恢复候选的判定

新版恢复筛选把故障、实际修正与后续进展分别定位。它首先查找适配器标记的异常或已知进度回退，然后要求后续出现规定的修正阶段，并能找到相对故障时发生变化的实际执行命令。最后检查修正之后的进展是否超过修正时的水平；普通异常要求超过故障前进度，进度丢失则允许恢复到此前水平。

这个顺序用于避免把修正之前已发生的进展，归给后来的动作。新故障也不能沿用前一次故障的修正记录。没有明确修正、进度缺测、适配器未知或数据损坏时，系统保留待复核状态。

输出类别为 recovered、attempted_not_recovered、no_failure_observed 和 needs_review。图环是独立特征，恢复不要求存在环；局部恢复也不要求整回合最终成功。旧 GRIT 原型中的 failure_unhandled 不再作为本次四分分类输出，不能把两个版本的标签混用。[L11]

这些规则形成的是几何启发式候选。动作变化和阶段顺序可以定位需要检查的片段，却不能证明行为意图或因果关系。正常重新摆放也可能产生进度回退，需要通过录像和独立标签识别误报。

4.4 送给模型的反馈

完整图和轨迹保留在证据目录；模型接收有界摘要，包括阶段、动作范围、实测时间、重复次数、进度变化及来源引用。当前环摘要最多保留八个候选，每个最多四次事件、每条路径最多八个阶段，并附最近十二个阶段。截断情况单独记录，不能把摘要中未出现的事件解释为不存在。[L4]

复盘优先按恢复事件的动作号定位三路录像；没有可靠映射时回退到普通抽帧，并记录原因。这次真实复盘使用六个时刻、三个视角，共十八帧。它能够帮助模型检查候选解释，但不提供连续接触过程的逐帧证明。

摘要不直接传递原始物体坐标。由私有分析状态推断的标签仍可能包含额外信息，因此数据使用约束不能只依赖“删掉坐标”。实现只允许 development/train 证据进入研究候选改进，拒绝 selection/final_test；旧 test 采集不自动改为训练数据。诊断 smoke 可以被复盘，但其 eligible_for_improvement=false 必须沿下游保留。

5. 恢复筛选与真实复盘结果

5.1 历史侧录的批量筛选

固定输入集合来自既有 75 个回合的动作重放侧录。处理前冻结 ID 集合，再核对原始动作、重放状态、采集文件和重放自身的原生结果。75 份侧录均完成本地 SDK 处理，未为筛选新增 GPU 回合，也没有调用模型 API。[L11, L13]

自动分类	回合数	解释
恢复候选	18	发现符合当前规则的故障、修正与后续进展
尝试恢复但未恢复	6	有修正尝试，未满足恢复进展条件
未观察到故障	47	当前记录与规则未标出故障，不等同于任务成功
待复核	4	保留缺证、歧义或规则不能确定的情况

十八个恢复候选中，十个重放回合最终失败，一个没有图环。这个分布说明局部恢复、图环与最终成功在数据上并不等价，但并未证明十八个候选全部正确。

笔筒子集共十五份侧录，其中恢复候选四个、尝试未恢复六个、未观察故障四个、待复核一个。其余任务的输出用于探索与规则调试，尚不是跨任务分类性能验收。

当前没有独立人工标签，accuracy、precision 和 recall 均为 null。正式的一百回合分类评价尚未采样或运行，应与另一项“二十五个开发回合加七十五个父子选择回合”的配置搜索分开。历史侧录已经用于规则调试，不再作为该分类评价的未曝光测试数据。



图 2：四类自动标签的数量。独立人工标签尚未取得，分布不代表分类准确率。

5.2 一次原生接入诊断

新的 taskgraph-rsi-20261019 实验完成一个笔筒诊断回合，使用 layout 5、Direct 路径。回合执行 940 次原生动作，在线 Astra 调用六十次，因调用上限结束；原生 Score 为 25/100，任务未成功。三路录像完整，原始配置、动作、结果和采集文件的哈希在最终处理前后保持一致。[L12]

记录	本次结果
完成原生回合	1
原生动作	940
在线 Astra / 复盘调用	60 / 1
原生 Score / success	25/100 / false
TaskGraph 处理	完成，本地 SDK
恢复分类	needs_review，恢复事件列表为空
环分析摘要	3 个图环，2 个故障环候选
复盘图像	6 个时刻 × 3 路相机，共 18 帧
已完成模型调用费用	\$5.415015，共 61 次调用
研究改进资格	false，诊断 smoke 不进入候选继承

后续 smoke 队列停止，第二个回合在首个动作前取消，第三个未启动。它们不计为完成实验，也不补零加入成功率。这一个诊断回合只验证接入与消费路径，不并入历史 75 回合的架构比较。

5.3 任务图、录像与复盘之间的差异

故障环摘要标出动作 614 至 789，以及 789 至 848 两段无新增进展的路径。第二段被阶段规则描述为抓取准备相关转换，但执行记录包含撤回、历史查询、放下和释放。复盘因此要求交叉检查标签与实际动作，没有把图上的环直接判作已经确认的失败或恢复。

复盘引用的执行反馈还显示，动作 776 附近的撤回出现停滞，残余位置误差约 13.4 mm、近期进展约 0.205 mm；后续另一条侧向与旋转清障运动完成，位置误差约 0.048 mm。这支持局部机器人运动恢复的判断，仍不足以证明物体放置已经恢复。接触未验证也不等于没有接触。

在末帧中，腕部图像出现超出笔筒轮廓的青色物体部分，但稀疏采样不足以确定准确插入数量、支撑几何或状态变化的原因。原生 Score 0.25 同样不能直接换算为“正确插入了一支笔”。

录像时间与仿真时间需要单独核对。复盘看到末帧为 37.600 秒，末次观测为 40.968 秒，因此提出检查 3.368 秒偏移。本报告对三路各 941 条 timeline 记录与视频实际帧率进行重算，记录中的仿真时间减视频时间始终为 3.368 秒（浮点容差内）。这解释了本条录像的时间原点差，不能自动证明所有观测、历史检索与动作端点都已同步。[L14]

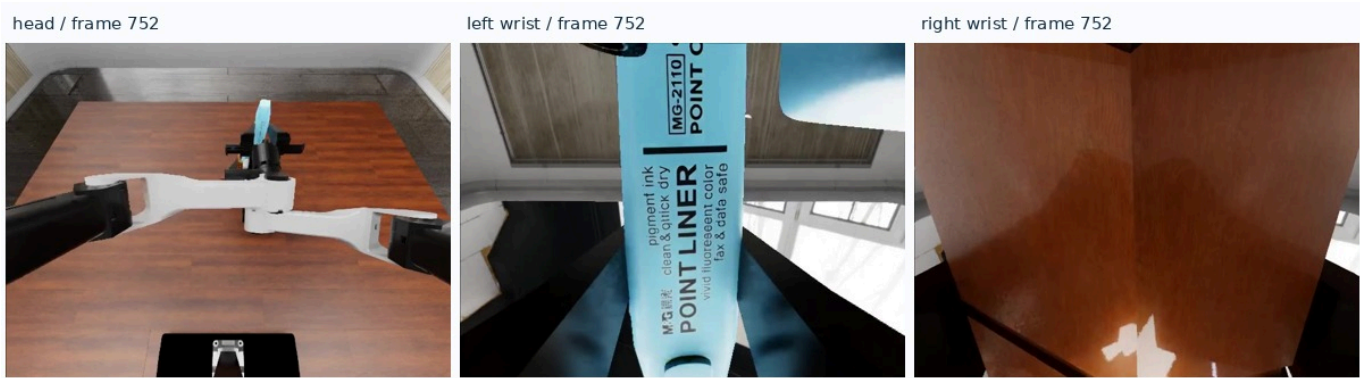


图 3：同一采样帧的头部和双腕录像画面。模型复盘用这些视角检查遮挡与支撑；单帧不能确认连续接触或最终容纳状态。

5.4 四项候选改进提案

一次真实 Astra 复盘产出了四项通用建议。它们已经保存为文本提案，没有在这次记录中物化为新的技能代码或配置，也未被执行或继承。其来源是上述不具研究改进资格的 smoke，后续若继续，需要在新开发协议中重新建立证据与候选身份。[L15]

候选	提出的改动	本次证据	下一步验证
C1：反馈来源	分别记录请求夹爪值、保持命令、回显、实测开度、附着和支撑证据	命令、回显与物体判断具有不同语义，支撑验证仍为空	命令保持与来源检查，独立验证附着和支撑
C2：阶段间核验	在搬运、释放、撤回与完成之间保留不确定性，必要时补充视角	752 与 940 采样帧中的筒口和笔端遮挡	误完成、额外观察与物体扰动的配对比较
C3：统一时间	对观测、动作端点、标定与检索历史保留时间和有效性	录像与末次仿真观测存在时间原点差，历史证据被再次使用	核对时钟与端点，测试历史在物体关系变化后的失效
C4：观察预算	根据对应点质量、射线几何、可见性和剩余调用量决定是否继续观察	末段定位被拒绝，回合到达调用上限	定位拒绝、信息增益、调用上限频率与原生成功

复盘还要求查看若干连续视频区间，并用独立标签核验物体状态。模型自己提出的解释与建议不等于人工分类标签，四项提案也不是四个已通过验证的子代。

5.5 接入完成与研究收益

当前已经完成的是原生记录到本地 TaskGraph 的处理，以及真实复盘对图摘要和图像的消费；恢复候选筛选与文本改进提案都有可核对产物。实现还把受限摘要接入代码与 Genome 候选入口，并保留数据用途约束。

尚待验证的是恢复分类是否准确、摘要是否改善可执行候选、候选是否优于父代，以及多轮继承是否提高独立测试表现。本次失败诊断回合可以证明链路能够消费失败证据，不能证明失败已经被候选修复。

6. 策略框架改进的实现与协议

6.1 可修改的对象

项目已有两类候选入口。受限技能代码搜索修改技能程序；共享配置搜索调整查询次数和历史图像分配。二者的搜索空间不同，需要分别冻结父代与允许变化的字段，不能把一次控制器修复同时解释成两类搜索的收益。

共享配置实现目前包含三项任务无关参数：

参数	父代值	搜索范围	影响的执行行为
read_only_limit	4	0 至 4	一次规划中允许的历史与几何查询数
recent_observations	3	1 至 3	自动上下文中的近期观测数量
wrist_event_views	2	0 至 2	历史腕部事件视角的分配

每个共享配置候选恰好改变一项，两个候选不得重复。当前三路图像仍由共同接口提供；减少某种历史视角可能为其他锚点腾出空间，不意味着总图像量必然同比减少。候选也不能修改模型权重、任务评分、动作上限或增加任务名称对应的固定坐标。[L2]

6.2 开发、选择与最终测试

已准备的共享配置实验对五个任务分别设置五个父代开发布局 and 五个选择布局。父代在开发阶段运行 25 回合，随后生成两个候选；父代和两个候选在选择阶段共运行 75 回合，一轮合计 100 个研究回合，工程 smoke 另计。

条件核对沿用 H/M/E/R/D：H 为框架、提示与记忆，M 为模型与输入语义，E 为环境、机器人、资产和评分，R 为资源与执行预算，D 为布局及数据用途。父子比较应明确哪些字段允许改变，哪些必须保持一致。

候选提案只能读取开发数据。父代与候选在选择阶段使用相同布局、模型、技能、控制边界及预算。选择要求每个任务的成功数和 Score 均不下降，汇总至少一项严格提高。合格候选按成功数、Score 和费用进一步排序；完全持平保留父代。下一轮引用必须能回到选中的配置和父子证据。

最终测试发生在候选选择完成、版本冻结之后。开发、选择、测试应按实际布局内容检查重复，仅更换随机种子不足以保证新场景。早期 cover_blocks 的重复布局问题表明，数据身份需要同时检查文件内容和几何近重复。[L1, L2]

6.3 已完成到哪一步

早期代码搜索留下了开发和部分选择尝试；v10 产生两个实际代码候选，并在 534 个离线决策边界上检查了执行契约，但没有完成对应的原生选择集。共享配置路径完成了准备工作和五任务工程 smoke，五个回合均成功。完整 100 回合候选研究与独立最终测试仍未完成。[L1, L2]

75 回合架构研究采用另一种推进方式：固定一轮源码和十五个研究槽位，每五个完成回合进行日志与视频复盘，整轮结束后再考虑工程修正。新版本先经过工程检查，再进入下一轮。这个过程产生了可追溯的软件演进，但没有把所有修正交给自主候选选择器决定。

因此，本文把实现流程、工程运行和研究收益分别记录。TaskGraph 反馈已进入复盘和候选入口，但本次 smoke 的四项文本建议没有改写上述历史选择结果。候选物化、父子评测与继承仍需要新的研究记录。

6.4 经验筛选与模型训练的边界

TaskGraph 的恢复筛选服务于诊断与研究提案。现有 SDK 还保留另一种训练候选选择方式，按结构覆盖、恢复和多样性计算增益：

$$G(i \mid S) = \alpha \cdot \text{coverage}(i \mid S) + \beta \cdot \text{recovery}(i) + \gamma \cdot \text{diversity}(i \mid S)$$

这里 i 是候选轨迹， S 是已选集合。该选择器只在合格训练数据中拟合结构频次，保留来源家族与划分约束。正向示范导出还要求成功终局及动作前公开观测对齐，与允许最终失败的局部恢复筛选不是同一种资格。

本次接入没有进行模型训练，也没有测量该选择器的训练收益。最终成功的整集仍可能含错误动作前缀，训练时需要另行决定分段监督方式。[L3]

7. 五任务架构实验设置

7.1 任务、布局与预算

每轮对五个任务各选择一个布局，三组在同一布局执行一次；五轮共形成 25 个任务与布局单元、75 个回合。三物体中央抓放是本项目扩展任务，其余任务沿用对应 RoboDojo 定义。任务设置覆盖指令选择、工具操作、支撑关系和多物体协作。[L9]

任务	主要要求	原生动作调用上限	Astra 调用上限
fill_pen_holder	扶筒、装入四支笔、放筒与回位	1100	60
general_pick_place_repeat	三个物体依次抬升并放入中央区域	600	60
align_blocks	使用三角板推动积木对齐	200	20
general_pickup	选择指令目标并抓取	200	20
stack_blocks	抓放、支撑与堆叠稳定	550	55

五轮使用开发布局 0 至 4，部分与更早研究存在重复使用。这组实验用于探索架构差异，没有宣称它们是未曝光的最终测试集。任务之间动作预算不同，因此不能把跨任务动作数直接当作相同难度下的效率比较。

7.2 对照条件

条件	高层决策	可用执行路径	在线 Astra
Astra+H ₀	Astra+回合记忆	固定 H ₀ 、几何工具与 IK	有
混合组	Astra+回合记忆	与 H ₀ 组相同工具，并可委托 Simate	有
独立 Simate	原始任务指令驱动的动作策略	Simate 连续动作块	无

astra_only 表示 Astra+固定 H₀，没有神经 VLA；它包含技能程序和控制底座。后续跨机型矩阵中的 Astra_Direct 使用相对腕部动作接口，属于不同条件。本文保持两者名称与分母分开。

两条 Simate 路径使用相同权重、归一化、动作前缀、历史设置和共同动作边界。混合组的子目标来自 Astra，独立组只编码原始任务指令。后端异常按异常记录，不能静默换成另一条策略路径后继续计为原条件的正常结果。

7.3 五轮与四个源码版本

轮次	布局	源码内容哈希前缀	该轮相关变化	H ₀ / 混合 / VLA 成功数
1	0	448a1e39	独立 VLA 条件、可选路由、整段执行反馈	2/5 · 2/5 · 2/5
2	1	d34e51e3	公开相机内参与实际 RGB 渲染对应修正	2/5 · 3/5 · 2/5
3	2	d8d7441d	姿态声明一致性、耗时与保持目标残差反馈	4/5 · 2/5 · 0/5
4	3	6bf94baf	保持约束与后续费用预留修复进入冻结版本	1/5 · 2/5 · 1/5
5	4	6bf94baf	沿用第四轮源码，布局改变	1/5 · 1/5 · 0/5

上述前缀是 manifest 的源码内容哈希，不是 Git commit 缩写。每个任务与布局单元中的三组使用对应轮次的冻结版本。前三个版本每任务只有一个布局，最后一个版本每任务两个布局，样本不足以分离源码修正、布局难度和执行波动的影响。[L1]

7.4 指标与统计单位

success 使用原生终局布尔值；Score 保留原生过程评价，从 0 至 1 转为百分制一次。每任务每组恰好五个回合，因此本研究按全部回合计算的均值，等于五个任务均值的等权平均。

配对比较以“轮次、任务、布局”为单位，将同一单元中的两种架构直接比较。成功率差描述完整任务结果，Score 差描述过程评价变化。我们同时报告双方成功、单方成功和双方失败的数量，以呈现均值掩盖的样本差异。

这 25 个单元跨任务、源码和已用于开发的布局，不宜当作一个冻结策略下独立同分布的 25 次试验。报告保留描述性配对计数，不把小样本区间或单次差值解释为普遍优势。逐回合数据与本次重算结果分别见 [CSV](#)、[分析记录](#)。

7.5 完整性与重试

统计从冻结的系列清单及恢复映射选取每个完成槽位，避免原目录、录像目录和归档别名被重复计算。原生结果、配置、源码和录像绑定到各自回合。原始中断尝试仍保留，其费用进入全部尝试的账目；它们不作为额外的完整策略样本。

任务失败按原结果保存，不因得分较低而选择最好的一次。没有终局的尝试保持缺失或异常状态。数据提取和本次配对重算都不触发模型或模拟器。[L1, L10]

8. 实验结果与解释

8.1 总体结果

架构	成功 / 完成回合	SR	平均 Score / 100
Astra+固定 H ₀	10/25	40%	47.47
Astra+H ₀ +可选 Simate	10/25	40%	56.13
独立 Simate	5/25	20%	21.60

混合组相对 H₀ 的平均 Score 高 8.67 点，完整成功数相同。这个结果意味着过程评价与终局完成发生了不同变化，不能只凭均分断言可靠性提高。相对独立 Simate，混合组多成功五个回合，但这一比较包含高层决策与工具集合的整体差异。

8.2 任务差异

任务	Astra+H ₀ : 成功数; 均分	混合: 成功数; 均分	独立 VLA: 成功数; 均分
放笔入筒	1/5; 32.00	2/5; 68.00	2/5; 42.00
三物体中央抓放	2/5; 53.33	0/5; 46.67	0/5; 0.00
指令抓取	3/5; 60.00	4/5; 80.00	0/5; 0.00
工具推积木对齐	3/5; 60.00	1/5; 20.00	0/5; 0.00
堆叠积木	1/5; 32.00	3/5; 66.00	3/5; 66.00

笔筒任务中，混合组和独立 Simate 都成功两次，混合组均分更高。现有样本说明两组在未完整成功的回合中取得了不同的过程评价，尚不能确定差异来自哪一个规划或控制机制。

中央抓放更清楚地暴露了过程分与完整性的差别。混合组均分为 46.67，却没有完成全部三个物体要求的回合；H₀ 完成了两次。仅展示某一次抓取或搬运成功，会遗漏剩余物体与终局条件。

对齐与堆叠呈现相反方向的组间差异：H₀ 在对齐上成功三次，混合组一次；堆叠则是 H₀ 一次、混合组三次。这支持继续研究按任务状态选择工具的条件，但每个任务只有五个布局，尚不足以拟合或验证稳定的路由规则。

8.3 配对结果



图 4：每格是已完成的配对单元数。两组成功率均为 40%，但只有四个单元双方都成功；相同总数由不同的成功集合组成。

左组相对右组	双方成功	仅左组成功	仅右组成功	双方失败
混合 / H ₀	4	6	6	9
混合 / 独立 Simate	4	6	1	14
H ₀ / 独立 Simate	0	10	5	10

混合组与 H₀ 在十二个单元上结果不同，双方各占六个。这说明“加入可选 VLA”改变了可完成的样本集合，却没有在这批数据上净增加成功数。H₀ 与独立 Simate 的成功集合没有重合，表明两条执行路径在当前样本中具有不同能力分布；实际混合策略仍需要决定何时调用哪条路径，并处理路由后的反馈。

Score 的配对结果也不均匀：混合组对 H₀ 为十胜、九平、六负，均值差为 8.67 点；对独立 Simate 为十四胜、十平、一负，均值差为 34.53 点。这里的胜负只比较同单元的原生分数，不意味着各次差值由同一种原因产生。

8.4 实际 VLA 使用与结果归因

混合组 25 个回合中有二十个实际执行 VLA。十次成功中，有五次 VLA 动作数为零：第二至第五轮的指令抓取，以及第三轮的堆叠。这些回合说明混合接口仍能选择 H₀/IK 路径，其成功不能归给 Simate 输出。[L10]

使用了 VLA 也不保证成功。第三轮笔筒记录了 1004 次 VLA 动作，原生 Score（0 至 1）为 0.10；同轮中央抓放记录 261 次 VLA 动作，Score 为三分之一；对齐记录 45 次，Score 为零。这些记录要求后续研究检查委托内容、低层遵循和返回高层的时机，不能只统计工具调用是否存在。

“使用过 VLA”的二十个回合由策略自行选择，任务状态和难度可能与未使用子集不同。因此不能把两个子集的成功率差作为 VLA 的随机对照效应。合适的后续消融需要固定任务、布局、源码与预算，再改变工具可用性或委托规则。

8.5 调用量与费用

从选定的 75 个完成槽位重算，H₀ 组共有 859 次 Astra 调用，中位数为每回合 33 次；混合组为 470 次，中位数十次；独立 Simate 为零。这些调用统计不含原始中断尝试与离线视觉复盘。[分析记录](#)

范围	Astra + H ₀	混合组	独立 Simate
选定完成回合的 Astra 调用总数	859	470	0
每回合 Astra 调用中位数	33	10	0
每回合原生动作数中位数	458	342	290
生成费用占用，含原始中断与恢复尝试	\$74.489065	\$44.112965	\$0 Astra API

混合组费用中，\$43.618896 已结算，\$0.494069 为未知请求预留。独立 Simate 仍使用本地推理与模拟器资源。三组动作数中位数还受到任务预算、提前成功和提前终止的影响，不能单独作为执行效率排名。

已有记录表明，长动作委托可以减少高层调用，但更少调用也可能来自更少修复或更早结束。要比较总成本，需要把成功率、动作模型计算、仿真时间、失败尝试和复盘开销放在同一协议下。本文的 API 费用表只描述已记录的研究支出。[L10]

9. 执行失败揭示的接口问题

9.1 夹爪命令与实测开度

早期记录中，闭合请求仅保持三个 tick，后续等待和抬升又把受物体阻挡的实测开度作为新目标，导致闭合命令丢失。控制器随后改为保留最后一次明确夹爪请求。这一问题发生在动作执行链路，若只查看上层“抓取”决策，很容易把它归为视觉定位或规划错误。[L1]

修复所支持的结论是命令语义恢复一致，而非所有抓取都已成功。夹爪受阻既可能说明接触目标，也可能说明碰到障碍；仍需结合图像、实测位置与后续物体运动判断持有关系。

9.2 姿态声明与实际四元数

历史工程回合出现过模型声明的工具轴方向与四元数实际方向相差约 60° 的情况。后续检查允许 5° 数值误差，明显矛盾时拒绝动作并提供一次额度内修复机会。它使语义描述与数值请求可以相互核验，避免矛盾指令进入执行器。[L1]

即使两种表述一致，目标仍可能不可达，或不适合接触。姿态一致性检查、IK 可达性、运动执行和任务判定应保留各自的失败原因，使模型下一次修正针对实际问题。

9.3 子目标遵循与反馈修正

v12 的一个成功笔筒回合中，Simate 两次选择了不同于高层指定颜色的笔，Astra 从后续图像中修正进度，最终仍完成任务。该回合的 72 次 VLA 请求最长为 184 token，没有超过 200-token 输入限制，因此这些错选不能直接归因为文本截断。[L1]

这条轨迹说明，反馈可以让高层适应低层已经发生的动作，但不能证明低层严格遵循子目标。后续应分别测量指定对象、实际操作对象和任务进展，在已有错误上继续规划与预先避免错误可能具有不同成本。

9.4 完成声明与继续执行

一个工程笔筒回合中出现 25 次 done 和十八次历史查询，最终原生 Score（0 至 1）为 0.40。首次 done 之后的四十九次调用花费 \$5.512331，占该回合费用的 85.4%；期间还发生了观察、运动和回位，不能全部记为无效开销。[L1]

同一审计对前三轮 45 个研究回合发现，八个包含 done 的回合中有七个后来成功，且声明后仍有实际动作。这意味着“首次声明后立即终止”会改变已有成功路径。更合适的研究问题是：如何为完成假设分配有限的验证动作，以及何时返回未确认状态。当前还没有这一机制的受控收益结果。

这些案例将失败定位到可检查的接口或事件，同时保留了竞争解释。它们可以生成下一轮改动假设，不能替代同条件父子比较。

10. 跨机型与真机工作

10.1 跨机型仿真

后续笔筒矩阵比较 Astra_Direct 与独立 Simate，计划覆盖双 ARX5、双 RealBot 和双 RM75 共 90 回合。本文沿用已封存的 10 月 19 日材料快照，其中审核完成 63 回合、31 组完整配对。该矩阵的模型接口、安装与适配条件不同于前述 H_0 架构研究，因此单列于附录，不合并分母。[L5]

另外的 measured 矩阵在该快照中仅有 3/90 完成。不同矩阵的进度、源码与结果分别保留，旧状态文件中的 running 不作为进程仍在运行的证据。

10.2 标定与初始化

10 月 19 日的一次 RealBot 标定仿真完成 3000 个原生动作、65 轮 Astra 指令，原生 Score 为零，success=false。三路录像各有 3001 帧；墙钟约 54.2 分钟，模型等待累计约 799.15 秒，API 结算 \$4.202765。[L6]

后续核查发现，模型动作之前左臂 J6/J7 已偏离待机配置，腕部偏移约 15.63 cm，独立诊断可以复现这一异常，原因尚未确定。该结果保留为完整失败记录，但不能单独归因于高层模型。正式比较需要先核对 reset 后的实测机器人状态，而不只检查相机图像和接口能否返回。

10.3 真机 harness 与技能层

RealBot 工作区从 H_0 分支扩展了关键点记忆、极线辅助、关系对齐和短动作队列。持物关键点可以随实测手腕更新，用于表达两手之间的对准关系；这一更新依赖物体与手之间近似刚性，滑移或重新抓取后需要重新观测。动作拒绝时返回的数值残差，则用于帮助模型区分姿态、可达性或执行问题。[L7]

该真机主驱动没有配置在线 VLA，结束原因中的 controller_done 也没有等价的原生任务评分。因此，本报告不据源码接入情况给出真机成功率或仿真到真实的迁移收益。

10 月 20 日在这一 harness 上加入可选的技能层 (--skills)。新增能力全部以 Astra 工具的形式出现，任务分解、目标选择和结果核对仍由模型完成：

- `detect_objects`: 几何查询, 在当前胸部 RGB-D 帧中找出桌上的彩色物体。直立物体按桌面平面网格聚类, 用 RANSAC 拟合筒口圆; 平放物体给出中心、轴向、长度和宽度。编号图在下次调用中提供给模型。
- `pick_object`、`grasp_holder`、`present_holder`、`insert_into_holder`、`pick_and_insert` 与 `place_holder`: 开始执行时按实测关节、最新检测和当前胸部帧重新规划, 展开为若干普通的 `move`、`grip` 和 `elbow` 原语。这些原语与模型直接下发的动作一样, 逐 tick 经过安全层。

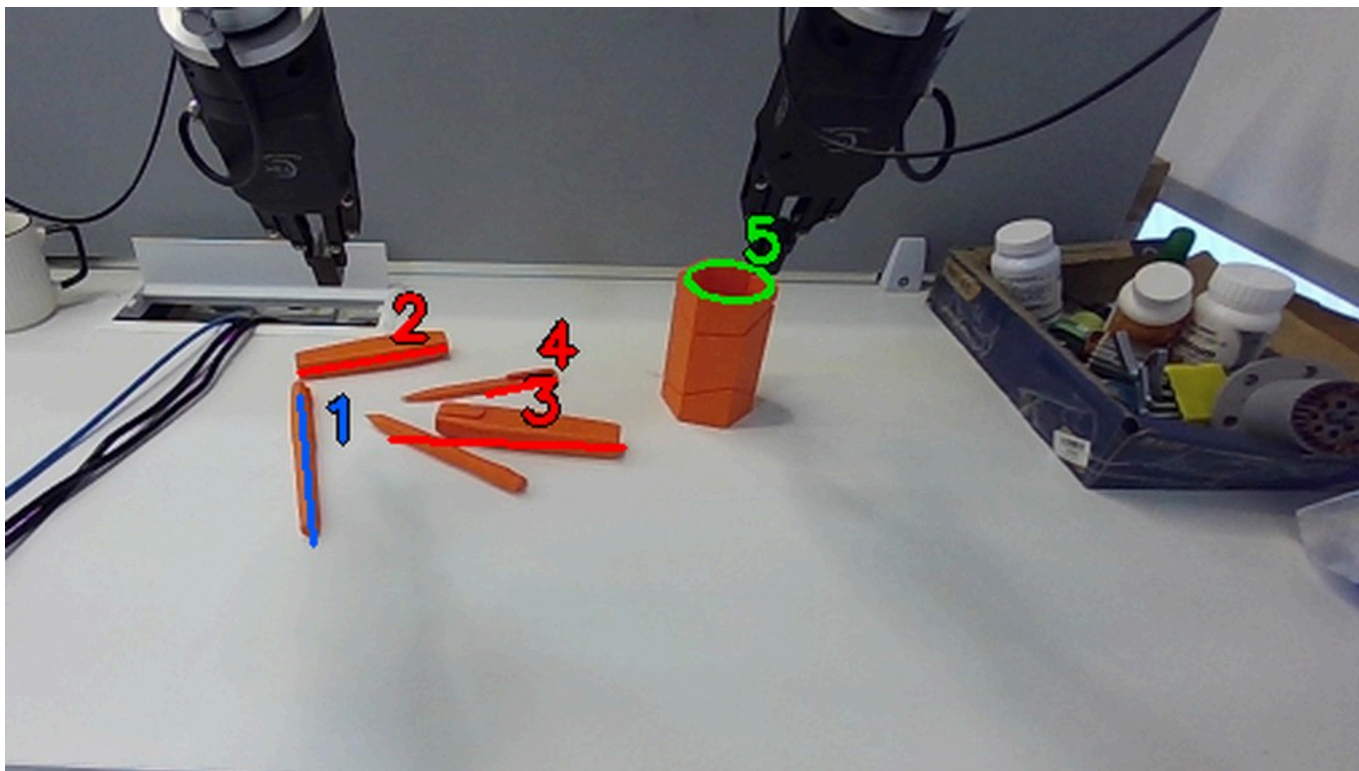


图 5: 第四次运行开始时 `detect_objects` 返回给模型的编号图 (裁去画面左侧墙面标识)。1-4 为平放的笔, 给出轴线; 5 为直立笔筒, 给出筒口圆。编号只用于本次对话, 下一次检测重新编号。

规划约束有五条: 腕部相机保持在手的上侧且不朝上看, 距桌面不少于 5 cm; 单段关节转角设上限, 较大转动拆成插值位姿; 双手距离小于 35 cm 时保持肘部外展, 与安全层的双手姿态规则一致; 夹爪闭合时指尖会前移约 25 mm, 规划按张开时的指尖位置计算, 闭合后用实测手腕位姿登记持物; 插入前先把持物提到安全高度再转向, 并以持物下端对准筒口轴线。这些约束都来自当天的失败记录。安全层同时加入腕部相机点, 要求它距桌面至少 3.5 cm。

10.4 真机插笔回合

场景为一个六棱柱笔筒 (对边约 70 mm、高约 124 mm) 和五支粗细不同的散放笔, 任务指令与仿真笔筒任务相同: 一只手持筒, 另一只手放入所有笔, 最后放回笔筒。当天共有四次 Astra 运行, 均使用 gpt-6-astra, 动作全部自动批准, 无遥操作。

运行	推理档 / 速度	tick	模型调用	费用	结束原因	插入支数
1 (14:46)	high / 5	108	15	\$1.71	取笔下探时指尖接近桌面, 操作员急停	0
2 (15:19)	medium / 10	48	6	\$0.64	夹筒动作反复被双手姿态规则拒绝, 操作员停止	0
3 (15:24)	medium / 10	79	6	\$0.66	急停后夹爪读数为零, 已夹住的笔筒被判为空抓, 操作员停止	0
4 (15:32)	medium / 10	752	47	\$5.80	第五支落在筒外后, 操作员停止	4

前三次运行暴露的缺陷在下次运行前修正：技能在双手靠近时保持肘部外展；夹爪读数低于预期时不再直接终止技能，而是标为未核实，由模型查看腕部图像。更早一次不经模型的脚本化持筒测试中，单步关节空间转动使腕部翻转，腕部相机擦碰桌面，操作员挪开底盘后重新标定了桌面。这次事件促成了腕部相机碰撞点和上面的姿态约束。

第四次运行中，模型先调用 `detect_objects`，再调用 `grasp_holder`（右手）和 `present_holder`，把笔筒端到胸前的递交位置，然后用 `pick_and_insert` 完成第一支。第二支的技能两次找不到满足约束的抓取，模型改用底层的 `grasp` 和 `move`，左手手指碰到筒口后退开。15:47 操作员提示模型回到技能流程，之后第二支以技能插入。第三支在张开夹爪时两次被双手姿态规则拒绝；15:55 的提示让模型先把左手移到 35 cm 以外，之后插入成功。第四支插入途中，安全层的双臂体积模型（持物按球体计算）判定两臂将要接触；模型退开后用 `insert_into_holder` 从更高位置重新规划并插入成功。这一段是连续的“故障、修正、恢复”记录。第五支插入时笔筒已较拥挤，技能仍对准筒口中心，笔落在笔筒右侧桌面。16:07 的提示允许右手调整笔筒姿态，模型开始倾斜笔筒后，操作员停止运行。

模型在回合内的修正可以从决策记录中逐条核对：每个决策都带有模型写下的理由与记忆。第二支的技能两次规划失败后，模型自己提出改用靠近笔端的倾斜抓取，先做 IK 探测再执行，但插入时手指碰到筒口而未成功。15:55 的提示之后，模型把“先抬高 100 mm、再退 200 mm，使两手相距 35 cm 以上”写入记忆，此后每次取笔前自行执行。第四支插入被安全层拦下时，模型读取拒绝原因（转动途中两臂间距将低于 0.5 cm），自行退开并从更高位置重新规划，插入成功；这一修正发生在任何相关提示之前。第五支落在筒外后，模型在右腕图像中找到粗笔左侧的空位并开始倾斜笔筒，这发生在 16:07 允许调整笔筒姿态的提示之后。

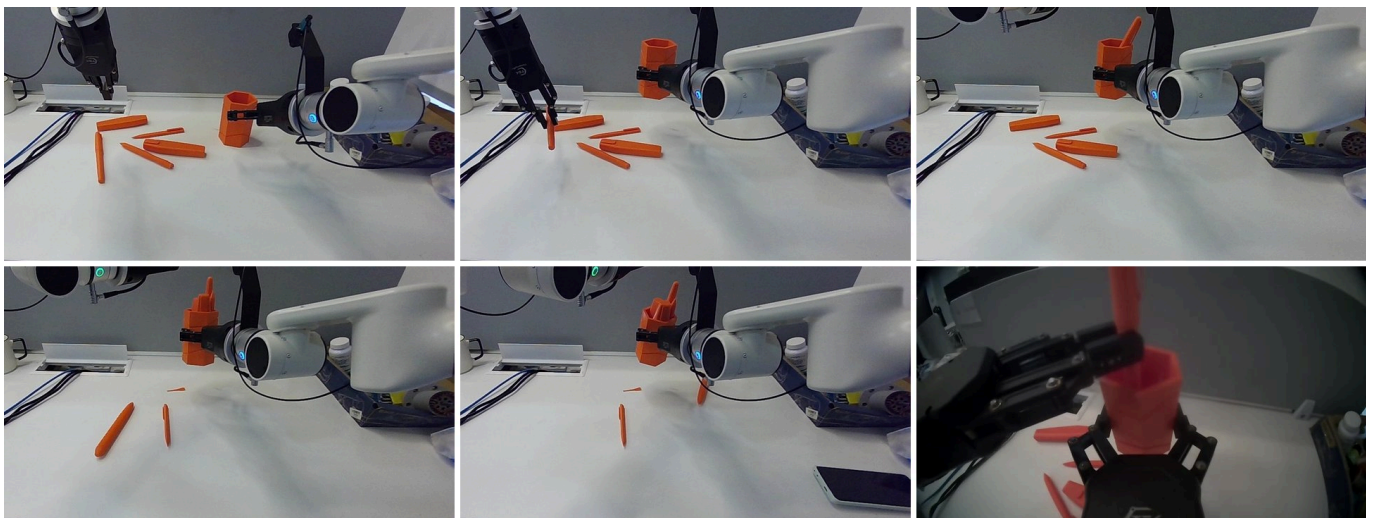


图 6：第四次运行。上排从左到右：右手持筒并端到胸前；左手从上下方采取笔；第一支立在筒内。下排：四支已在筒内；第五支落在笔筒右侧桌面；第一支插入时的右腕视角。均取自 MCAP 原始帧，画面不加文字。同一回合的三段右腕近景见网站真机一节：第一支，2 倍速、第四支被拦下后恢复，8 倍速、第五支失败，8 倍速。

全程执行 748 个 tick，安全层拒绝 4 次。操作员没有遥控，也没有动手搬动物体，干预只有三条文字提示和最后的停止。笔是否插入以右腕和胸部图像为准，由模型和操作员核对，没有独立的原生评分。该回合按单次演示报告；三条提示属于人工介入，若用于训练或比较，须单独标注。[L16]

运行之间的修正由工程人员完成。为检查模型能否从同样的失败证据中得出这些修正，我们在事后做了一次盲审：对每次失败或停止的运行，只给 gpt-6-astra 当时运行的 harness 代码、原始运行日志（含它自己的决策文本与安全层拒绝信息）和 3–6 张图像，不提供工程诊断、提交说明或之后的代码；另一个独立代理先清除可能泄露答案的措辞。每例调用一次、不重试，两名独立评审对照实际代码改动判定。四个病例中，模型都指出了与工程修正相同的失效机理；在运行 1、2、4 中提出了实质相同的核心改法（按两个指尖与实际张开量规划接触；所有技能都按双手情形做姿态检查；规划时使用运行时的两臂碰撞检查）；运行 3 中它同样判定读数不可信，但主张先小幅试抬确认附着再登记，比工程修正更保守。八份评审均为“部分一致”，没有一例覆盖工程修正的全部改动。模型另外指出了若干工程

尚未处理的问题，其中持物判定 `held()` 的“被阻挡”比较方向错误经代码核对属实。该结果说明模型能从失败记录中提出可用的 harness 改动，不说明这些修正由模型完成；运行 4 对应的工程修正尚未上机。网关在请求前注入了一段与任务无关的系统提示，约 4k tokens。[L18]

10.5 真机记录与仍缺的证据

第四次运行的回合目录约 45 GB。其中 MCAP 包含头部双目与两腕相机的 H264、30 Hz 胸部原始 JPEG、约 190 Hz 的双臂与躯干状态；每个 tick 保存命令目标、安全层结论和到位误差；每次模型请求与输出均保存，包括图像；另有约十次 `detect_objects` 的编号图，以及不含密钥的运行时代码快照。同一回合胸部视角的 8 倍速剪辑约 2.4 分钟，取自 MCAP 原始帧，裁去画面左侧的墙面标识，不加文字；另从右腕 H264 流解码三段近景，对应第一支（成功）、第四支（恢复）和第五支（失败）。各数据流按 ROS 消息时间戳对齐，剪辑区间与倍率可由剪辑脚本复现。

这些材料对应缺口清单中的 P0 连续全景与同回合近景、D01 同步观测与实测状态、D02 版本身份，并提供一段恢复片段（第四支）和一个失败负例（第五支）。仍缺：逐支的独立结果标注、多回合成功率、无提示条件下的对照、真机 TaskGraph 标签，以及技能层本身的消融。本稿之后已修正的工程缺陷，包括筒口空位估计、在规划中使用双臂体积模型、取笔前先远离对侧手，尚未经真机回合验证。具体条目见 [真机素材缺口清单](#) 与 [采集跟踪表](#)。配对改进录像属于后续研究材料，须在候选物化与协议冻结后采集。

11. 局限与下一阶段研究

TaskGraph 已完成处理与消费链路的接入，本次材料仍没有独立分类标签。恢复候选需要盲标核验，尤其要检查进度回退是否真是故障、修正命令与恢复是否具有正确时间顺序，以及原实验视频能否支持重放状态的解释。没有标签时，候选数量与处理完成率不能代替 precision 或 recall。

新的真实复盘也有范围限制。它只使用十八帧抽样，诊断回合不具研究改进资格；分类规则在该原生回合后修正并重算，最终 hook 尚未再跑一个新的 GPU 回合。后续应使用新 experiment ID 冻结源码、规则与数据划分，再执行 development 回合、物化候选和父子配对选择。

四项文本提案为后续研究提供了可测试的方向，但当前没有新的 H_k 性能。比较 TaskGraph 摘要的价值，需要固定基础模型、上下文与提案预算，对照原有复盘输入和增加任务图后的输入；先核对实际消费与候选差异，再比较被选择策略的独立结果。候选选择还应保留等预算随机提案基线。[R4]

历史五轮架构结果跨四个源码版本，每个任务与版本的样本很少。它适合分析工具分工与接口问题，无法为 TaskGraph 新接入或某个工程修正估计因果收益。是否使用 VLA 由策略自行选择，也需要受控路由消融处理选择偏差。

真机结果只有一个回合，且包含三条人工文字提示。技能层针对笔筒任务编写，它的作用与模型本身的作用尚未分开测量；第二至第五支暴露的问题有一部分由提示弥补，一部分已在代码中修正但尚未在新回合中验证。下一步应在提示通道关闭的条件下重复多个回合，逐支记录结果与介入，并保留失败回合。

本阶段的结果是执行经验能够被保存、结构化并送入真实复盘，同时产出恢复候选和明确的改动假设。下一阶段的工作是对这些候选逐项核验，并完成冻结条件下的选择与独立测试。

附录 A. 补充研究快照

下表只包含跨机型矩阵中已经审核的回合，各机器人均在该矩阵的仿真条件中运行。双 RealBot 两组样本数不齐，不计算成完整配对的成功率差。

机器人	Astra_Direct 成功 / 已完成	Simate 成功 / 已完成
双 ARX5	0/11	3/11
双 RealBot	0/11	0/10
双 RM75	0/10	0/10

这些是部分样本的描述，无法支撑跨机型迁移有效或模型普遍优劣的结论。来源、原始矩阵和历史 snapshot 见 [data.json](#)。

附录 B. 复核方式与数据口径

报告使用已经冻结的 75 回合 CSV。analyze_results.py按轮次、任务和布局检查配对完整性，计算总体、逐轮、成功配对、Score 配对及调用统计，结果写入 analysis.json。输入 CSV 的 SHA-256 随分析记录保存。此过程只重算已有数据，不产生新的实验样本。

每条回合记录可以回到原始 episode、原生结果和配置。源码哈希说明执行字节身份，布局哈希说明声明的场景身份；它们不能证明物理结果确定，也不能单独证明历史上从未曝光。实验费用按请求 ID 去重，区分选定完成槽位、原始中断、复盘和工程运行。

媒体来源记录历史成功回合三路视频的路径与哈希；新接入诊断的三路录像另见 TaskGraph 证据快照。视频视角属于同一回合，不增加样本数。TaskGraph 重放侧录与原视频如果不是同一次采集，需保持各自的结果和时钟，不能把原视频当成新重放的物理实录。

网站另提供 75 回合的 225 路完整历史录像、3 路接入诊断录像、76 份可交互任务图与 4 部流程视频，索引见 video-library.json，逐部视频的来源、剪辑范围和讲解见 录像与任务图浏览说明，派生视频的来源与渲染代码哈希见 process-video-provenance.json。同一回合的多个视角不增加样本数；流程讲解与跨轮片段不计为新实验或候选收益。

TaskGraph 数据快照保留本次筛选、真实诊断、四项建议、时间映射和输入哈希；提取脚本只读取已有产物。evidence/ 保存可独立查看的摘要、复盘输出与验证记录。

正文保留实现和实证所需的信息；会议分工、域名和部署安排见 [资料与会议记录及交付说明](#)。

参考与证据索引

- R1. [PhysicalRSI 项目页](#)。方法与页面设计参考。
- R2. [PhysicalRSI 架构说明](#)。执行与开发两种时间尺度。
- R3. [PhysicalRSI RoboDyna pilot](#)。候选选择、负结果与迁移研究。

- R4. PhysicalRSI Core 开发与实验记录；研究记忆。
- L1. 历史完整技术报告 [../../physicalRSI/technical_report/REPORT.md]；本稿 75 回合数据。方法与工程案例采用其对应历史版本。
- L2. 配置优化协议 [../../physicalRSI/harness_optimization/GENOME_ITERATION.md]；实际交付状态 [../../physicalRSI/DELIVERY_STATUS.md]。
- L3. TaskGraph SDK 方法 [../../taskgraph/SDK_DESIGN.md]。
- L4. RoboDojo TaskGraph 接入与故障环候选 [../../taskgraph/ROBODOJO_INTEGRATION.md]；图与视频联动 [../../taskgraph/RESULT_VIEWER.md]。接入实现与真实消费已完成，收益仍以独立实验为准。
- L5. 跨机型比较 [../../eval_result/Astra_vs_VLA/matrix_15_steps3000_final/comparison.html]；输入文件哈希见 data.json。
- L6. 标定仿真实验与初始化纠正 [../../realbot_calibrated_astra_20261019.md]。
- L7. 真机与仿真 harness 对照 [../../double7R_Arm/HARNESS_COMPARISON_20261019.md]。
- L8. GRIT v0 本地说明：/root/autodl-tmp/grit/README.md。用于方法定义，未引用未经核实的检测指标。
- L9. 五轮三组研究协议 [../../physicalRSI/FIVE_ROUND_THREE_ARM_PROTOCOL.md]。
- L10. 75 回合完整结果、暴露与费用说明 [../../physicalRSI/V13_COMPLETE_RESULTS.md]。
- L11. TaskGraph 回合后恢复筛选与 RSI 接入 [../../taskgraph/RSI_RECOVERY.md]。
- L12. 接入状态；原生数据完整性验证。
- L13. 75 份侧录筛选摘要；待独立标签的分类评价。
- L14. 任务图、复盘与时间映射快照。
- L15. 真实复盘原文与四项候选建议；任务图反馈摘要。
- L16. RealBot 真机插笔回合摘要：四次运行的统计、第四次运行的逐段时间线、人工介入、数据位置与哈希。
- L17. RealBot H₀ 与技能层代码：第四次运行时的运行时快照 code_snapshot_run4.tgz（SHA-256 4aa584f2...3645e，不含密钥），随回合数据保存；本地备份为机器人 ~/claude_ros2 当日快照（git HEAD b40bb37 加未提交的技能层改动）。
- L18. Harness 盲审记录：四个病例的协议、请求 ID、用量、模型完整输出、评审结论与模型额外指出的问题。

公开参考的版本索引与阅读范围见 调研记录。