

2026 丘成桐中学科学奖

计算机科学

记忆写入可能破坏未来证据 可变智能体记忆的递归验证契约

参赛学生：谢子涵
学校：华东政法大学附属松江高级中学
省 / 州：上海
国家 / 地区：中国
指导教师：林海博、陈森
指导教师所在单位：华东政法大学附属松江高级中学
研究方向：智能体记忆、序贯决策、状态抽象

研究报告

记忆写入可能破坏未来证据： 可变智能体记忆的递归验证契约

谢子涵

华东政法大学附属松江高级中学，上海

摘要 用摘要替代仍可访问的证据源，即使保留了当前答案，也可能改变智能体日后能够验证的内容。对任意有限预测深度 (prediction depth) k ，我们构造一个有限时域过程 (finite-horizon process)，其中成对历史 (paired histories) 在深度 k 以内不可区分，但两条历史的最优初始记忆动作相反，且最优动作与备选动作之间的动作价值间隔 (action margin) 保持固定。动作执行和观测传递均无延迟；即使已有事实证据 X 被完整保留且不额外收费，这种分离仍然存在。我们量化将需要不同最优动作的历史合并为同一表示所造成的损失，并刻画未来访问成本未知时记忆必须保留的信息。在一族访问成本上取平均时，压缩带来的额外遗憾 (additional regret) 恰好正比于压缩前后生命周期后验概率 (lifecycle posterior) 的平方失真 (squared distortion)，即两个后验概率之差的平方的期望。这一损失可分解为契约形成阶段的损失和后续编码损失；在后验概率均匀分布的模型中，当记忆最多可存储 K 个符号时，成本平均遗憾 (cost-averaged regret) 和最坏成本下的遗憾 (worst-cost regret) 均有精确的容量界。当需要保留完整的受控过程时，经典递归闭包 (recursive closure) 给出充分契约；在已知有限模型中，划分细化 (partition refinement) 达到稳定即可停止。受控学习实验检验了相关区别能否被保留，或能否根据有噪声的历史审计 (past audits，即以往的可用性检查) 推断出来。文件与超文本传输协议 (Hypertext Transfer Protocol, HTTP) 实验直接检验后续证据访问，并设置使动作反转消失的对照。一个小规模 SQLite 实验测量了保留活动快照 (live snapshot) 的存储负担，并展示完整离线备份如何避免这一负担。仅保留当前事实，并不能保证后续决策合理。

关键词 智能体记忆；持久记忆 (persistent memory)；验证；序贯决策；递归状态抽象 (recursive state abstraction)；互模拟 (bisimulation)；有限容量表示；大语言模型智能体

Memory Writes Can Destroy Future Evidence: Recursive Verification Contracts for Mutable Agent Memory

Xie Zihan

Songjiang Senior High School Affiliated to East China University of Political Science and Law, Shanghai, China

Abstract Replacing a live evidence source with a summary can change what an agent can verify later, even when the current answer is preserved. For every finite prediction depth, a finite-horizon process is constructed in which paired histories remain indistinguishable through that depth but require opposite optimal initial memory actions, with fixed action-value margins. Actions and observations are immediate; the separation can persist even when existing factual evidence is retained in full and no additional fee is charged. The loss caused by merging histories that require different optimal actions is quantified, and the information that memory must retain when future access costs are unknown is characterized. Averaged over a family of access costs, the additional regret from compression is exactly proportional to the expected squared difference between lifecycle posteriors before and after compression. This loss separates into loss at contract formation and loss during subsequent encoding. With a uniform posterior and a finite storage alphabet, sharp capacity bounds hold for both cost-averaged and worst-cost regret. When the complete controlled process must be preserved, classical recursive closure provides a sufficient contract; in a known finite model, refinement can stop when the partition becomes stable. Controlled learning experiments test whether the relevant distinctions can be retained or inferred from noisy past availability audits. File and Hypertext Transfer Protocol (HTTP) experiments directly test later evidence access, including controls that remove the action reversal. A small SQLite experiment measures the storage burden of retaining a live snapshot and shows how a complete offline backup avoids it. Preserving the current fact alone does not guarantee a sound later decision.

Keywords agent memory; persistent memory; verification; sequential decision making; recursive state abstraction; bisimulation; finite-capacity representation; large language model agents

目录

0 引言	4
1 背景与相关工作	4
2 问题表述：可变验证契约	5
3 固定浅层深度的局限	7
4 有限记忆：深度、容量与策略遗憾	8
5 充分契约：精确闭包与决策证书	12
6 学习决策所需的区分信息	14
7 文件、网络与数据库中的证据访问	17
8 讨论	21
9 结论	21
研究过程与致谢	22
参考文献	23
附录 A 完整理论细节	24
附录 B 有限容量结果的证明	28
附录 C 受控学习实验	31
附录 D POSIX 权威证据源到期协议	33
附录 E 网络签名权能凭证生命周期协议	34
附录 F SQLite 快照保留协议	36
附录 G 计算验证	36

0 引言

智能助手可能在尚不知道下一次查询内容时，就总结对话、整合工具结果，或用紧凑的持久记录替代长历史^[1-2]。这类写入通常按记录保留了多少信息来评价，但是写入也可能影响记录之外的资源，譬如已打开的句柄可能被关闭，证据源可能会过期，可恢复的记录可能被覆盖。同时，近期研究已记录长期记忆整合中的遗漏、损坏、来源信息（provenance）丢失、权威性坍塌（authority collapse）以及性能退化等问题^[3-7]。

用两条历史即可说明这一机制。两条历史包含相同的当前事实与证据源的当前观测。如果智能体等待，那么证据源在一条历史中仍可查询，但在另一条中就会消失。因此，只存储当前答案的记忆可能将两条历史合并，尽管第一条历史中等待最优，第二条历史中提交快照最优。缺失的变量不是当前事实，而是记忆动作所诱导的未来验证概率律。会过期的会话、临时工具句柄、被删除的文件和可变缓存都具有这种结构。

本文用递归验证契约（recursive verification contract）描述可变智能体记忆（mutable agent memory）在写入后需要保留的信息。第一个结果表明，任何预先指定的有限预测深度都不足以对所有有限系统保证后续决策最优。即使动作执行与观测传递均无延迟，仍可构造这样的反例，而且两类历史中最优动作与备选动作的价值差不随预测深度增加而缩小。即使完整保留已有事实证据 X 且不收取额外费用，上述反例仍然成立：主目标与辅助目标竞争唯一的活路（live route）名额，由此产生机会成本（opportunity cost）。

第二个结果刻画了未来访问成本未知时，记忆需要保留的信息。在保留既有证据的模型中，保留生命周期后验概率是在整个成本族上保持最优风险的充要条件。压缩造成的额外成本平均遗憾与该后验概率的平方失真成正比，比例系数为 $g/2$ ，其中 g 表示持久型与易失型通路的未来信息价值之差。这一损失可分解为契约形成阶段的损失和后续编码阶段的损失。若后验概率服从均匀分布，以 K 表示可用的存储符号数，则平均遗憾与最坏成本下的遗憾均有精确的 K^{-2} 容量律。这些结果明确了写入时需要保留的信息，并量化了省略这些信息所造成的损失。

第 4、5 节区分面向特定决策的记忆和完整的递归充分性。后者使用已有的状态抽象方法；成本族分析也建立在经典二元统计试验恒等式之上。学习实验保持当前信息不变，检验监督是否有助于恢复后续决策所需的生命周期信息。采用文件、网络和数据库协议的实验检验失去或保留证据源访问权的影响。保留预览、扩大通路预算以及复制完整证据源的对照，说明这种权衡何时消失。

1 背景与相关工作

1.1 智能体长期记忆

在近期的系统设计中，记忆不再被当作一个静态检索库，而被看作具有生命周期的对象^[1]。分层记忆把写入安排在查询之前，并让不可变证据充当恢复层；WorldMemArena 则借助智能体与世界的交互来考察记忆^[8]。整合带来的退化、可信整合、来源漂白（provenance laundering）、权威性坍塌、类型化记忆、事务式写入，以及把“预测下一步”当作写入准则的局限，都已有工作涉及^[3-7,9-10]。本文据此研究证据可用性变化时所需的预测深度，以及记忆必须保留的区别。

1.2 信息状态与状态抽象

充分信息状态（sufficient information states）、受控感知（controlled sensing）、预测状态表示（predictive state representations）、概率互模拟（probabilistic bisimulation）与近似行为度量（approximate behavioral metrics），都为概括受控过程提供了理论方法^[11-14]。把待处理信号一并写进状态，延迟控制模型同样能恢复马尔可夫结构（Markov structure）^[15-16]；而在固定的决策问题中，单周期动作充分压缩（single-cycle action-sufficient compression）关心的则是码率与遗憾之间的取舍^[17]。本文所用的状态抽象与决策论工具，正来自这些框架。

一个持久记忆系统要在未来任务尚不清楚时就决定留下什么，而这次写入本身又可能决定哪条验证通路继续存在。定理 3.1 指出：即使动作执行与观测传递均无延迟，任何预先选定的有限预测深度都可能无法区分需要不同最优动作的历史，且两类历史中的动作价值间隔保持固定。递归闭包对应于划分细化达到稳定的状态：继续比较各动作下的观测与后继类别分布，也不再分裂任何类别。在保留任务所需基础标签的有限受控模型中，这样的稳定划分足以保持后续决策价值。第 5 节给出这一结果，其中的细化方法取自经典状态抽象文献。

2 问题表述：可变验证契约

2.1 最小持久型 / 易失型见证例

先从包含完整机制的最小见证例 (witness) 出发。记忆动作有提交 (commit, C) 与等待 (wait, W) 两种，用 V 表示动作价值。提交建立可靠度为 0.80 的快照，并支付 0.06 的成本，因此

$$V(\text{提交}) = 0.80 - 0.06 = 0.74. \quad (1)$$

等待保留原有证据通路。在持久型 (durable) 历史中，后续验证的可靠度为 0.95；在易失型 (fragile) 历史中，证据源消失，智能体只能退回可靠度为 0.60 的预测。分别以 h^D, h^F 表示这两类历史， π^* 表示最优策略， m_D, m_F 表示各自最优动作与备选动作的价值间隔。因此

$$\pi^*(h^D) = \text{等待}, \quad m_D = 0.95 - 0.74 = 0.21, \quad (2)$$

$$\pi^*(h^F) = \text{提交}, \quad m_F = 0.74 - 0.60 = 0.14. \quad (3)$$

两条历史的当前事实和当前验证状态相同，区别仅在于证据通路的未来存续情况。图 1 概括了这一分支结构。

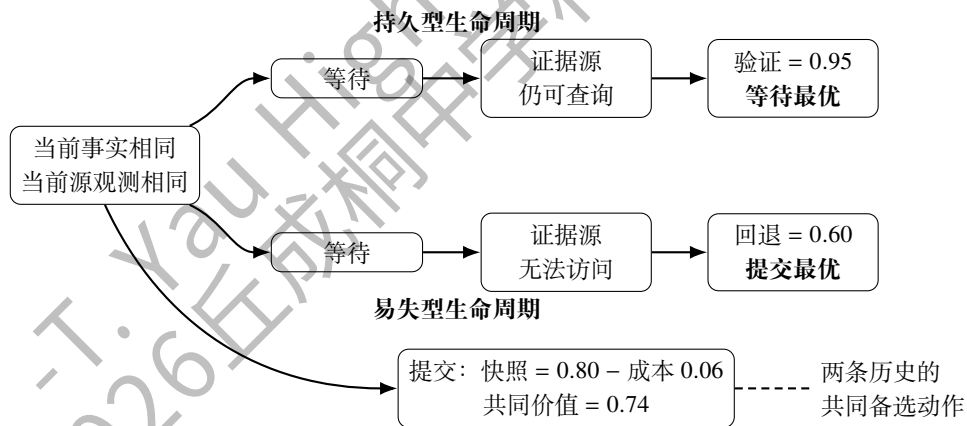


图 1 同一当前契约下的两种隐藏证据生命周期

注：两条历史具有相同的当前事实和证据源观测。等待使生命周期差异在之后显现，而提交在两条历史中的价值均为 0.74，从而使最优初始记忆动作反转。

2.2 一般递归契约

记忆写入时，用 h 表示写入前的历史，以 $\mathcal{A}$ 表示可用动作集， $a \in \mathcal{A}$ 表示记忆 / 控制动作。 O_{t+1} 是下一次观测， $h(a, o)$ 是执行 a 并观测到 o 之后的历史。基础契约 $S_0(h)$ 存储当前任务要求区分的信息。在下述反例中，它保留当前事实，却省略证据源的生命周期信息；这一选择仅用于构造反例，一般契约应保留什么信息，取决于它需要支持的决策。对于后续决策问题，用 $Q^*(h, a)$ 表示最优动作价值。

“隐藏”指契约省略了生命周期信息。在持久型 / 易失型见证例中，完整历史能够区分两种生命周期，基础契约却丢失了这一区别。若原始历史仅提供有噪声的线索，则应根据后验概率决策，不能直接采用已知生命周期时的理想动作。第 4.4 节分析这种后验决策，第 6.4 节检验如何从审计证据推断所需信息。三者的关系如下：

$$\text{真实生命周期} \rightarrow \text{原始历史提供的证据} \rightarrow \text{契约与编码实际保留的信息} \quad (4)$$

原始历史可能不足以完全识别真实生命周期，契约形成与编码又可能丢失历史中已有的信息。这两类不确定性的来源不同，需要分别处理。

定义 2.1 (递归签名). 给定基础契约 S_0 ，对 $d = 0, 1, \dots$ ，递归签名 (recursive signature) 定义为

$$S_{d+1}(h) = (S_0(h), \{\mathcal{L}(O_{t+1}, S_d(h(a, O_{t+1})) \mid h, a)\}_{a \in \mathcal{A}}). \quad (5)$$

其中 $\mathcal{L}(X \mid \cdot)$ 表示给定历史与动作时，随机变量 X 的条件概率律 (conditional probability law)。若两条历史的 S_d 签名相同，则称它们在深度 d 上不可区分。

这一递归定义可按层次理解。 $S_0(h)$ 记录当前任务要求区分的信息； S_1 进一步记录每个动作下一次观测与后继基础类别的联合分布； S_2 则继续区分这些后继历史所支持的一步统计试验。因此，递归签名同时追踪未来观测及观测之后仍可用于决策的信息。

递归签名保存“观测—后继签名”的联合概率律。未来策略依据观测选择动作，因而观测与后继状态的对应关系会影响决策。即使两个边缘分布分别相同，不同的对应关系仍可能导致不同的后续最优动作。

如果某个记忆动作在允许的后续策略下，改变了某个未来权威验证事件的概率，就称其具有验证干预作用。这将被动映射 $h \mapsto m$ 与同时改变证据通道的写入区分开来；后者还会改变日后用来核验 m 的途径。

是否必须保留来源信息，取决于所声明的任务、可用动作和后续过程模型。决策决定哪些区别重要；表示容量限制能够保留多少区别。

2.3 保留已有证据时的生命周期依赖

下面考察保留已有事实证据 X 作为回退信息的变体。固定已有事实证据 $X = x$ ，并令持久型生命周期 D 与易失型生命周期 F 下关于目标的后验分布相同。两个动作都无损保留 X 。提交保存这一记录，但放弃未来访问；等待还保留一条未来验证通路，并在预留通路时支付持有成本 (holding cost) c 。令 R_0 为仅使用 X 的贝叶斯风险 (Bayes risk)， R_m 为使用 X 和生命周期 m 下可用的未来统计试验时的贝叶斯风险。记其信息价值 (information value) 为 $\Delta_m = R_0 - R_m \geq 0$ 。风险与成本使用相同的计量单位，以 L_m 记生命周期 m 下的总风险。

命题 2.2 (保留已有证据时的动作分离). 在上述假设下，

$$L_m(\text{提交}) = R_0, \quad L_m(\text{等待}) = R_m + c, \quad \text{等待 严格最优当且仅当 } c < \Delta_m. \quad (6)$$

持久型历史中的等待与易失型历史中的提交同时严格最优，当且仅当

$$\Delta_F < c < \Delta_D. \quad (7)$$

这样的区间存在，当且仅当持久型试验具有严格更低的贝叶斯风险，即 $R_D < R_F$ 。两个动作间隔分别为 $\Delta_D - c$ 和 $c - \Delta_F$ 。成本 $c^* = (\Delta_D + \Delta_F)/2$ 使较小的间隔最大化，此时两个间隔均为 $(\Delta_D - \Delta_F)/2$ 。

将两种总风险相减即可证明。 Δ_m 非负, 是因为未来的决策读出器 (reader) 可以忽略新证据; 无需假设访问失败会抹去 X 。持有成本在选择等待时发生, 而不是仅在日后发起某个可选查询时才收取。

若在给定生命周期 m 和 X 后, 通路以概率 p_m 成功访问, 且成功事件独立于目标和验证结果, 令成功验证后的风险为 R_V , 则 $\Delta_m = p_m(R_0 - R_V)$, 式 (7) 化为一个显式的存续概率阈值。一般命题不要求这种独立性: 成功或失败本身透露的信息已包含在 R_m 中。

也可以不额外收费而得到同样的结论。假设写入边界处必须将唯一的活动通路名额分配给原目标或一个独立的辅助目标。令辅助目标在不保留与保留其通路时的贝叶斯风险分别为 B_0 和 B_1 。保留原通路得到总风险 $R_m + B_0$; 释放它并转向辅助目标得到 $R_0 + B_1$ 。减去共同常数 B_1 后, 便以内生机会成本 $c = B_0 - B_1$ 恢复式 (6)。保留的活动通路数上界是访问约束, 与后文的表示字母表大小不同。

3 固定浅层深度的局限

增加若干层未来预测, 并不能在任何固定深度 k 上提供普遍保证。下面的构造保持动作间隔固定, 从而可以量化合并初始历史时的遗憾。

定理 3.1 (任意深度验证阶梯). 对任意整数 $k \geq 1$, 存在一个有限时域、采用终端效用、信号延迟为零的可变记忆决策过程, 以及历史 h_k^D, h_k^F , 使得

$$S_j(h_k^D) = S_j(h_k^F) \quad \forall j \leq k, \quad S_{k+1}(h_k^D) \neq S_{k+1}(h_k^F), \quad (8)$$

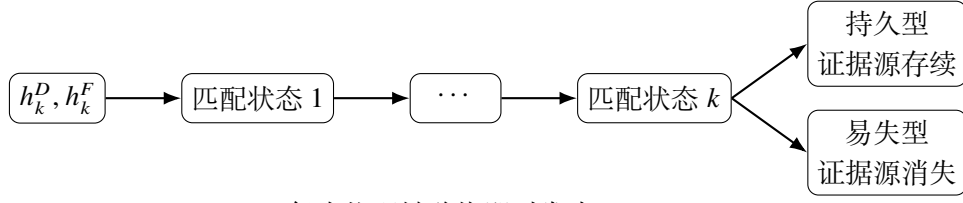
而它们在完整时域内的最优初始动作相反, 且具有式 (2)、(3) 中固定的动作间隔 0.21 和 0.14。每个动作均即时执行, 每次后继观测均在下一个决策时刻之前送达。

同一构造也给出在保留已有证据时仍成立的分离: 将其最终验证对替换为命题 2.2 的构造, 即可保留全部已有事实证据 X , 并在任意阶梯深度得到间隔 $\Delta_D - c$ 和 $c - \Delta_F$ 。即使旧事实记录完好无损, 未来访问的差异仍会导致不同的动作。

推论 3.2 (不删除预览且不额外收费时的分离). 对任意 $k \geq 1$, 存在一个信号延迟为零的有限过程, 在两个动作下都保留已有事实记录 X , 不收取额外费用, 却仍有一对历史在 S_k 及其之前的层次上相同、首次在 S_{k+1} 上可区分, 且其初始动作相反、动作间隔分别为 $1/8$ 和 $1/16$ 。写入动作在两个独立的验证任务之间分配唯一的活动通路名额。

为得到这些常数, 取两个等概率二元目标, 并使用总 Brier 损失 (Brier loss): 各目标的损失为预测概率与二元真值之差的平方, 再对两个目标求和。第一个目标已有准确率为 $3/4$ 的预览, 因此 $R_0 = 3/16$ 。其最终通路或者揭示真值 ($R_D = 0$), 或者在保留预览的情况下过期 ($R_F = 3/16$)。辅助目标的先验风险为 $B_0 = 1/4$, 未来观测准确率为 $3/4$, 故 $B_1 = 3/16$ 。于是机会成本为 $1/16$, 动作间隔为 $1/8$ 和 $1/16$ 。成对匹配的中间状态给出与定理 3.1 相同的深度归纳; 构造细节见附录 A.1。

构造思路. 将对生命周期敏感的验证转移放在 k 个成对匹配状态之后。在初始等待之后, 每个中间状态都只能执行零奖励的推进。最后一对终端前状态在递归深度为一时首次可区分; 每一对匹配的前驱恰好将这次分离向后推移一个签名层次。由于终端价值保持为 0.95 和 0.60, 原有动作间隔不会随 k 缩小。完整归纳证明见附录 A.1。图 2 展示了由此产生的递归分离。



每次物理转移均即时发生。

首次与决策相关的分离仅在递归深度 $k + 1$ 出现。

图 2 任何固定的有限验证深度都可能遗漏首次与决策相关的分离

注：持久型和易失型历史在深度 k 以内相同，首次在 $k + 1$ 处分离，且不存在传感器或执行器延迟。图中只画出等待分支：两条历史共同的提交分支及其固定值 0.74 已略去，正是它与等待分支的比较使初始动作发生反转。

与延迟控制的关系。 定理 3.1 不依赖信号延迟：所有动作即时执行，后继观测均在下一次决策前送达，标准延迟控制中的状态增广方法仍然适用。本定理关注的是记忆契约可能在写入时省略的生命周期变量；该变量对最优决策的影响只有在更深的后续过程中才会显现。两者的形式化区分见附录 A.1。

3.1 固定动作间隔下的浅层近碰撞

策略反转既不依赖精确相等，也不依赖趋近动作间隔为零的决策边界。为描述近碰撞 (near-collision)，定义一步试验距离

$$d_1(h, h') = \max_{a \in \mathcal{A}} \text{TV}(J_h^a, J_{h'}^a), \quad (9)$$

其中 TV 为总变差距离 (total variation distance, TV)， $J_h^a = \mathcal{L}(O_{t+1}, S_0(h(a, O_{t+1})) \mid h, a)$ 是定义 2.1 中以动作为条件的一步联合验证概率律。

命题 3.3 (浅层距离不存在正的安全半径)。对任意 $\epsilon > 0$ ，存在历史 h^D, h^F ，满足

$$0 < d_1(h^D, h^F) < \epsilon, \quad (10)$$

但它们在时域为二时的最优动作相反，动作间隔仍精确等于 0.21 和 0.14。

构造。从精确的持久型 / 易失型见证例出发，在一步观测中分别添加参数为 $1/2 + \delta$ 和 $1/2 - \delta$ 的伯努利干扰变量。该变量独立于原过程且与奖励无关，并在下一次决策前被丢弃。它使一步试验不再完全相同，且 $d_1 = 2\delta$ ，但既不改变后续价值，也不改变未来证据通道。取 $0 < 2\delta < \epsilon$ 即可证明。附录 A.2 给出了严格的乘积测度论证。

命题 3.3 表明，即使动作价值间隔保持为固定正数，一步距离足够小仍不能保证所有系统中的策略安全：深层验证差异并未出现在一步试验中。第 5.2 节在有限状态、奖励有界的折扣控制模型中，递归计入后续转移以控制动作价值之差。该证书依赖递归信息及上述模型条件。

4 有限记忆：深度、容量与策略遗憾

用 $\mathcal{P}_k$ 记 S_k 所诱导的划分。 S_k 本身并未被列为 S_{k+1} 的一个分量，但可以由后者恢复出来：先取出 S_0 ，再逐层把较高深度签名投影到较低深度。因此这一列划分彼此嵌套。历史先映射为契约，再经有限编码器 e 得到存储表示 Z ；两个阶段都可能丢失信息：

$$h \xrightarrow{\text{契约深度 } k} S_k(h) \xrightarrow{\text{有限编码器 } e} Z. \quad (11)$$

若 $S_k(h) = S_k(h')$ ，则区别在编码之前就被删除，扩大码本 (codebook) 无法重建它。这称为契约混同 (contract aliasing)。若 $S_k(h) \neq S_k(h')$ ，但 $e(S_k(h)) = e(S_k(h'))$ ，则有限表示造成了容量混同 (capacity aliasing)。这两个阶段的损失一般不能相互补偿。

4.1 当前核心保护下的深度宽度

假设当前任务划分 $\mathcal{P}_0$ 被无损存储。对于一个受保护的划分块 $B \in \mathcal{P}_0$ ，令 $w_k(B)$ 为其中深度为 k 的递归子类数，并定义

$$w_k = \max_{B \in \mathcal{P}_0} w_k(B). \quad (12)$$

以 w_∞ 记各深度宽度的上确界。由于递归细化只会分裂划分块，

$$1 = w_0 \leq w_1 \leq \dots \leq w_\infty. \quad (13)$$

定理 4.1 (深度—容量阈值). 一个精确保留当前核心、残差字母表 (residual alphabet) 大小为 K 的表示，能够识别每个深度为 k 的递归类，当且仅当

$$K \geq w_k. \quad (14)$$

此外，对任意整数 $L \geq 0$ ，存在信号延迟为零的验证系统及受保护的初始划分块 B ，使得

$$w_0(B) = w_1(B) = \dots = w_L(B) = 1, \quad w_{L+1}(B) = 2, \quad (15)$$

其中新出现的分离对应相冲突的动作，且具有式 (2)、(3) 中固定的动作间隔。

阈值来自对每个受保护划分块内子类的计数。延后发生的跃变是关于初始块 B 的局部陈述，而不是关于全系统最大值 w_k 的陈述。事实上，若全局 $w_1 = 1$ ，就有 $\mathcal{P}_1 = \mathcal{P}_0$ ，从而不会再发生后续细化。在阶梯构造中，更接近终端验证的划分块比初始块更早分裂。

4.2 决策关键宽度

完整的递归可区分性可能强于某个下游任务的实际需要。假设每个深度为 k 的子类都有唯一最优动作，且该动作在子类内处处相同。定义 $\chi_k(B)$ 为受保护块 B 内这类不同动作的数量，并令

$$\chi_k = \max_B \chi_k(B). \quad (16)$$

命题 4.2 (决策关键宽度 (decision-critical width)). 在上述各子类内唯一最优动作恒定的假设下， K 符号残差表示能够在每个深度为 k 的子类上保留下游最优动作，当且仅当 $K \geq \chi_k$ ，并且

$$\chi_k \leq w_k. \quad (17)$$

只要 $\chi_k < w_k$ ，就可能出现只保住最优动作、却重建不出完整递归契约的情形。以下例子以 w 记递归类别数、 χ 记不同最优动作数。持久型 / 易失型见证例不是这种情形，它展示的是两种宽度相等： $w = \chi = 2$ ，两个递归类别要求相反动作，因此保住正确动作与区分这两类所需容量相同。例如，考虑三个递归类别，其中两类都应等待、第三类应提交。于是 $w = 3$ 、 $\chi = 2$ ，用两个残差符号就能保住动作，却无法区分全部三个类别。这个例子说明 $\chi_k < w_k$ 的情形可以实现，并不对应某个具体协议。

若某个深度为 k 的子类内部包含需要相反最优动作的历史，上述假设便不成立。此时，无论 K 多大，仅对 S_k 编码都无法恢复已被合并的区别。

4.3 精确的合并遗憾

设一个受保护块内的递归子类 C_i 具有先验概率质量 p_i 。令 $Q_i(a) = Q^*(h, a)$ ，其中 $h \in C_i$ ，并假设该最优动作价值在各子类内恒定；编码只限制控制器可利用的信息，不改变所声明的受控过程。残差编码器 e 将它们分组为 $G_z = \{i : e(i) = z\}$ 。

这里的 $Q_i(a)$ 需要满足共同后续策略的可实现性条件。给定编码 z 与动作 a 后，控制器无法识别组内的真实类别 i ，因此需要由同一个后续策略实现组内各类的动作价值：

$$\forall z, a \exists \rho_{z,a} \forall i \in G_z : \quad \mathbb{E}_i[U(a, \rho_{z,a})] = Q_i(a), \quad (18)$$

其中 $U(a, \rho_{z,a})$ 是先执行动作 a 、再采用策略 $\rho_{z,a}$ 所得的效用， $\mathbb{E}_i$ 按类别 i 的概率律取期望。 $\rho_{z,a}$ 只能使用保留的编码 z 和此后允许获得的观测，不能访问已被编码省略的类别信息。量词顺序要求先选定共同策略，再要求它对组内每个类别都成立。满足这一条件时，只观察残差编码的最优控制器所承受的遗憾为

$$R(e) = \sum_z \left[\sum_{i \in G_z} p_i V_i - \max_a \sum_{i \in G_z} p_i Q_i(a) \right], \quad R_K^* = \min_{|im e| \leq K} R(e), \quad V_i = \max_a Q_i(a). \quad (19)$$

本节首先刻画一次记忆动作选择的遗憾。完整历史下的最优动作价值给出可实现价值的上界，而共同后续策略的可实现性保证组内各类都能达到相应上界。对给定编码器，若每个编码组内的每个初始动作均满足式 (18)，则式 (19) 给出相对于该动作价值基准的最优遗憾。

若基准取自完整历史下的多步最优动作价值，而共同可实现条件不成立，则式 (19) 仅为实际总遗憾的下界。原因在于，压缩除了限制初始动作选择时可用的信息，还可能限制后续策略的实现。这一下界判断依赖上述完整信息价值基准，不适用于任意给定的价值表。

本文的阶梯构造对所有被比较的编码器均满足式 (18)，因此在编码器集合上最小化式 (19)，同样得到实际最优遗憾。

条件不成立时的落差有具体例子。取两条等概率历史 h_D, h_F ，二者在初始时刻的最优动作都是“继续”，而不是价值为零的“停止”。选择“继续”后，还需根据类型选择一次终端动作。保留全部历史时，两类都能作对这次选择，因此

$$Q_D(\text{继续}) = Q_F(\text{继续}) = 1. \quad (20)$$

现在把两条历史压成同一个编码，此后又不产生任何新观测，那么第二次选择最多只能猜对一半，实际可得的期望价值降至 $1/2$ 。若直接把 1 代入式 (19)，得到的合并遗憾是零，与实际损失并不相符。保留初始最优动作，并不保证其后续价值可实现；式 (18) 要求同一后续策略在组内各类中都能实现相应价值。

对于等概率出现、被赋予同一个残差编码的一对持久型和易失型历史，

$$R_1^* = \min\{0.5 \times 0.21, 0.5 \times 0.14\} = 0.07. \quad (21)$$

对于随机或连续编码器，同一思路给出加权的分布重叠公式，因此结果不依赖连续嵌入的精确相等（见附录 B.4）。

对于定理 4.1 中等概率的初始历史对，这两类损失给出一个简单的联合规律。对具有 K 个残差符号的 S_d 编码器，若 $d \leq L$ 或 $K = 1$ ，最小遗憾为 0.07；若 $d \geq L + 1$ 且 $K \geq 2$ ，最小遗憾为零。分离之前输入相同；分离之后，两个编码即可区分最优动作。附录 G 包含该规律的精确有限状态检查。

保留已有证据的变体还赋予合并损失另一种解释。对于命题 2.2 中取成本 c^* 的等概率历史对, 采用同一动作所产生的最小遗憾为

$$R_1^* = \frac{\Delta_D - \Delta_F}{4} = \frac{R_F - R_D}{4}. \quad (22)$$

这里的损失完全由未来信息价值的差异决定; 两条历史仍保留相同的已有事实证据。

4.4 面向未知访问成本的契约

对于一个已知成本, 记忆只需保留相应的最优动作。若契约需要支持一族未来访问成本, 要求就更强。固定命题 2.2 中的事实记录和两个信息价值, 令 $g = \Delta_D - \Delta_F > 0$ 。以 M 表示生命周期, H 表示可用历史, $\eta = P(M = D | H)$, Z 表示在后续成本揭示前由 H 生成的记忆, 取值空间记为 $\mathcal{Z}$ 。其后验概率为 $\Pi = P(M = D | Z)$ 。假设存在一个共同的最终读出器, 无需事先知道 M , 即可达到以生命周期为条件的贝叶斯风险; 双目标构造满足这一条件。

决策分为三个阶段: 编码器先形成 Z ; 随后给定外生成本 c ; 决策器再依据 (Z, c) 选择控制动作。本节比较同一表示支持整个成本族的能力。若通路在成本揭示前已被不可逆地分配, 决策器就无法根据成本重新选择; 这种情形不属于本节的模型。

由 H 形成表示 Z 是编码阶段, 占用名额、关闭通路或释放快照则属于后续控制动作。本节分析编码阶段需要保留的后验信息。

以下用 $R_Z(c)$ 表示基于 Z 的最优决策相对于已知生命周期的理想决策器 (oracle) 所产生的期望遗憾。这里说的“保持最优风险”是指基于 Z 与基于完整历史 H 的最优决策具有相同风险, 等价于对每个成本都有 $R_Z(c) = R_H(c)$, 并不要求二者为零; 完整历史本身也可能无法完全识别生命周期, 这部分由原始历史信息不足造成的不确定性不属于压缩新增的损失。

模型还假设: 给定生命周期与固定事实记录后, 其余历史线索不再提供关于目标及未来验证结果的信息。因此, 历史只改变生命周期后验, 两种生命周期对应的风险表在所考虑的历史间保持固定。结合共同读出器条件, 风险可按后验线性混合, 得到下述阈值判据。审计情形见附录 B.4。

定义成本平均量

$$\mathcal{A}(Z) = \frac{1}{g} \int_{\Delta_F}^{\Delta_D} R_Z(c) dc. \quad (23)$$

这里对声明的决策族取平均, 并不假设实际部署中的访问成本均匀分布。对成本敏感的贝叶斯风险积分, 是经典二元统计试验分析工具^[18]。在这里, 它给出保留已有证据模型中的精确压缩目标。

后验概率的作用可由归一化成本阈值说明。令 $t = (c - \Delta_F)/g \in [0, 1]$ 。给定 H 时, 生命周期后验为 η 。由命题 2.2 的条件风险, 提交风险减去等待风险所得的有符号差值为

$$L(\text{提交} | H) - L(\text{等待} | H) = \Delta_F + g\eta - c = g(\eta - t). \quad (24)$$

因此, 等待严格最优当且仅当 $\eta > t$; 当 $\eta = t$ 时, 等待与提交并列最优。成本固定时, 阈值 t 也固定, 最优动作只取决于 η 位于阈值的哪一侧, 因此保留一个动作标签即可, 无须保存 η 的具体数值。成本遍历整个区间时, 阈值 t 相应遍历 $[0, 1]$ 。任意两个不同的后验取值之间都存在一个阈值, 使二者对应相反的最优动作。因此, 要在整个成本族上保持最优风险, 就不能合并不同取值的 η 。

平方失真由这族损失的积分得到。给定记忆后验 $\Pi = p$, 在归一化阈值 t 处, 相对于已知生命周期的理想决策器, 条件遗憾为 $g \min\{p(1-t), (1-p)t\}$; 两条直线在 $t = p$ 处相交。对阈值积分, 后验恰为 p 时这一段成本族的平均损失是

$$\int_0^1 g \min\{p(1-t), (1-p)t\} dt = \frac{g}{2} p(1-p). \quad (25)$$

再对 (H, Z) 取期望，并用条件方差分解

$$\mathbb{E}[\Pi(1 - \Pi)] - \mathbb{E}[\eta(1 - \eta)] = \mathbb{E}[(\eta - \Pi)^2], \quad (26)$$

即得式 (27) 中压缩造成的额外损失。保留 η 足以支持阈值判据，因为该判据只依赖 η ；阈值遍历整个区间时，不同的 η 则必须可区分。完整证明见附录 B.5；附录 B.4 将同一个归一化成本阈值记为 τ ，用于分析审计情形。

定理 4.3 (成本平均遗憾与后验不确定性成正比). 在上述假设下，

$$\boxed{\mathcal{A}(Z) = \frac{g}{2} \mathbb{E}[\Pi(1 - \Pi)]}, \quad \boxed{\mathcal{A}(Z) - \mathcal{A}(H) = \frac{g}{2} \mathbb{E}[(\eta - \Pi)^2]}. \quad (27)$$

因此，压缩记忆能够在该区间内每一个成本处保持最优风险，当且仅当它几乎处处保留 η 。若 η 在 $[0, 1]$ 上均匀分布，最优 K 符号记忆的额外成本平均遗憾精确为

$$\inf_{|Z| \leq K} \{\mathcal{A}(Z) - \mathcal{A}(H)\} = \frac{g}{24K^2}. \quad (28)$$

将后验概率划分为等宽区间，对上述平均准则和最坏成本准则均为最优：

$$\inf_{|Z| \leq K} \sup_{c \in [\Delta_F, \Delta_D]} \{R_Z(c) - R_H(c)\} = \frac{g}{8K^2}. \quad (29)$$

允许随机化编码器也无法改进这两个最优值。

这些恒等式区分了给定原始历史后仍存在的不确定性与压缩新增的损失，也区分了三种信息保留要求：支持一个固定成本时，保留一个最优动作标签即可；支持整个成本族时，需要保留后验概率；若要重建完整的受控过程，则可采用递归稳定契约，即继续按后续转移细化也不再分裂类别的契约。第 5 节将给出其正式定义及充分性条件。式 (29) 的保证覆盖每一个成本，两个准则的证明见附录 B.5。这些精确的 K^{-2} 容量律限于所指定的均匀后验模型。

推论 4.4 (契约损失与容量损失的分离). 在定理 4.3 的条件下，令保留的契约为 $U = S_k(H)$ ，且 Z 仅由 U 生成，允许随机化。令 $\eta_k = P(M = D | U)$ 、 $\Pi = P(M = D | Z)$ ，则

$$\mathcal{A}(Z) - \mathcal{A}(H) = \frac{g}{2} \left\{ \underbrace{\mathbb{E}[(\eta - \eta_k)^2]}_{\text{契约损失}} + \underbrace{\mathbb{E}[(\eta_k - \Pi)^2]}_{\text{容量损失}} \right\}. \quad (30)$$

证明. 马尔可夫链 $M - H - U - Z$ 给出 $\eta_k = \mathbb{E}[\eta | U]$ 和 $\Pi = \mathbb{E}[\eta_k | Z]$ 。给定 (U, Z) 时，残差 $\eta - \eta_k$ 的条件均值为零。因此展开 $(\eta - \Pi)^2$ 后，式 (27) 中的交叉项消失。□

式 (30) 量化了式 (11) 中两个阶段的损失。增大码本可以减小第二项，却无法恢复契约已经删除的后验区别。反过来，细化嵌套契约可以减小第一项，但本身不能保证固定大小的编码器保留新增信息。对于这一成本族，保留生命周期后验概率已经充分，未必需要重建每一个递归类。

这些界讨论的是写入之后仍可获得的信息。保留可恢复的证据源，可以避免模型中的损失；否则，存储状态必须保留下游决策所需的生命周期区别。

5 充分契约：精确闭包与决策证书

上述结果在不要求重建每个未来统计试验的情况下，刻画了与决策相关的损失。如果契约必须保留完整的、已声明的受控过程，递归稳定性则给出一个充分的终点。我们先用记忆契约的符号表述经典有限状态划分细化结果^[19]，再给出近似停止证书。

5.1 精确递归闭包

设 S 、 $\mathcal{A}$ 和 O 分别为有限状态集、动作集和观测集，受控马尔可夫核为 $\kappa(o, s' | s, a)$ 。基础标签为 $\ell: S \rightarrow \mathcal{L}$ ，其中 $\mathcal{L}$ 是标签集， $\mathcal{P}_0$ 是 ℓ 诱导的划分。对其细化 $\mathcal{P}$ ，以 $[s]_{\mathcal{P}}$ 记包含状态 s 的划分块，并定义

$$\sigma_{\mathcal{P}}(s) = \left([s]_{\mathcal{P}}, \ell(s), \left(\sum_{s' \in B} \kappa(o, s' | s, a) \right)_{a \in \mathcal{A}, o \in O, B \in \mathcal{P}} \right). \quad (31)$$

令 $F(\mathcal{P})$ 为按 $\sigma_{\mathcal{P}}$ 相等关系诱导的划分。若一个划分保留基础标签，且同一划分块中的状态在每个动作下，向每个“观测—划分块”对转移的概率质量均相同，则称其递归稳定。

定理 5.1 (递归闭包刻画). 序列 $\mathcal{P}_{j+1} = F(\mathcal{P}_j)$ 到达 $\mathcal{P}_0$ 唯一的最粗递归稳定细化 $\mathcal{P}^*$ 。 $\mathcal{P}_0$ 的每一个递归稳定细化也都细于 $\mathcal{P}^*$ 。若 $|S| = n$ 、 $|\mathcal{P}_0| = b_0$ ，严格细化的轮数至多为 $n - b_0$ 。

加入 $[s]_{\mathcal{P}}$ 保证细化只分裂而不合并。以 S'_j 记后继状态的签名。当 $S_0 = \ell$ 时， $\mathcal{P}_j$ 恰好是 S_j 诱导的划分： (O, S'_j) 的概率律相同，等价于向每个“观测—划分块”对转移的概率质量相同。因此，定理 3.1 讨论的是同一序列的有限前缀，而这里描述其终点。不存在适用于所有有限系统的统一深度，但每个固定有限系统最终都会稳定。证明见附录 A.3。

为得到决策保证，基础标签必须保留动作可行性和奖励；在有限时域问题中，还必须保留剩余时域与终端效用。此时，反向归纳或折扣贝尔曼不动点保证稳定划分块上的动作价值恒定。完整模型的状态需要包含影响未来奖励或验证的持久变量。压缩表示则需要保留这些变量对已声明决策造成的区别；若要求维护完整受控过程，还需保留它们对后续奖励与验证概率律造成的区别，但不必逐项存储原始变量。这里的粗细关系相对于该受控模型和固定的观测标签而言；已知的可逆重标记可以先作规范化处理。

给定有限受控模型，按下述步骤计算 $\mathcal{P}_0$ 的最粗递归稳定细化。

- 1) 从基础划分 $\mathcal{P}_0$ 出发。
- 2) 在每个当前划分块内部，比较所有动作下向各个“观测—划分块”对转移的概率质量。
- 3) 按这些概率质量的差异把划分块拆开；块数只增不减，因此已分开的类别不会被重新合并。
- 4) 若整个划分不再变化就停止，否则回到第 2 步。

停止条件要求全局划分稳定；某个初始块不再分裂，并不足以停止迭代。第 4 节已经说明，局部宽度可以连续多轮保持为一，而更靠近终端验证的块早已开始分裂。任意深度反例与闭包算法分别描述了同一细化序列的有限前缀与稳定终点。有限停止保证适用于上述已知的有限受控模型。环境状态有限，并不自动意味着历史的充分信息状态也有限——在一般的部分可观测环境下，信念状态或充分信息状态的集合可能是无限的，本节结论不能直接套用。

若 $\mathcal{P}_0$ 已经稳定，就不需要细化。因此，仅有长时域或隐藏状态不足以造成上述失效；当前契约块必须隐藏后续概率律的差异。

5.2 近似闭包也能认证决策

当预测只是近似时，精确细化可能不切实际。考虑有限折扣马尔可夫决策过程 (Markov decision process, MDP)，其奖励 $r(s, a)$ 有界，转移概率为 $P(\cdot | s, a)$ ，折扣因子满足 $0 < \gamma < 1$ 。取满足 $\delta_a(s, t) \geq |r(s, a) - r(t, a)|$ 的有界伪度量 (pseudometric) δ_a 。沿用 Ferns et al.^[14] 的互模拟度量构造，对状态上的有界伪度量 d 定义算子 $\mathcal{B}$ ：

$$(\mathcal{B}d)(s, t) = \max_{a \in \mathcal{A}} [\delta_a(s, t) + \gamma W_d(P(\cdot | s, a), P(\cdot | t, a))], \quad (32)$$

其中 W_d 为以 d 为运输代价的有限状态 1-Wasserstein 距离 (1-Wasserstein distance)。

命题 5.2 (折扣情形的近似证书). 在配备上确界范数 (supremum norm) 的有界伪度量空间上, $\mathcal{B}$ 是 γ -压缩映射 (contraction), 并有唯一不动点 d^* . 从 $d^{(0)} = 0$ 出发, 令 $d^{(k+1)} = \mathcal{B}d^{(k)}$, 则

$$0 \leq d^*(s, t) - d^{(k)}(s, t) \leq \frac{\gamma^k \Delta_0}{1 - \gamma}, \quad \Delta_0 = \sup_{s, t, a} \delta_a(s, t). \quad (33)$$

此外, $|Q^*(s, a) - Q^*(t, a)| \leq d^*(s, t)$. 若 s 有唯一最优动作, 且最优与次优动作的价值间隔为 $m(s)$, 则

$$m(s) > 2 \left[d^{(k)}(s, t) + \frac{\gamma^k \Delta_0}{1 - \gamma} \right] \implies \pi^*(s) = \pi^*(t). \quad (34)$$

该动作价值界还给出一个反向分离结论: 若 s, t 的唯一最优动作 a_s, a_t 不同, 间隔分别为 $m(s), m(t) > 0$, 则

$$d^*(s, t) \geq \frac{m(s) + m(t)}{2}. \quad (35)$$

这说明, 若两个状态具有不同的唯一最优动作, 则在能够控制动作价值差异的递归距离下, 二者的距离存在正下界。证明见附录 A.4。应用于终端效用被推迟的阶梯构造时, 需按所选折扣约定重新计算动作间隔。

附录 A.4 给出了证书的证明。如果有限深度的上界已低于动作间隔的一半, 就无需计算完整不动点。该折扣结果与阶梯构造中的固定间隔不矛盾: 阶梯使用未折扣终端效用; 若折扣因子固定且奖励一致有界, 其贡献就会随深度增加而减小。

6 学习决策所需的区分信息

理论讨论契约能够支持的最优决策; 神经编码器也可能额外保留其目标函数并未要求的信息。实验在保持当前信息不变的条件下比较生命周期敏感监督与浅层监督, 再改变浅层目标数和训练预算。最后一个实验用不确定的审计证据替代显式生命周期线索。

6.1 固定一步契约

首先固定当前与一步信息, 使策略改善只能来自学习到的残差。在一个受控两步验证过程中, 定义

$$S_1(h) = [p(h), J_{1,w}(h), J_{1,c}(h)]. \quad (36)$$

其中 $p(h)$ 是待写入内容在当前条件下正确的概率, $J_{1,w}, J_{1,c}$ 是生成器在等待与提交下的一步联合概率向量。该向量被精确且完全相同地复制到两个实验组中, 不含可训练参数。各组还使用相同的两层因果 Transformer (causal Transformer) 架构学习四维残差 $R_\theta(h)$, 其中 θ 为编码器参数, 模型隐藏维度为 $d_{\text{model}} = 48$, 注意力头数为四, 前馈网络 (feed-forward network, FFN) 宽度为 96, 训练 400 次更新。两组残差编码器使用相同的当前 / 一步训练目标, 只有递归组额外使用两步试验目标。残差训练结束后, 另行训练的非线性 Q 解码器 (decoder) 读取 $[S_1, R_\theta]$ 。

训练集、验证集和测试集由同一生成器独立采样得到。报告所用的测试集由 300 个成对基础样本构成, 每个实验组各用五个优化器随机种子; 模型选择一律不看测试结果。将对内最优动作不同的历史对称为决策分离对。表 1 列出策略性能, 并给出直接预测头的 KL 散度 (Kullback-Leibler divergence) 诊断。

表 1 受保护 S_1 实验，使用五个优化器随机种子

指标	浅层残差	递归残差
受保护 S_1 的对内差异	0	0
两步策略遗憾	0.023324 ± 0	0.000345 ± 0.000135
决策分离样本的动作准确率	0.500 ± 0	0.996 ± 0.004
决策分离对内均正确的比例	0.000 ± 0	0.993 ± 0.008
直接预测头的两步 KL 散度	1.422 ± 0.168	0.0603 ± 0.0043

注：两组的当前 / 一步统计量完全一致。在所比较的实验设置下，浅层组的后续策略碰撞仍然存在，而递归残差训练使决策解码器能够区分这些历史。

策略改善发生时，受保护的当前 / 一步统计量没有改变。KL 一行是直接预测头的诊断：浅层组的两步预测头未经过训练，因此该项比较没有对两组的冻结表示另行训练相同的两步预测解码器，不能据此判断可恢复信息的多少。两组的策略遗憾均使用另行训练的解码器评估。在实际部署中，受保护统计量本身也需要估计或保留。

6.2 增加浅层目标或生命周期敏感目标

下一项测试将成对历史共享的浅层目标数取为 $k \in \{1, 2, 4, 8, 16, 32\}$ ，并在对照组之间匹配新增监督。每对长度为 20 个词元 (token) 的历史共享一个由生成器抽取、与生命周期无关的数值坐标，以及干扰比特和全部噪声词元，只有一个显式生命周期词元不同。该坐标分箱后作为输入词元，用于生成各个浅层目标。编号 $1, \dots, k$ 的每个浅层目标都是同一共享数值坐标的平滑函数，且由历史对共享，因此持久型与易失型历史在目标编号 k 以内具有相同目标。第一个新增目标等概率取 $\{0.95, 0.60\}$ 。在可观测干扰对照组中，该值由成对共享的干扰比特决定；在生命周期敏感目标组中，该值由生命周期决定：持久型取 0.95，易失型取 0.60。由此，对照组与该组在额外损失、目标边缘分布及方差、目标数量设置和二元线索可见性上均匹配，只有线索与未来证据的耦合不同。

这些是构造的监督目标，而不是由模拟转移核计算得到的完整递归概率律 S_j 。实验检验的是：共享目标数量增加时，揭示一个生命周期敏感目标会产生什么影响。此外，其连续瓶颈 (bottleneck) 未声明有限字母表，因而瓶颈维度不是定理 4.1 中的容量 K 。

本组实验中 k 是浅层目标的数量，即目标编号 $1, \dots, k$ 的最大值。递归深度则指定义 2.1 中签名 S_j 的层次，由附录 G 的有限状态阶梯检验。增大 k 只增加共享浅层目标的数量，不改变系统的递归深度。

下游动作价值固定为

$$Q(D, \text{等待}) = 0.95, \quad Q(F, \text{等待}) = 0.60, \quad Q(\cdot, \text{提交}) = 0.74.$$

对于任意对内不变表示，最优共同动作为等待，从而平均碰撞遗憾精确为

$$R_{\text{collision}} = \frac{0 + 0.14}{2} = 0.07. \quad (37)$$

单层因果 Transformer ($d_{\text{model}} = 24$ 、四个注意力头、FFN 宽度 48) 将历史映射到三维瓶颈。训练使用 900 个成对基础样本、220 次更新，每次更新均匀抽取一个浅层目标编号。在训练瓶颈表示上拟合的岭回归解码器 (ridge decoder)，对留出样本 (held-out samples) 预测两个动作价值；线性生命周期探针 (lifecycle probe) 作为辅助指标。图 3 概括了预测目标数量实验的留出集结果。

测试使用 300 个留出的成对基础样本，每个实验组、每个 k 设置使用五个优化器随机种子，共 $6 \times 3 \times 5 = 90$ 次运行。独立验证集用于模型和分析检查，测试样本不用于调参。

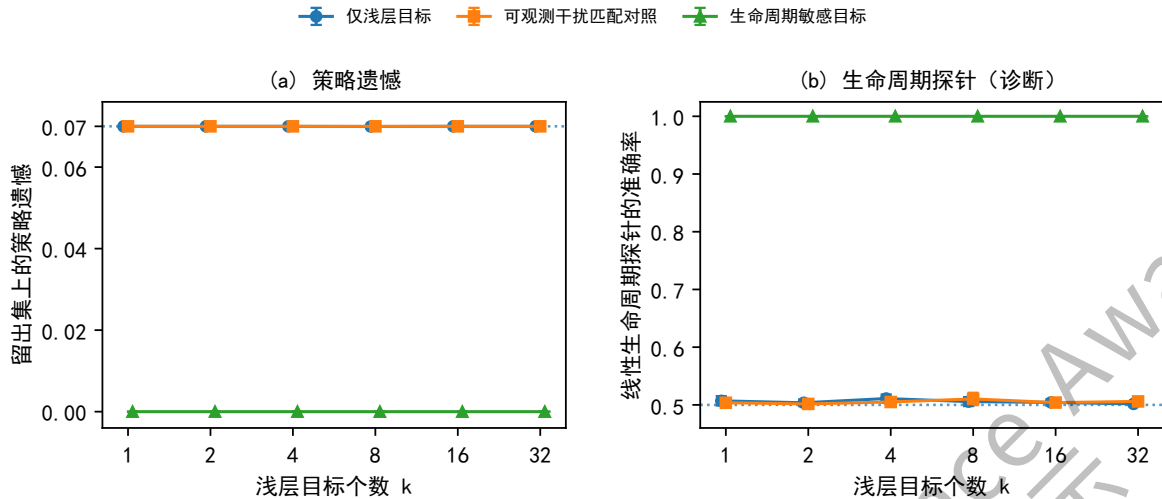


图3 预测目标阶梯

注：(a) 直至 $k = 32$ ，浅层目标与可观测干扰匹配目标给出的遗憾都贴近对内不变参考值 $R = 0.07$ ；生命周期敏感目标组则在每一种目标数量设置上都把遗憾压到零。(b) 辅助线性生命周期探针浅层与对照表示的准确率与随机猜测相当，对生命周期敏感目标组的表示可达 1.00。误差线表示五个优化器随机种子对应测试指标的一个标准差。

留出集结果。 在全部六种目标数量设置上，30 次生命周期敏感目标组运行的策略遗憾均为零，动作准确率均为 1.00。浅层组和干扰匹配对照组的遗憾接近对内不变参考值，五种子均值与 0.07 的最大偏差小于 5×10^{-5} 。

辅助线性生命周期探针在浅层组的平均准确率为 0.502–0.511，在匹配对照组为 0.502–0.510，生命周期敏感目标组每次运行均为 1.00。生命周期标签存在于输入中，因此浅层训练原则上可以保留它。但在所测试的目标与瓶颈下，只有生命周期敏感监督使其能被决策解码器可靠使用。

6.3 对架构和训练预算的敏感性

另两项受控两步实验分别使用门控循环单元（gated recurrent unit, GRU）和两层 Transformer。GRU 与受保护契约实验使用同一数据生成器。两项实验的设置均与预测阶梯不同，用于补充检验监督效果，不能直接比较架构优劣。GRU 补齐了实验组合，形成监督方式与训练预算的二乘二对照。图 4 既给出汇总后的遗憾，也给出逐个随机种子的结果。

GRU。 更新 320 次时，浅层与递归的遗憾分别为 0.05659 ± 0.00071 与 0.03258 ± 0.02659 。五个配对种子中，四个在递归训练下的遗憾降低，一个略有增加。在这五个种子中，有两个的遗憾仍接近浅层组。更新 700 次时，两者分别为 0.04828 ± 0.00730 与 0.00284 ± 0.00076 ，五个配对种子全部改善。两组更新次数相同，700 次时的收益仍然存在；320 次时收益不稳定。

小型因果 Transformer。 带四维瓶颈的两层因果 Transformer 上，浅层与递归的两步 KL 散度分别为 0.0592 ± 0.0024 与 0.0175 ± 0.0122 ；两步遗憾则从 0.0508 ± 0.0032 降到 0.0000 ± 0.0000 。浅层 Transformer 的生命周期探针准确率为 0.574 ± 0.038 ，表明在该架构下仍能从表示中解码出部分生命周期信息；这一观察并不保证递归充分性。

在本次较高训练预算的匹配对照中，所测试模型从相应监督中获得了收益。GRU 的监督收益随预算变化，浅层 Transformer 也可能保留部分生命周期信息，因此这些结果不能支持“遗忘必然发生且与架构无关”的一般结论。

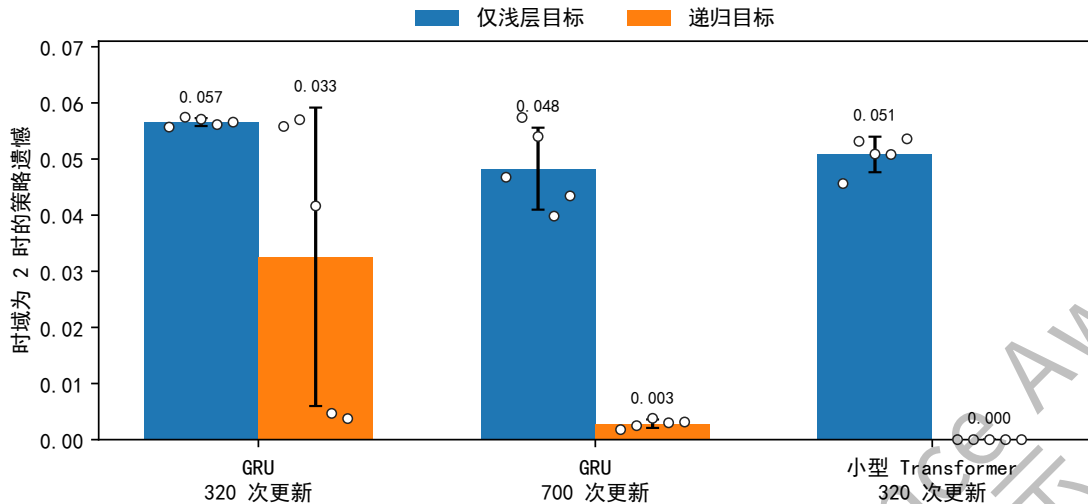


图 4 各组内部更新次数配平后的两步策略遗憾

注：GRU 给出 320 次与 700 次更新两种预算的结果，其中 320 次时递归训练的种子间波动明显偏大。小型 Transformer 为 320 次更新。柱高为均值，误差线为一个标准差，散点画出全部五个优化器随机种子；位置重合的点在水平方向相互错开。

这些学习实验中的遗憾衡量的是指定的拟合读出器，而不是连续嵌入上所有可能读出器中的最优者。接近随机猜测的探针，或接近成对碰撞参考值的遗憾，都不能证明嵌入已经不可逆地丢失生命周期信息。

6.4 没有显式标记时推断生命周期

前述神经网络输入通过词元显式提供生命周期。为检验能否转而从不确定证据恢复相关决策，我们在推论 3.2 的双目标构造中令生命周期不可观测。令 π 为观察历史审计后属于持久型的后验概率。预留主通路选择辅助通路的条件风险分别为 $7/16 - 3\pi/16$ 和 $3/8$ 。因此决策阈值为 $\pi > 1/3$ ，而不是分类阈值 $1/2$ 。

样本为每个历史提供了八个审计比特：给定生命周期后，这些比特相互独立；持久型与易失型条件下的成功概率分别为 $3/4$ 与 $1/4$ 。除了反映生命周期之外，它们不再包含目标方面的信息。模型看到的只是成功次数，并不接收生命周期词元。逻辑回归 (logistic regression) 在 4096 个样本上完成拟合，标签取自事后观测到的最终可用性；评估则在类别均衡、与之互不重叠的另外 4096 个样本上进行。干扰对照在训练集和测试集中分别独立打乱审计记录的行序，保留经验边缘分布，切断记录与可用性的关联。审计历史由受控生成器产生。

在 n 次审计中有 s 次成功，且两种生命周期先验相等时，已知似然给出的后验概率为 $\pi(s) = [1 + 3^{n-2s}]^{-1}$ 。当 $n = 8$ 时，拟合模型与该贝叶斯规则在每个测试样本上选择相同动作。相对于已知生命周期的理想决策器，平均遗憾为 0.00546；无审计策略和打乱审计模型的遗憾均为 0.03125。给定训练集，基于 Hoeffding 不等式 (Hoeffding's inequality)^[20] 得到的拟合策略相对于始终等待策略的遗憾改善量的 95% 区间为 $[0.02181, 0.02977]$ 。精确的总体贝叶斯遗憾为 0.00526：剩余损失与两种审计概率律的重叠相符，并不意味着必须提供显式生命周期标记。附录 B.4 推导该风险，并界定后验估计误差造成的额外损失。该实验假设审计似然稳定且具有信息量，没有检验如何在分布偏移下发现这些似然。

7 文件、网络与数据库中的证据访问

前述学习实验使用受控生成器。以下通过实际文件和网络机制检验证据可用性：可移植操作系统接口 (Portable Operating System Interface, POSIX) 下的文件生命周期，以及采用确定性读出器、

基于传输控制协议 (Transmission Control Protocol, TCP) 和 HTTP 的权能凭证服务。保留预览和通路预算的两项扩展用于考察决策权衡, 最后用 SQLite 观察快照保留与存储成本之间的关系。

前两个协议明确规定保留规则: 等待只保留活动引用, 提交保留校准后的快照。选择等待后, 最终读出器无法获得之前的预览; 访问失败时返回固定预测 $1/2$ 。若读出器还保留预览, 则易失型历史中的等待可以退回快照预测, 与提交并列。因此, 第 7.3 节加入第三个协议, 保留预览并检验命题 2.2 中的成本阈值。

7.1 放弃活动文件引用

专用证据源进程创建权威临时文件。持久型来源记录额外保留一个已打开的文件描述符; 易失型来源记录只保留路径名。测量时间线为

当前读取 $\rightarrow$ 记忆动作 $\rightarrow O_1 \rightarrow$ 最终写入并解除链接 $\rightarrow O_2$ 。

对每个给定动作, 实验均显式测量下一次观测 O_1 , 并要求持久型与易失型历史的观测结果相同。在等待下, 两条仍有效的访问路径均返回相同预览 y_0 ; 在提交下, 两者均返回相同的序列化快照, 并丢弃所有活动描述符。之后, 证据源才写入最终真值 G , 调用 `fsync`, 并解除路径名链接。POSIX 规定, 只要仍有已打开的引用, 解除最后一个路径名的链接并不会销毁文件内容^[21]。

此处不使用学习型读出器。令 $G \sim \text{Bernoulli}(1/2)$ 、预览翻转概率 $q = 0.25$, 并使用二元 Brier 损失, 则

$$L_{D,W} = 0, \quad L_{D,C} = L_{F,C} = q(1-q) = 0.1875, \quad L_{F,W} = 0.25. \quad (38)$$

对任意 $0 < q < 1$ 且 $q \neq 1/2$, 均有 $0 < q(1-q) < 1/4$, 故动作反转并不是针对 0.25 调出来的结果。

实验设置独立验证集, 并留出 1024 对样本作为测试集。时间线在每个记忆动作之后、证据源最终更新之前显式测量各动作下的 O_1 , 从而分别观测浅层结果相同与后续证据不同这两个事件。POSIX 留出集指标见表 2。

表 2 POSIX 留出测试

指标	持久型描述符	易失型路径名
两类历史的当前观测及各动作下 O_1 相同	1024/1024	1024/1024
解除链接后成功读取最终权威值	1024/1024	0/1024
等待下的平均 Brier 损失	0	0.25
最优动作	等待	提交
观测动作间隔	0.18799	0.06201

注: 两类历史的当前观测及每个给定动作下的一步观测 O_1 均相同; 差异体现在后续证据是否仍可访问。

7.2 HTTP 权能凭证的访问损失

为检验该机制是否依赖已打开文件描述符的特定语义, 我们使用独立证据源进程开展第二项系统实验, 进程仅通过 TCP/HTTP 通信。服务拥有权威记录, 并提供两个引用: 公开定位符和使用基于 SHA-256 的密钥哈希消息认证码 (keyed-hash message authentication code using SHA-256, HMAC-SHA256) 签名的持久型权能凭证 (capability)。最终更新之前, 两者返回相同预览。等待时, 持久型来源记录保留签名凭证, 易失型来源记录只保留公开定位符; 提交时, 两者均序列化相同的预览 / 校准转移模型, 并放弃活动证据通路。显式时间线仍依次测量当前读取、动作条件 O_1 、证据源最终更新和 O_2 。

最终更新改变权威值, 并使公开定位符过期, 返回 HTTP 状态码 410; 签名凭证则仍有权读取最终值。因此, 同一个确定性 Brier 读出器给出相同的总体排序 $0 = L_{D,W} < q(1-q) < 1/4 = L_{F,W}$ 。在含 1024 对样本的留出测试前使用独立验证集。TCP/HTTP 留出集指标见表 3。

表 3 TCP/HTTP 留出测试

指标	持久型权能凭证	仅公开定位符
两类历史的当前观测及各动作下 O_1 相同	1024/1024	1024/1024
过期后成功读取最终权威值	1024/1024	0/1024 (HTTP 410)
等待下的平均 Brier 损失	0	0.25
最优动作	等待	提交
观测动作间隔	0.18652	0.06348

注：签名凭证与会过期的公开定位符在当前读取和 O_1 时均返回相同结果；服务器完成最终更新后，二者的访问结果才出现差异。

两个协议里，两类历史在当前状态与动作条件的一步观测 O_1 上完全一致，差别只在后续还能否访问权威证据源。HTTP 测试通过环回 TCP 在独立进程间运行，说明该效应不依赖 POSIX 解除链接语义。这些本地测试确定了所述协议的行为，而不是该行为在已部署智能体中的出现频率。

7.3 保留预览：未来信息的作用

这一扩展沿用同一套 HTTP 权能凭证机制，改动的是记忆与决策规则。两类等待历史此时都把完整预览 y_0 留下来。当易失型定位符过期，读出器改用校准预测 $P(G = y_0) = 0.75$ ，与提交所得完全相同。选择等待需要支付一笔事先声明的持有成本 c ；它是为预留未来访问而设的效用参数，而不是实测到的网络费用或运行时间，无论最终验证结果如何都要支付。

当 $q = 0.25$ 时，快照风险为 $R_0 = q(1 - q) = 0.1875$ 。持久型验证成功时 $R_D = 0$ ；易失型验证失败后预览仍完好，因此 $R_F = R_0$ 。由命题 2.2，整个区间 $0 < c < 0.1875$ 内都应有严格相反的最优动作。主要成本设置 $c = 0.06$ 下，预测间隔分别为 0.1275 和 0.06。

新增的 512 对样本全部保持两类历史的当前观测及各动作下 O_1 相同。持久型凭证在每个样本上都返回最终真值，易失型定位符在每个样本上都返回 HTTP 410；每个易失型预测都与其保留的快照预测精确一致。平均快照 Brier 损失为 0.1875。表 4 报告相应总损失。成本网格只是对相同观测重新评分，并不构成相互独立的数据集。

表 4 保留预览的 HTTP 测试，共 512 对样本

持有成本 c	持久型等待	易失型等待	提交（两类）
0	0	0.1875	0.1875
0.03	0.03	0.2175	0.1875
0.06	0.06	0.2475	0.1875
0.09	0.09	0.2775	0.1875
0.15	0.15	0.3375	0.1875
0.1875	0.1875	0.3750	0.1875
$c = 0.22$	0.22	0.4075	0.1875

注：表项为平均 Brier 损失加上事先声明的持有成本，数值越小越好。成本为零时，易失型历史下两个动作并列；成本超过持久型通路的信息价值时，两类生命周期中的最优动作均为提交。

持有成本为零时，易失型历史下的两个动作给出相同预测，因而并列最优。当 c 增至 0.22 时，两类生命周期下均以提交为最优动作，与阈值预测一致。主要设置下，持久型的动作间隔为 0.1275，跨样本对的描述性 95% 正态区间是 [0.1087, 0.1463]；易失型的动作间隔恒等于 c ，因为两个动作在每个样本上给出相同预测。这说明，即使保留预览，动作分离仍可发生。该扩展要求维持未来访问具有显式成本，其允许范围由理论预先给定。

7.4 不额外收费时的证据通路竞争

本实验检验不额外收取费用时是否仍会发生动作分离。证据源服务按推论 3.2 的双目标问题构建。每次写入仅保留一条活动通路：等待动作保留主通路；提交动作保留主目标预览，并将活动通路名额分配给辅助目标。服务器拒绝同时保留两条通路的请求，并撤销未选中通路的访问权。即使在写入前复制了有效令牌，也无法恢复已撤销的访问权。全部已有事实观测仍对读出器可见。

持久型主通路返回最终真值；易失型主通路过期。辅助通路返回独立目标的有噪声观测，准确率为 3/4。评分是两个 Brier 损失之和，不含服务费用或时间惩罚。在双名额对照中，两条通路都保持可用。对于易失型历史，等待此时可以同时使用旧主目标预览和辅助观测，因此应在每个样本上与提交并列。

表 5 比较精确总体风险与新增的 512 对样本测试。所有超出单名额预算的尝试均被拒绝，所有未选通路均拒绝原有令牌，且配对历史在两种生命周期下的当前观测及各动作下的 O_1 均相同。单名额动作间隔为 0.09766 和 0.06934。联合覆盖概率至少为 95% 的两个 Hoeffding 区间分别为 [0.05140, 0.14391] 和 [0.03663, 0.10204]，均严格为正。双名额对照中，512 对样本的易失型损失全部并列。附录 E.1 给出精确枚举与有限样本界。

表 5 保留预览且不额外收费时的预算受限验证

活动通路预算	动作与生命周期	总体风险	观测均值
一个名额	等待，持久型	0.2500	0.25000
一个名额	等待，易失型	0.4375	0.41699
一个名额	提交，任一类型	0.3750	0.34766
两个名额	两条通路，持久型	0.1875	0.18066
两个名额	两条通路，易失型	0.3750	0.34766

注：总风险为两个 Brier 损失之和。机会成本来自辅助通路自身的信息价值；通路约束一旦解除，易失型历史中的严格偏好随之消失。

活动通路的分配约束由此产生满足理论阈值关系的机会成本。该约束作用于写入时的活动访问权分配；即使已有事实证据 X 被保留，权衡仍然存在。双通路对照表明，同时保留两条通路可以消除这一特定权衡。

7.5 保留活动快照与完整复制的比较

上述访问预算由测试服务强制执行。为考察现有存储引擎本身产生的资源成本，我们使用预写日志（write-ahead logging, WAL）模式下的 SQLite。活动读事务保留其历史快照，但当独立写入进程改变数据库时，可能阻止检查点（checkpoint）完成^[22]。实验测量实际 WAL 文件和 SQLite 检查点响应，不强加通路配额，也不额外添加效用费用。

SQLite 实验考察不同保留方式的存储代价。POSIX 与 HTTP 协议检验后续能否继续访问证据源，SQLite 查询的则是写入前已经存在的旧载荷，因此完整离线备份可以通过复制保存所需证据。该实验不构成另一组持久型 / 易失型动作反转，也不能据此推断完整备份能够替代尚未产生的未来证据。

数据库共 64 行，每行包含一条不变的当前事实和 1024 字节的证据载荷（payload）。记忆动作有三种：保留读事务与当前事实；只保留事实；建立完整离线备份后释放事务。独立进程分 8、16 或 32 批写入，每批替换全部载荷，并尝试一次截断式检查点。检索阶段才选定一行，查询其写入前的旧载荷。结合两个预先选定的数据种子，共有 18 种配置。

表 6 报告保留产物的字节数加上检查点尝试后的 WAL 峰值字节数。所有配置都保留了当前事实。活动事务和离线备份在各组六个配置中均恢复了所查询的旧载荷，而最新证据源返回不同载荷。只要旧事务仍被保留，每次检查点均报告忙碌；释放旧事务后，WAL 可以截断为零字节。另两组始终允许检查点完成。

表 6 SQLite 快照实验

保留状态	可查询旧载荷	8 批 (字节)	32 批 (字节)
活动事务 + 事实	是	725 162	2 900 522
仅事实	否	10	10
离线备份 + 事实	是	98 314	98 314

注：两个种子给出的字节计数完全一致。资源计数为保留产物与检查点之后 WAL 峰值之和，其中各组均含 10 字节的事实记录；各组共用的源数据库不计在内。三种事先声明的工作负载都已保留在数据中，16 批的设置同样在内。

备份对照在不保留源 WAL 的情况下保存了证据。在这些工作负载中，其固定存储占用小于活动事务的占用，因而提供了保留活动引用之外的另一种选择。仅凭字节数不能确定效用参数 c ；这需要针对任务评价存储成本。协议细节见附录 F。

8 讨论

记忆即使保留了当前答案，也可能丢失选择未来证据通路所需的信息。任何固定深度的浅层契约都无法对所有有限系统排除这种失效。在成本族模型中，契约形成与有限编码造成的额外遗憾可分解为两个非负项（推论 4.4）。受保护 S_1 实验表明，在当前与一步信息保持不变时，改变残差监督仍可降低所测试读出器的后续策略遗憾。

最初的 GRU 对照采用了不同的训练更新次数。补齐匹配预算对照后，递归监督在 320 次更新时收益不稳定，在 700 次更新时五个种子均有改善（图 4）。这表明其收益受训练预算影响。

若整合后无法再访问权威证据源，存储记录就需保留下游任务所需的生命周期信息。来源字段不必逐个保留：决策关键宽度有可能明显小于完整的递归宽度。需要保留的是这些量对所声明决策造成的区别；如果还要求维护完整的受控过程，还需保留影响后续奖励与验证概率律的区别，但仍不必逐项保存产生这些区别的原始变量。

保留预览的扩展表明：把已有事实证据 X 留着，并不能抵消未来证据通路的价值。该信息价值决定了动作反转的成本区间。对于整个成本族，定理 4.3 说明精确的压缩损失与生命周期后验概率的平方失真成正比，比例系数为 $g/2$ 。活动通路实验通过通路竞争产生机会成本；SQLite 实验测的则是物理存储开销，并说明离线复制在什么条件下可以绕开这一开销。零成本、高成本、双名额与完整备份这几组对照，各自对应一种让权衡消失的途径。

闭包定理要求状态集有限；近似证书假设采用折扣回报且奖励有界。成本族恒等式还要求以生命周期为条件的信息价值固定，并存在共同且可实现的最终读出器。神经网络实验的输入显式包含生命周期线索，审计扩展则假设观测概率律稳定且含有信息量。这些结果尚不能说明已部署智能体能够可靠推断生命周期，连续瓶颈的维度也不等于有限码本约束。系统实验检验的是所指定机制的行为，未估计该机制在实际部署中的出现频率或效用成本。后续实证研究可进一步考察：当审计似然发生漂移时，如何比较测得的存储开销与未来信息价值。

9 结论

持久记忆写入可能同时压缩记录并改变证据通路。即使某种状态表示足以支持当前任务或未来验证的任意固定有限前缀，当写入改变后续证据访问时，该表示仍可能无法保证后续决策最优。

在成本族模型下，若要在每一个访问成本处都保持最优风险，记忆就必须留住生命周期后验概率。压缩造成的损失可以拆成契约损失与容量损失两部分。当完整的受控过程都需要保留时，递归闭包给出的是一个充分契约。学习实验表明，在一步统计量保持不变时，生命周期敏感监督仍能改善所测试读出器的决策性能；系统实验表明，证据访问约束能够按照理论预测产生或消除决策权衡。因此，确定写入内容时，需要同时考虑未来证据通路的可用性。

研究过程与致谢

研究过程

本研究起源于多目标跟踪中的标识变化问题。我们注意到，身份连续性与关联事实的更新时间尺度不同。持续传输有助于加快信息更新，却也可能使一次身份切换造成长期保留的错误记录。这一观察引出了本文的记忆问题。

理论研究从持久型 / 易失型见证例出发，分析未来查询未知时写入应保留的信息，继而研究任意深度分离、递归闭包与有限容量遗憾。学习实验比较浅层、干扰匹配及生命周期敏感监督；受保护 S_1 实验固定当前 / 一步信息，考察可学习递归残差对后续决策的影响。

POSIX 与 TCP/HTTP 实验以确定性读出器检验证据可用性的变化。后续实验保留旧预览，以通路竞争代替显式成本，并检验动作反转消失的条件。GRU 匹配预算对照同时报告高预算下的稳定收益和低预算下的不稳定结果。

学生与指导教师贡献

参赛学生贡献：本人负责本报告的研究、程序实现与写作。沿用的理论工具与本项目得到的结果分列如下。

本文的图、表和全部数值结果，除文中另有说明者外均由本研究自行得到；凡涉及外部方法、标准或已发表结论，均在首次使用时注明出处。

沿用部分。划分细化、互模拟度量、Wasserstein 距离，以及二元统计试验的风险积分与阈值判据均为已有工具，首次使用处已引用。有限状态递归闭包、最粗稳定细化与 $n - b_0$ 轮严格细化停止界沿用经典细化论证；本文将其用于记忆契约，并说明稳定终点与任意深度失效的关系。

本项目自行得到的部分。持久型 / 易失型见证例与保留已有证据的阈值变体；任意深度验证阶梯及其不删除预览、不额外收费的变体；浅层距离不存在正安全半径的近碰撞构造；深度—容量阈值与延迟跃升的局部性结论；决策关键宽度及其与完整递归宽度的区别；成本平均遗憾与最坏成本遗憾的两条精确容量律，以及契约损失与容量损失的分解；三套系统协议的设计与实现；以及上述结果的中文表述与报告撰写。

实现、实验与计算。受控学习实验中的残差编码器、动作价值解码器与生命周期探针，POSIX 文件生命周期协议、TCP/HTTP 权能凭证协议与 SQLite 快照保留协议，以及有限状态与精确有理数验证脚本，由本人负责实现与复跑；训练、验证与测试数据由配套脚本生成，本人负责数值整理与图表编排。支持本文结果的代码、原始数据与验证脚本随作品提交。

指导教师贡献：林海博教师指导选题与数据分析，帮助将多目标跟踪中的可变地址现象表述为可证明的记忆安全问题；就受控学习实验的对照组、随机种子、留出集划分及三套系统协议的可复现性提出建议，并核对实验数值与正文的一致性。陈森教师指导理论研究，与本人讨论递归签名、划分闭包及容量—遗憾分析的证明，审阅前提条件、适用范围与结论措辞。两位教师共同指导写作，审阅结构、术语与论证。指导采用面谈和线上讨论，平均每周一至两次，关键阶段增加，均未收费。两位教师不参与代码编写与数值计算；文中错误由参赛学生负责。

外部协助与基础设施：本文报告的受控学习与系统实验未使用外部数据集。提交前的实验复跑在参赛学生的个人计算机上完成。程序基于 Python、PyTorch、NumPy、Matplotlib 等公开发行的开源软件包，以及 SQLite 与操作系统自带的 POSIX 文件接口；POSIX 与 TCP/HTTP 实验在同一台机器上的多个进程之间通过本地回环通信，不涉及任何第三方服务器。学校与指导教师所在单位未提供额外的经费、设备或数据支持。

AI 工具使用：使用大语言模型辅助文字润色、代码与文档编写及一致性核对。

参考文献

- [1] Zhu Qiming, Chen Shunian, Yu Rui, et al. From lossy to verified: a provenance-aware tiered memory for agents [PP/OL]. arXiv (2026-02-20) [2026-09-12]. <https://arxiv.org/abs/2602.17913>.
- [2] Zhou Wei, Zhou Xuanhe, Han Shaokun, et al. Are we ready for an agent-native memory system? [PP/OL]. arXiv (2026-06-23) [2026-09-12]. <https://arxiv.org/abs/2606.24775>.
- [3] Zhang D, Lin Yanshan, Wu Zhengkun, et al. Useful memories become faulty when continuously updated by LLMs [PP/OL]. arXiv (2026-08-29) [2026-09-12]. <https://arxiv.org/abs/2605.12978>.
- [4] Yang Tianyu, Paul S, Srinivasan V, et al. TRUSTMEM: learning trustworthy memory consolidation for LLM agents with long-term memory[PP/OL]. arXiv (2026-06-23) [2026-09-12]. <https://arxiv.org/abs/2606.25161>.
- [5] Xu Jinghan, Xiao Yiyong, Shao Wanru, et al. Memory provenance laundering in LLM agents: a non-amplification firewall for persistent memory[PP/OL]. arXiv (2026-07-31) [2026-09-12]. <https://arxiv.org/abs/2607.29167>.
- [6] Zhan Qiuyang, Zhang Rui, Guo Sheng, et al. When memory becomes authority: benchmarking authority collapse at the memory consolidation boundary[PP/OL]. arXiv (2026-08-04) [2026-09-12]. <https://arxiv.org/abs/2608.01679>.
- [7] Jin Zhengda, Wang Bingbing, Li Jing, et al. Mitigating provenance-role collapse in long-term agents via typed memory representation[PP/OL]. arXiv (2026-05-25) [2026-09-12]. <https://arxiv.org/abs/2605.25869>.
- [8] Liu Chengzhi, Yang Yuzhe, Pu S X, et al. WorldMemArena: evaluating multimodal agent memory through action-world interaction[PP/OL]. arXiv (2026-06-01) [2026-09-12]. <https://arxiv.org/abs/2605.29341>.
- [9] Li Xiaoyang, Wang Yiqi, Lu Haohui, et al. MemTX: transactional belief commit for stateful agent memory[PP/OL]. arXiv (2026-07-28) [2026-09-12]. <https://arxiv.org/abs/2607.23929>.
- [10] Li Ziyide. Prediction is not memory: dual-timescale gated profile writing for persistent user modeling[PP/OL]. arXiv (2026-07-03) [2026-09-12]. <https://arxiv.org/abs/2607.24798>.
- [11] Subramanian J, Sinha A, Seraj R, et al. Approximate information state for approximate planning and reinforcement learning in partially observed systems[J/OL]. Journal of Machine Learning Research, 2022, 23 (12): 1-83. <https://www.jmlr.org/papers/v23/20-1165.html>.
- [12] Krishnamurthy V. Partially observed Markov decision processes: from filtering to controlled sensing[M/OL]. Cambridge: Cambridge University Press, 2016. DOI:10.1017/CBO9781316471104.
- [13] Littman M L, Sutton R S, Singh S. Predictive representations of state[C/OL]//Advances in neural information processing systems: v.14. MIT Press, 2001: 1555-1561. <https://proceedings.neurips.cc/paper/2001/hash/1e4d36177d71bbb3558e43af9577d70e-Abstract.html>.
- [14] Ferns N, Panangaden P, Precup D. Metrics for finite Markov decision processes[C]//Proceedings of the 20th conference on uncertainty in artificial intelligence. Arlington, VA: AUAI Press, 2004: 162-169.
- [15] Nath S, Baranwal M, Khadilkar H. Revisiting state augmentation methods for reinforcement learning with stochastic delays[C/OL]//Proceedings of the 30th ACM international conference on information and knowledge management. New York: Association for Computing Machinery, 2021: 1346-1355. DOI:10.1145/3459637.3482386.
- [16] Wang Wei, Han Dongqi, Luo Xufang, et al. Addressing signal delay in deep reinforcement learning[C/OL]//International conference on learning representations. 2024. https://proceedings.iclr.cc/paper_files/paper/2024/hash/45cc967d7899616f51993b7b363d35b5-Abstract-Conference.html.
- [17] Walsh M. Support sufficiency as action-sufficient compression: a single-cycle rate-regret formulation[PP/OL]. arXiv (2026-05-28) [2026-09-12]. <https://arxiv.org/abs/2606.09858>.
- [18] Reid M D, Williamson R C. Information, divergence and risk for binary experiments[J/OL]. Journal of Machine Learning Research, 2011, 12 (22): 731-817. <https://jmlr.org/papers/v12/reid11a.html>.
- [19] Larsen K G, Skou A. Bisimulation through probabilistic testing[J/OL]. Information and Computation, 1991, 94 (1): 1-28. DOI:10.1016/0890-5401(91)90030-6.
- [20] Hoeffding W. Probability inequalities for sums of bounded random variables[J/OL]. Journal of the American Statistical Association, 1963, 58 (301): 13-30. DOI:10.1080/01621459.1963.10500830.
- [21] The Open Group. The open group base specifications issue 7, 2018 edition: unlink, unlinkat[EB/OL]. (2018) [2026-09-12]. <https://pubs.opengroup.org/onlinepubs/9699919799/functions/unlink.html>.
- [22] SQLite Development Team. Write-ahead logging[EB/OL]. [2026-09-12]. <https://www.sqlite.org/wal.html>.

附录 A 完整理论细节

本附录把正文用到的有限构造与相应证明补全。划分细化、Wasserstein 距离与互模拟均为已有工具；本文使用这些工具分析记忆写入如何改变后续验证过程。

A.1 任意深度阶梯与信号延迟的区分

固定整数 $k \geq 1$ 。构造持久型历史 h_k^D 与易失型历史 h_k^F ，使二者具有相同的当前任务契约和深度 k 以内的递归签名，但最优初始记忆动作不同。构造中没有信号延迟：每个决策时刻选定的动作立即执行，随后的受控转移即时发生，观测在下次决策前送达。过程中没有待执行的动作队列或待送达的观测队列。

令提交终止验证通路，并获得价值

$$V_C = 0.80 - 0.06 = 0.74. \quad (39)$$

令等待在有限时域终端效用问题中，进入一条由成对匹配的普通状态构成的确定性链。中间奖励为零。初始选择之后，每个链上状态只有一个强制的“推进”动作；等价地，可以将每个动作标签映射到同一转移。因此，智能体无法通过在阶梯中途提交而重新规划。每条链边都是即时受控转移，而不是延迟送达事件。两条链的基础标签和观测直到最终验证转移前都一致。在最终转移处，持久型通路给出终端效用 0.95，易失型通路给出 0.60。故

$$V_W^D = 0.95, \quad V_W^F = 0.60, \quad (40)$$

初始最优动作分别为等待与提交，间隔为 0.21 和 0.14。

为证明签名结论，令 z_j^D, z_j^F 表示一对匹配链状态，该状态对与第一个来源敏感转移之间有 j 层匹配前驱。在最后一对终端前状态 z_0^D, z_0^F 处，基础标签相同，故 $S_0(z_0^D) = S_0(z_0^F)$ ；但下一次转移到不同验证标签，因此 $S_1(z_0^D) \neq S_1(z_0^F)$ 。归纳地，若 z_{j-1}^D, z_{j-1}^F 首次在深度 j 分离，则确定性的成对匹配前驱 z_j^D, z_j^F 在深度 $j-1$ 以内具有相同的后继概率律，并首次在深度 $j+1$ 分离。选择链长，使初始历史具有 k 层匹配递归层次，就得到

$$S_j(h_k^D) = S_j(h_k^F) \quad (j \leq k), \quad S_{k+1}(h_k^D) \neq S_{k+1}(h_k^F). \quad (41)$$

终端价值不依赖 k ，故初始动作间隔保持固定。这证明了定理 3.1。

这一构造还说明，精确细化深度可以随系统规模增长：终端验证区别在每轮细化中至多沿一条确定性阶梯边向后传播。

对于保留已有证据的扩展，让事实证据 X 沿两条链保持不变。在初始等待处收取持有成本 c ，随后强制进行零成本推进；提交则用基于 X 的贝叶斯预测终止过程。最终的等待试验具有贝叶斯风险 R_D 或 R_F 。例如，错误概率为 $1/4$ 的二元预览在 Brier 损失下给出 $R_0 = 3/16$ ；最终精确揭示真值给出 $R_D = 0$ ，保留预览但通路过期给出 $R_F = 3/16$ 。观测和后验状态均只有有限多个。在终端验证边之前，成对基础标签、观测、动作成本以及 X 的副本都一致。因此相同的签名归纳成立，初始间隔 $R_0 - R_D - c$ 和 $R_F + c - R_0$ 与链长无关。这里保留的是已有事实证据，而不是被省略的生命周期元数据：若把元数据加入 S_0 ，两条历史在深度零就可区分。

对于推论 3.2，用失去辅助验证机会替代持有费用。初始写入不可撤销地分配唯一的活动通路名。辅助目标独立，其先验 Brier 风险为 $1/4$ ，验证后风险为 $3/16$ 。因此终端总风险为：持久型等待下 $1/4$ ，易失型等待下 $7/16$ ，任一种历史的提交下 $3/8$ 。风险之差给出间隔 $1/8$ 和 $1/16$ ，且没有额外费用。在两条成对匹配、强制推进的转移链末端，执行初始动作所选的未来验证试验。每个初始动作下，两类生命周期的中间标签和观测一致，全部原有事实观测仍留在状态中。最终主目标验证区

分两条链，而辅助试验在两条链中的概率律相同。因此签名首次分离仍在 $k+1$ 。若有两个活动通路名额，保留两种机会就会消除易失型一侧的严格偏好，第 7.4 节对此作了测试。

为什么这不是信号延迟定理。 递归验证深度不同于延迟控制 MDP 中研究的动作 / 观测延迟。形式上，一个无延迟受控过程的纯有界信号延迟封装，是指其增广状态为 (x, q) 的封装，其中 x 为投影后的基础过程状态， q 为包含待执行动作和 / 或待送达观测的有限队列。出队影响投影后的基础转移何时发生，以及观测何时送达，却不改变投影后的转移概率律本身。增广环境状态 (x, q) 具有受控马尔可夫性。在部分可观测情况下，给定延迟历史的 (x, q) 后验分布因而是精确的信念信息状态，这是延迟 MDP 中使用的标准状态增广归约，而非本文的新结果。

我们的阶梯在每个时刻均有 $q = \emptyset$ ，因此这种增广不增加队列信息。两条历史的差别在于浅层记忆契约省略的证据源生命周期状态。该状态不影响前 k 层递归签名，却改变后续验证概率律，进而改变初始记忆决策。因此，物理信号延迟既不是定理的陈述对象，也不是不充分性的来源。该构造属于有限受控过程 / 部分可观测马尔可夫决策过程（partially observable Markov decision process, POMDP）。其对记忆的影响是：当记忆写入可以改变未来证据通路的可用性时，将持久写入契约验证到任意预先指定的有限预测深度，都不能提供与系统无关的后续安全证书。

A.2 固定动作间隔下的浅层近碰撞

从精确的持久型 / 易失型见证例出发。令 J^a 为动作 a 下共同的一步条件试验。对任意 $\epsilon > 0$ ，选取 $0 < \delta < \min\{\epsilon/2, 1/2\}$ ，并定义辅助的一步干扰变量 N ；给定历史时，它独立于原过程的全部状态、奖励、转移和验证变量。在持久型历史下令

$$N \sim \text{Bernoulli}(1/2 + \delta), \quad (42)$$

在易失型历史下令

$$N \sim \text{Bernoulli}(1/2 - \delta). \quad (43)$$

将这两种干扰变量概率律分别记为 P_D^N, P_F^N ，则一步试验变成乘积概率律 $J^a \otimes P_D^N$ 或 $J^a \otimes P_F^N$ 。由于第一个因子相同，且总质量为一，

$$\text{TV}(J^a \otimes P_D^N, J^a \otimes P_F^N) = \text{TV}(P_D^N, P_F^N) = 2\delta. \quad (44)$$

干扰变量在下一次决策前被丢弃，不出现在任何奖励或未来转移核中。因此，每个后续价值都与原见证例精确相同：

$$Q_D(W) = 0.95, \quad Q_D(C) = 0.74, \quad (45)$$

$$Q_F(W) = 0.60, \quad Q_F(C) = 0.74. \quad (46)$$

最优动作仍相反，间隔仍为 0.21 和 0.14，而

$$0 < d_1 = 2\delta < \epsilon. \quad (47)$$

这证明了命题 3.3。配套数值验证脚本评估 $\epsilon = 10^{-1}, \dots, 10^{-8}$ ，并在浮点精度内核对了两个固定间隔。

A.3 递归划分闭包

假设 $S, \mathcal{A}, \mathcal{O}$ 为有限集， $\kappa(o, s' | s, a)$ 为 $\mathcal{O} \times S$ 上的时齐马尔可夫核， $\ell: S \rightarrow \mathcal{L}$ 为有限值基础标签。在我们的有限模型中，采用共同动作字母表不失一般性：依赖状态的动作可行性可由 ℓ 及一个表示非法动作结果的吸收态编码。令 $\mathcal{P}_0$ 为 ℓ 诱导的划分。

对于细于 $\mathcal{P}_0$ 的划分 $\mathcal{P}$ ，以 $[s]_{\mathcal{P}}$ 表示当前划分块，并定义

$$\sigma_{\mathcal{P}}(s) = \left([s]_{\mathcal{P}}, \ell(s), \left(\sum_{s' \in B} \kappa(o, s' | s, a) \right)_{a \in \mathcal{A}, o \in \mathcal{O}, B \in \mathcal{P}} \right). \quad (48)$$

令 $F(\mathcal{P})$ 为按这一签名相等关系诱导的划分。坐标 $[s]_{\mathcal{P}}$ 防止已经分离的划分块再次合并，因此 $F(\mathcal{P})$ 始终细于 $\mathcal{P}$ 。若细于 $\mathcal{P}_0$ 的划分 $\mathcal{Q}$ 满足：同一 $\mathcal{Q}$ 块内的状态在每个动作下，向每一对 (o, C) ($C \in \mathcal{Q}$) 转移的概率质量都相同，则称其递归稳定。特别地，若 $\mathcal{P}_0$ 本身具有这种动作—观测可聚合性，则其签名在每个当前块内恒定，且 $F(\mathcal{P}_0) = \mathcal{P}_0$ 。所以递归细化并非不可避免；只有当前契约隐藏了后续概率律不同的状态时，才必须细化。持久型 / 易失型族通过依赖证据源生命周期的未来验证，违反了这一基准条件。

因此，迭代 $\mathcal{P}_{j+1} = F(\mathcal{P}_j)$ 是有限格上的单调细化链，最终到达不动点 $\mathcal{P}^*$ 。为证明最粗性，取任意细于 $\mathcal{P}_0$ 的递归稳定划分 $\mathcal{Q}$ 。我们归纳证明 $\mathcal{Q}$ 细于每个 $\mathcal{P}_j$ 。基础情形显然。若 $\mathcal{Q}$ 细于 $\mathcal{P}_j$ ，则每个 $B \in \mathcal{P}_j$ 都是若干 $\mathcal{Q}$ 块的并。同一 $\mathcal{Q}$ 块内两个状态向每个 $\mathcal{Q}$ 块转移的质量相同；对这些质量求和，得到每个 (a, o) 下向各 $B \in \mathcal{P}_j$ 转移的质量也相同。两个状态本就属于相同的当前 $\mathcal{P}_j$ 块，因此其 $\sigma_{\mathcal{P}_j}$ 签名相同，在 $F(\mathcal{P}_j)$ 下仍位于同一块。故 $\mathcal{Q}$ 细于 $\mathcal{P}_{j+1}$ 。到达不动点后， $\mathcal{Q}$ 细于 $\mathcal{P}^*$ 。于是 $\mathcal{P}^*$ 是 $\mathcal{P}_0$ 唯一的最粗递归稳定细化。

这是概率互模拟与状态最小化中标准的有限划分细化论证^[19]。本项目特有的对象是验证契约，其转移概率律本身也可能被记忆动作改变。

与递归签名的等价性。 假设有限受控状态 s 是相应历史的精确马尔可夫表示，且 $S_0(s) = \ell(s)$ 。每个 S_j 都可由 S_{j+1} 恢复：先直接恢复 S_0 ，再通过较低深度的签名映射，将各后继概率律投影为相应的低深度概率律。因此，签名划分相互嵌套。令 $\mathcal{P}_j$ 为 S_j 诱导的划分。结论在 $j = 0$ 时成立。若在深度 j 成立，则对于两个状态 s, t ，概率律

$$\mathcal{L}(O_{t+1}, S_j(s_{t+1}) | s, a) \quad \text{与} \quad \mathcal{L}(O_{t+1}, S_j(s_{t+1}) | t, a) \quad (49)$$

在每个动作下都相同，当且仅当对于每个观测 o 和每个块 $B \in \mathcal{P}_j$ ，两个核赋予 (o, B) 相同质量。再加上前一层划分块 / 基础标签相同，恰好等价于 $\sigma_{\mathcal{P}_j}$ 相同。因此 S_{j+1} 诱导 $F(\mathcal{P}_j) = \mathcal{P}_{j+1}$ 。由归纳，有限深度签名与划分构造一致。

A.4 有限停止与折扣近似递归

若 $|S| = n$ 且 $|\mathcal{P}_0| = b_0$ ，每次严格细化使块数至少增加一，而任意划分最多有 n 个块。因此，至多经过

$$n - b_0 \quad (50)$$

轮严格细化就达到精确闭包。任意深度阶梯不与有限终止矛盾：它让系统随 k 变化，因此排除的只是与系统无关的统一常数深度。

对于近似折扣控制，设 S 和 $\mathcal{A}$ 有限，奖励 $r : S \times \mathcal{A} \rightarrow \mathbb{R}$ 有界，转移概率为 $P(\cdot | s, a)$ ，且 $0 < \gamma < 1$ 。任何即时验证观测都可以并入后继状态，因此在有限情形下这种记法不损失信息。对每个动作 a ，假设 $\delta_a : S \times S \rightarrow [0, \infty)$ 是有界伪度量，且

$$\delta_a(s, t) \geq |r(s, a) - r(t, a)|. \quad (51)$$

令 $\mathcal{P}$ 为有限集 S 上有界伪度量的集合，并赋予上确界范数。它是有限维空间 $\mathbb{R}^{S \times S}$ 的闭子集，因而完备。简记 $P_s^a = P(\cdot | s, a)$ 。对 S 上的概率分布 μ, ν ，以 $\Pi(\mu, \nu)$ 表示以二者为边缘分布的耦合集 (coupling set)。对于 $d \in \mathcal{P}$ ，令

$$W_d(\mu, \nu) = \min_{\lambda \in \Pi(\mu, \nu)} \sum_{x, y \in S} \lambda(x, y) d(x, y) \quad (52)$$

为有限状态 1-Wasserstein 传输成本，并定义

$$(\mathcal{B}d)(s, t) = \max_{a \in \mathcal{A}} [\delta_a(s, t) + \gamma W_d(P(\cdot | s, a), P(\cdot | t, a))]. \quad (53)$$

只要 d 是伪度量，拉回 $W_d(P_s^a, P_t^a)$ 就是 (s, t) 上的伪度量；伪度量的和与逐点最大值仍是伪度量，因此 $\mathcal{B} : \mathcal{P} \rightarrow \mathcal{P}$ 。对两个有界底层成本函数，

$$|W_d(\mu, \nu) - W_{d'}(\mu, \nu)| \leq \|d - d'\|_\infty, \quad (54)$$

因为将任一成本下的最优耦合代入另一成本即可得到该界。因此

$$\|\mathcal{B}d - \mathcal{B}d'\|_\infty \leq \gamma \|d - d'\|_\infty. \quad (55)$$

由 Banach 不动点定理 (Banach fixed-point theorem)，存在唯一不动点 d^* 。由于 $\delta_a \geq 0$ ， $\mathcal{B}$ 对底层成本单调；当 $d^{(0)} = 0$ 时，迭代满足 $0 \leq d^{(0)} \leq d^{(1)} \leq \dots \leq d^*$ 。此外，

$$0 \leq d^*(s, t) - d^{(k)}(s, t) \leq \|d^* - d^{(k)}\|_\infty \leq \frac{\gamma^k}{1 - \gamma} \|d^{(1)} - d^{(0)}\|_\infty. \quad (56)$$

因为 $d^{(1)}(s, t) = \max_a \delta_a(s, t)$ ，有 $\|d^{(1)} - d^{(0)}\|_\infty = \Delta_0 := \sup_{s, t, a} \delta_a(s, t)$ ，从而证明式 (33)。

下面证明动作价值界。从 $V_0 \equiv 0$ 出发，令 $Q_{n+1}(s, a) = r(s, a) + \gamma \mathbb{E}_{P_s^a} V_n$ ， $V_{n+1}(s) = \max_a Q_{n+1}(s, a)$ ，作折扣价值迭代。归纳证明 V_n 关于 d^* 为 1-Lipschitz 函数 (1-Lipschitz function)。基础情形显然。若 V_n 为 1-Lipschitz，则任意耦合 $\lambda \in \Pi(P_s^a, P_t^a)$ 给出 $|\mathbb{E}_{P_s^a} V_n - \mathbb{E}_{P_t^a} V_n| \leq \mathbb{E}_\lambda d^*(X, Y)$ ；对耦合取最小值得

$$|Q_{n+1}(s, a) - Q_{n+1}(t, a)| \leq |r(s, a) - r(t, a)| + \gamma W_{d^*}(P_s^a, P_t^a) \quad (57)$$

$$\leq \delta_a(s, t) + \gamma W_{d^*}(P_s^a, P_t^a) \leq d^*(s, t). \quad (58)$$

对 a 取最大值，说明 V_{n+1} 也为 1-Lipschitz。奖励有界的折扣价值迭代一致收敛到 V^* ，故该性质传递到极限，同一推导得到

$$|Q^*(s, a) - Q^*(t, a)| \leq d^*(s, t) \quad \forall a. \quad (59)$$

令 a^* 为 s 的唯一最优动作， $m(s) > 0$ 为它与次优动作的间隔。若 $d^*(s, t) < m(s)/2$ ，则对任意 $a \neq a^*$ ，

$$Q^*(t, a^*) \geq Q^*(s, a^*) - d^*(s, t), \quad (60)$$

$$Q^*(t, a) \leq Q^*(s, a) + d^*(s, t), \quad (61)$$

故 $Q^*(t, a^*) - Q^*(t, a) \geq m(s) - 2d^*(s, t) > 0$ 。代入 d^* 的有限深度上界，即证明式 (34)。

还可得到一个有用的反向分离界。若 s, t 的唯一最优动作 a_s, a_t 不同，间隔分别为 $m(s), m(t)$ ，则

$$m(s) + m(t) \leq [Q^*(s, a_s) - Q^*(s, a_t)] + [Q^*(t, a_t) - Q^*(t, a_s)] \quad (62)$$

$$\leq 2d^*(s, t), \quad (63)$$

因此

$$d^*(s, t) \geq \frac{m(s) + m(t)}{2}. \quad (64)$$

对于动作间隔为 0.21 和 0.14 的折扣实例，此下界为 0.175。若要将其用于终端效用被推迟的构造，须先检查所选折扣约定下的实际动作间隔。

附录 B 有限容量结果的证明

B.1 受保护核心下的深度宽度

假设当前任务划分 $\mathcal{P}_0$ 得到无损表示。对于受保护划分块 $B \in \mathcal{P}_0$ ，定义

$$w_k(B) = |\{C \in \mathcal{P}_k : C \subseteq B\}|, \quad w_k = \max_{B \in \mathcal{P}_0} w_k(B). \quad (65)$$

由于 $\mathcal{P}_{k+1}$ 是 $\mathcal{P}_k$ 的细化，故 $w_{k+1} \geq w_k$ 。

定理 4.1 的必要性可如下证明。取达到 w_k 的划分块 B 。 B 内的受保护编码相同，因此，若残差编码的符号数 $K < w_k$ ，必有两个深度为 k 的子类被映射到同一残差符号，无法精确识别深度为 k 的类别。对于充分性，在每个 B 内为各深度为 k 的子类分配不同的局部残差标签即可；由于 B 已知，不同受保护划分块之间可以复用标签。因此，精确阈值为 $K = w_k$ 。

对于局部宽度的延迟跃升，为阶梯的每一阶段赋予独立的受保护标签，并令 B 恰好包含初始的持久型与易失型历史对。此时 $w_0(B) = \dots = w_L(B) = 1$ ，而 $w_{L+1}(B) = 2$ ，相反最优动作的固定间隔由式 (2) 和 (3) 给出。 $L = 0$ 时，直接以对终止结果敏感的历史对作为初始历史对。全局宽度在第一轮已达到二，因为最后一个终止前划分块在该轮分裂。全局最大宽度不可能从一延迟跃升： $w_1 = 1$ 意味着 $F(\mathcal{P}_0) = \mathcal{P}_0$ ，此后所有划分均保持不变。

B.2 决策关键宽度

固定一个下游决策，假设其最优动作 $a^*(C)$ 唯一，且在每个深度为 k 的子类 C 内恒定。若同一子类中的两段历史需要不同动作，仅有各自最优动作的唯一性并不够。对于受保护划分块 B ，定义

$$\chi_k(B) = |\{a^*(C) : C \in \mathcal{P}_k, C \subseteq B\}|, \quad \chi_k = \max_{B \in \mathcal{P}_0} \chi_k(B). \quad (66)$$

若两个子类具有相同的最优动作，它们可以共享一个残差编码而不产生动作遗憾；若它们需要不同的唯一最优动作，则不能共享。因此， B 内最少需要的局部编码数恰为 $\chi_k(B)$ 。标签可在不同受保护划分块之间复用，故全局充要条件为 $K \geq \chi_k$ 。由 $\chi_k(B) \leq w_k(B)$ 可得 $\chi_k \leq w_k$ 。

若最优动作不唯一，编码单元 G 的遗憾为零，当且仅当 $\bigcap_{i \in G} A_i^* \neq \emptyset$ ，其中 $A_i^* = \arg \max_a Q_i(a)$ 。此时，最少编码数等于将各子类划分为这类相容组所需的最少组数。正文采用唯一最优动作，是因为持久型与易失型见证例以及学习实验中的决策指标均具有严格的动作间隔。

当 $\chi_k \leq K < w_k$ 时，完整的递归契约无法重建，但特定下游任务的动作仍可保持精确。在持久型与易失型见证例中，两个递归子类需要相反动作，故 $w = \chi = 2$ 。

B.3 有限容量下的精确合并遗憾

在一个受保护划分块内，设递归子类 C_i 的概率质量为 p_i ，类内恒定的动作价值为 $Q_i(a)$ ，最优价值为 $V_i = \max_a Q_i(a)$ 。沿用第 4.3 节的共同后续策略可实现条件式 (18)：对每个编码组 G_z 与动作

a , 存在同一个只依赖保留编码与后续合法观测的策略 $\rho_{z,a}$, 同时实现组内各类的价值。确定性残差编码器 e 诱导组 $G_z = \{i : e(i) = z\}$ 。给定编码 z , 贝叶斯最优控制器选择

$$a_z \in \arg \max_a \sum_{i \in G_z} p_i Q_i(a). \quad (67)$$

该组在完全信息下贡献的价值为 $\sum_{i \in G_z} p_i V_i$, 而仅使用编码时的最优价值为 $\max_a \sum_{i \in G_z} p_i Q_i(a)$ 。对所有编码组的这两个价值之差求和, 即得式 (19)。在至多使用 K 个编码的编码器上取最小值得到 R_K^* ; 增大 K 不会使这一最优值增加。

以下二类显式公式针对本文的两动作问题 $\mathcal{A} = \{\text{等待}, \text{提交}\}$ 。对于一般动作集, 应继续使用式 (19), 对全部候选动作求最优值; 仅比较两个类别各自的最优动作, 可能高估合并损失, 因为其他候选动作可能具有更高的组内平均价值。

对于先验概率质量为 $q, 1-q$ 的两个子类 D, F , 设其最优动作相反, 误选对方最优动作的遗憾分别为 α, β 。单个残差编码迫使两类共用一个动作。选择持久型的最优动作, 代价为 $(1-q)\beta$; 选择易失型的最优动作, 代价为 $q\alpha$ 。随机化不能改善仿射目标, 故

$$R_1^* = \min\{q\alpha, (1-q)\beta\}. \quad (68)$$

取 $q = 1/2$ 、 $(\alpha, \beta) = (0.21, 0.14)$, 得到 0.07。

B.4 随机或连续编码器

以下沿用上一小节的两动作设定 $\mathcal{A} = \{\text{等待}, \text{提交}\}$, 对一般动作集同样应回到式 (19)。令 P_D^Z 和 P_F^Z 为均衡持久型与易失型历史对的表示分布, 以 $\lambda = P_D^Z + P_F^Z$ 作为二者的控制测度 (dominating measure), 记 $p_D = dP_D^Z/d\lambda$ 、 $p_F = dP_F^Z/d\lambda$ 。在表示值 z 处, 选择持久型的最优动作产生易失侧的损失密度 $\beta p_F(z)/2$; 选择易失型的最优动作则产生持久侧的损失密度 $\alpha p_D(z)/2$ 。贝叶斯控制器逐点选择较小者, 得到仅依赖表示 Z 的最小期望遗憾

$$R_{\text{rep}}^*(Z) = \frac{1}{2} \int \min\{\alpha p_D(z), \beta p_F(z)\} d\lambda(z) = \frac{\alpha + \beta - \int |\alpha p_D - \beta p_F| d\lambda}{4}. \quad (69)$$

这一形式不要求分布相对于 Lebesgue 测度 (Lebesgue measure) 存在密度。精确编码碰撞对应极端情形 $P_D^Z = P_F^Z$; 若表示分布互相奇异, 则不可恢复的遗憾为零。

不确定的生命周期与审计分布重叠。 当 Z 是观测到的审计历史而非编码器输出时, 同一加权重叠公式仍然适用。假设审计依赖于生命周期, 但在给定生命周期与已有事实后, 与目标及未来验证结果条件独立。另假设存在一个共同且可实现的最终读出器, 无须事先知道生命周期, 便能达到以生命周期为条件的贝叶斯风险。双目标构造满足这一条件: 响应要么揭示真值, 要么使共同预览保持不变。在这些条件下, 审计改变后验概率 π , 但不改变信息价值 Δ_D, Δ_F , 条件风险因而线性混合。对于 $g = \Delta_D - \Delta_F > 0$ 和 $\Delta_F < c < \Delta_D$, 提交风险与等待风险之差为 $g(\pi - \tau)$, 其中

$$\tau = \frac{c - \Delta_F}{g}. \quad (70)$$

以 $R(\hat{\pi})$ 和 $R(\pi)$ 分别表示按估计后验与真实后验决策的期望风险。用估计后验概率 $\hat{\pi}$ 代替 π , 仅在两者位于 τ 两侧时, 才会相对后验最优动作产生超额风险。在该事件上, $|\pi - \tau| \leq |\pi - \hat{\pi}|$, 故

$$R(\hat{\pi}) - R(\pi) \leq g \mathbb{E}|\hat{\pi} - \pi|. \quad (71)$$

这是代入估计后验概率所得到的决策界，不同于相对已知生命周期的理想决策器而言不可消除的遗憾。在双目标构造中， $g = 3/16$ ， $\tau = 1/3$ 。

对于 n 次审计，假设给定生命周期后，各次审计结果相互独立；两类生命周期下的成功概率分别为 $3/4$ 和 $1/4$ ，其二项计数分布记为 $P_D^{(n)}, P_F^{(n)}$ 。以 R_n^* 记审计后相对已知生命周期理想决策器的最小期望遗憾。将错误动作损失 $\alpha = 1/8$ 、 $\beta = 1/16$ 代入上述重叠公式，得到

$$R_n^* = \frac{1}{2} \sum_{s=0}^n \min\{\alpha P_D^{(n)}(s), \beta P_F^{(n)}(s)\} \leq \frac{\sqrt{\alpha\beta}}{2} \left(\frac{\sqrt{3}}{2}\right)^n. \quad (72)$$

不等式由 $\min(x, y) \leq \sqrt{xy}$ 及二项式定理得到：似然乘积平方根之和为 $(2\sqrt{(3/4)(1/4)})^n$ 。因此，重复的有效审计即使不能完全识别生命周期，也能使相对理想决策器的遗憾指数衰减。精确枚举得到 $R_0^* = 0.03125$ 、 $R_8^* = 0.00526$ 、 $R_{16}^* = 0.00131$ 和 $R_{32}^* = 0.00010$ 。若没有观测，或观测与生命周期独立，相对于已知生命周期的理想决策器，遗憾仍为 $1/32$ 。这里的 n 是过去观测的数量，不是递归预测深度：必须先有包含信息的证据，记忆契约才可能保留其后续影响。

第 6.4 节的拟合检验预先选定 $n = 8$ ，使用带截距、正则化参数 $C = 1$ 且仅以计数为输入的逻辑回归模型。训练样本与测试样本分别生成；最终可用性在训练中作为标签，在测试中仅供评估器使用，从不作为测试特征。拟合策略在持久型上出错 58 次，在易失型上出错 242 次，故遗憾为 $(58/8 + 242/16)/4096 = 0.00546265$ 。对固定的拟合策略，相比始终等待，每个样本的改进位于 $[-1/8, 1/16]$ 。对于均衡分层下的 $N = 4096$ 个独立样本，Hoeffding 不等式给出的主要改进区间半宽为 $(3/16)\sqrt{\log(2/0.05)/(2N)}$ 。该区间以训练样本为条件，并非对所有拟合模型的保证。

B.5 成本族恒等式与后验容量

固定定理 4.3 中的模型及共同读出器条件。令 $t = (c - \Delta_F)/g \in [0, 1]$ 。基于 Z 的控制器在后验概率 $\Pi = p$ 下，相对理想决策器的条件遗憾为

$$g \min\{p(1-t), (1-p)t\}.$$

两条直线在 $t = p$ 处相交。分别在 $[0, p]$ 与 $[p, 1]$ 上积分，得到 $gp(1-p)/2$ 。损失非负且有界，因此可以交换积分与期望，得到式 (27) 的第一个恒等式。

由于编码器仅使用 H 及与生命周期独立的随机化， $M-H-Z$ 构成马尔可夫链，且 $\Pi = \mathbb{E}[\eta | Z]$ 。条件方差分解给出

$$\mathbb{E}[\Pi(1-\Pi)] - \mathbb{E}[\eta(1-\eta)] = \mathbb{E}[(\eta - \Pi)^2],$$

从而证明第二个恒等式。额外成本平均遗憾为零当且仅当 $\eta = \Pi$ 几乎处处成立，因而保留成本族中每个成本下的后验最优动作风险。反之，若每个成本下的风险均得到保留，则成本平均超额遗憾为零。相对理想决策器的成本平均遗憾为零是更强的要求：它需要 $\Pi \in \{0, 1\}$ 几乎处处成立，即完全识别生命周期。

对于均匀后验概率下的容量结论，最小化超额遗憾等价于最小化 $\eta \sim \text{Unif}[0, 1]$ 的平方误差量化损失。将每个后验概率分配给最近的编码代表点已经最优，随机分配不能进一步改善。相应量化单元为区间。长度为 ℓ 的单元，其积分平方误差最小值为 $\ell^3/12$ ，在中点处达到。至多 K 个单元的总长度为一，由凸性可得 $\sum \ell^3/12 \geq 1/(12K^2)$ 。等长区间达到等号；再乘以 $g/2$ ，即得式 (28)。这一模型可以通过先均匀抽取 η ，再令 $M | \eta \sim \text{Bernoulli}(\eta)$ 来实现。

对于最坏成本界，记 $x_+ = \max(x, 0)$ ，并以 g 归一化超额遗憾。在阈值 t 处，归一化后的超额遗憾为

$$e(t) = \mathbb{E}[(\eta - t)_+] - \mathbb{E}[(\Pi - t)_+] = \frac{(1-t)^2}{2} - \mathbb{E}[(\Pi - t)_+] \geq 0,$$

其中非负性来自条件 Jensen 不等式 (conditional Jensen's inequality)。若 Π 有 $k \leq K$ 个不同取值 $p_1 < \dots < p_k$, 则在 $[0, p_1]$ 上 $e(t) = t^2/2$, 在 $[p_k, 1]$ 上 $e(t) = (1-t)^2/2$, 相邻取值之间则有 $e'' = 1$ 。记 $\epsilon = \sup_t e(t)$ 。两端区间各自的长度至多为 $\sqrt{2\epsilon}$ 。对于长度为 ℓ 的内部区间, 其中点函数值等于两端点函数值的平均数减去 $\ell^2/8$ 。由非负性可得 $\ell \leq \sqrt{8\epsilon}$ 。对全部区间长度求和, 得到 $1 \leq k\sqrt{8\epsilon}$, 故即使允许随机编码器, 仍有 $\epsilon \geq 1/(8K^2)$ 。使用长度为 $1/K$ 、中点为 p 的等长区间 $[a, b]$ 时, 超额遗憾在 $a \leq t \leq p$ 上为 $(t-a)^2/2$, 在 $p \leq t \leq b$ 上为 $(b-t)^2/2$, 达到该界。编码器在成本指定之前固定; 允许编码依赖成本的二动作编码属于另一种契约。

对八次审计的九个可能计数的全部 512 个二元编码器, 精确分段线性积分均验证了方差恒等式。随后, 有序动态规划给出 $K = 1, 2, 4, 9$ 时最小成本平均超额遗憾分别为 0.0190713、0.0016514、0.0000731 和零。九个计数的后验概率互不相同, 因此, 要精确保留整个成本族, 需要九个编码; 若只考虑其中一个预先选定的成本, 两个编码便已足够。独立的精确检验验证了 $K = 1, \dots, 16$ 时均匀后验概率的两个最优值, 并以 128 个随机区间编码器检验恒等式与极小极大下界 (minimax lower bound)。将九个审计计数粗化为三个契约单元, 再枚举这些单元的全部八个二元编码器, 也通过精确积分验证了推论 4.4 中的两项恒等式。这些检验只是有限实例, 不能代替证明; 离散审计的数值也不被宣称服从均匀后验概率下的收敛率。

附录 C 受控学习实验

C.1 任务构造与指标

受控学习实验使用配对历史: 同一对中的两条历史在各动作下具有相同的即时统计试验目标, 但决定后续验证试验的生命周期变量不同。神经网络实验通过输入词元显式提供该变量。单步目标预测即时统计试验, 递归目标还要求预测后续统计试验。下游指标均基于学习到的瓶颈表示计算。

对于给定决策问题, 主要指标为策略遗憾和动作准确率, 辅助诊断包括统计试验分布的 KL 散度及生命周期探针准确率。各实验采用不同的探针架构: 预测目标阶梯与小型 Transformer 使用线性探针, GRU 与受保护 S_1 实验使用非线性探针。受保护 S_1 实验的 KL 散度来自直接预测头, 架构检验则使用另行拟合的解码器。因此, 这些诊断指标不能直接互换。

C.2 预测目标阶梯

预测阶梯采用以下浅层目标数量设置:

$$k \in \{1, 2, 4, 8, 16, 32\}.$$

对于每个成对基础样本, 长度为 20 个词元的序列具有对内共享的数值坐标、干扰位及噪声流; 只有一个生命周期词元不同。在编号 $1, \dots, k$ 的每个浅层目标上, 软目标都是共享数值坐标的确定性光滑函数, 因而持久型与易失型目标在目标编号 k 以内完全相同。第一个额外目标均衡地取 $\{0.95, 0.60\}$ 中的值。

这些软目标是人为构造的替代目标, 并非通过递归转移模拟生成的完整 S_j 分布。精确有限状态阶梯另在附录 G 中检验。

三个实验组分别为: (i) 浅层组, 每次更新均匀抽取一个编号在 $1, \dots, k$ 内的浅层目标; (ii) 可观测干扰匹配对照组, 在相同浅层损失之外, 接收由对内共享干扰位决定的额外目标; (iii) 生命周期敏感目标组, 在相同浅层损失之外, 接收由生命周期决定的额外目标。由于干扰位在各成对基础样本之间均衡, 生命周期又由构造保证均衡, 对照组与该组额外目标的边缘分布及方差完全相同。两种线索均以显式输入词元给出, 唯一差别是额外目标与持久型、易失型区别之间的关联。

下游动作价值为

$$Q(D, W) = 0.95, \quad Q(F, W) = 0.60, \quad Q(\cdot, C) = 0.74.$$

若表示在历史对内不变，共同选择等待的平均成对遗憾为 $(0 + 0.14)/2 = 0.07$ ，共同选择提交则为 0.105。因此，0.07 是分析判据中使用的精确成对碰撞下界。

编码器是单层因果 Transformer ($d_{\text{model}} = 24$)，四个注意力头，前馈层宽 48，序列长度 20，瓶颈为三维。训练用 AdamW，共 220 次更新，批大小 128，学习率 0.003，权重衰减 0.004，训练集含 900 个成对基础样本。岭回归解码器以训练得到的瓶颈表示为输入，拟合两个已知动作价值，并在 300 个留出的成对基础样本上评估。线性生命周期探针仅用于辅助诊断，决策性能由下游动作指标评价。

训练、验证与测试数据分别独立生成。报告的测试集包含 300 个留出的成对基础样本；每组、每个 k 设置均使用五个优化器随机种子，共 90 次运行。

在全部六种目标数量设置上，浅层组与干扰匹配组的五种子平均遗憾，与表示对内不变时的参考值 0.07 之差均不超过 5×10^{-5} ；全部 30 次生命周期敏感目标组运行的遗憾均为零，动作准确率均为一。浅层组与对照组的生命周期探针准确率接近随机猜测水平，生命周期敏感目标组的每次运行均为 1.0。学习到的嵌入未必严格对内不变；这些结果针对拟合出的读出器，不是对所有可能读出器的信息论下界。

输入已包含生命周期线索，实验只评估所测试读出器能否利用该信息，不能证明所有浅层架构都会将其丢失。边缘分布与方差匹配、且在历史对内共享的额外监督，未使读出器可靠利用生命周期信息；加入对后续证据存续敏感的目标后则可可靠利用。理论证明不依赖这些学习结果。

C.3 GRU 结果

最初的 GRU 比较采用浅层训练 320 次更新、递归训练 700 次更新。为补全对照，我们增加了递归训练 320 次更新和浅层训练 700 次更新两个组合。数据生成方式、优化器设置、五个优化器随机种子及读出流程均保持不变；原有两组结果保留，新增两组各使用五个种子。表 7 汇总全部四种组合。

表 7 GRU 在监督目标与更新预算两个维度上的比较

更新次数	目标	两步策略遗憾	两步 KL 散度
320	浅层	0.05659 ± 0.00071	0.18854 ± 0.00643
320	递归	0.03258 ± 0.02659	0.13370 ± 0.07070
700	浅层	0.04828 ± 0.00730	0.16649 ± 0.02924
700	递归	0.00284 ± 0.00076	0.04185 ± 0.01177

注：表内数值都是同一组五个优化器随机种子上的均值 $\pm$ 标准差。

在 320 次更新下，以递归训练为正向改进的成对遗憾差分别为 -0.00012 、 0.00044 、 0.01545 、 0.05146 和 0.05282 。更新 700 次时，五个改进量一律为正，落在 0.03682 到 0.05491 之间。递归训练的非线性生命周期探针准确率，在 320 次更新时为 0.715 ± 0.259 ，到 700 次更新时升为 1.000 ± 0 ；两种预算下浅层探针的表现都接近随机猜测。每个冻结表示都单独配一个两步解码器。综合来看，预算较长时收益是稳定的，预算较短时训练本身不够稳；这组实验并没有给出任何普遍的收敛率结论。

C.4 小型因果 Transformer

架构对照实验用的是两层因果 Transformer ($d_{\text{model}} = 24$)，四个注意力头，前馈层宽 48，序列长度 24，学习瓶颈四维，可训练参数合计 11 328 个。每次运行使用 1200 对训练样本和 400 对测试样本，更新 320 次。实验共使用五个独立随机种子，并且不提供任何直接的持久性标签监督。表 8 汇总这五个种子的结果。

按随机种子配对计算，两步 KL 散度的改进约为 0.04172 (95% 的 t 区间为 $[0.02529, 0.05816]$ ，配对检验 $p \approx 0.00213$)；两步策略遗憾的改进为 0.05081 (95% 区间为 $[0.04688, 0.05475]$ ，配对检验

表 8 小型因果 Transformer 上五个随机种子的均值 $\pm$ 标准差

指标	单步	递归
生命周期探针	0.574 ± 0.038	1.000 ± 0
两步 KL 散度	0.0592 ± 0.0024	0.0175 ± 0.0122
两步策略遗憾	0.0508 ± 0.0032	0.0000 ± 0

$p \approx 3.6 \times 10^{-6}$)。浅层 Transformer 偶然保留了部分持久性信息，因此，结果支持的是浅层目标不保证递归充分性，而不是浅层目标必然抹去这些信息。

这些区间刻画的是在一套固定生成的训练集与测试集下，优化器随机种子带来的波动；至于独立采样环境之间、或者已部署智能体任务之间存在多大不确定性，它们并未给出度量。

C.5 精确 S_1 保护下的残差

本实验沿用与 GRU 实验同一套成对数据生成器，训练、验证与测试数据各自独立生成，互不重叠。

每段历史的受保护向量为

$$S_1(h) = [p(h), J_{1,w}(h), J_{1,c}(h)] \in \mathbb{R}^{33}. \quad (73)$$

这个向量完全确定、不含可训练参数，并以逐位相同的方式复制给两个实验组。两组的残差编码器都是两层因果 Transformer ($d_{\text{model}} = 48$)，四个注意力头，前馈网络宽 96，瓶颈四维。两组都跑 400 次 AdamW 更新，使用的当前有效性与单步统计试验目标也相同；只有递归组额外预测两步的等待与提交试验。持久性标签不进入训练损失。编码器冻结之后，两组各配一个事先规定好的同构解码器（具有两层隐藏层、宽度分别为 64 与 32 的多层感知机），只用训练样本把 $[S_1, R_\theta]$ 映射为两个时域为 2 的动作价值。

在新生成的测试集中，最优动作不同的历史对占全部历史对的比例为 0.55；下文称这些历史对为决策分离对。解析碰撞下界与五种子浅层遗憾均四舍五入为 0.023324，绝对差小于 4×10^{-9} 。浅层组对每一对历史都预测相同动作，因此其预测动作分歧率为零；在决策分离对的全部成员中，预测动作正确的比例为 0.500。用 R_{H2} 记两步策略遗憾，其中 H2 表示时域为 2 (horizon 2)。五个递归种子的结果为

$$R_{H2} = 0.000345 \pm 0.000135, \quad (74)$$

递归组预测对内动作不同的历史对占全部历史对的比例为 0.580 ± 0.005 。在决策分离对的全部成员中，预测动作正确的比例为 0.996 ± 0.004 ；在全部决策分离对中，对内两个动作均预测正确的比例为 0.993 ± 0.008 。残差两步 KL 散度从 1.422 ± 0.168 降至 0.0603 ± 0.0043 。辅助生命周期探针准确率分别为 0.510 ± 0.008 和 1.000 ± 0 。

两步 KL 散度由直接预测头计算，浅层组的两步预测头未经过训练。因此，这项诊断没有对两组冻结表示另行拟合相同的预测解码器，不能据此比较可恢复信息的多少。辅助生命周期探针采用非线性模型。精确的 S_1 是检验受保护核心假设所设置的控制条件，实际部署中仍需估计或保留相应统计量。策略评估表明，在当前与单步信息固定时，递归残差仍可改善所测试读出器的决策性能。

附录 D POSIX 权威证据源到期协议

实际系统实验显式测量当前状态、动作执行后但最终写入前的观测 O_1 ，以及后续观测 O_2 ，从而将浅层相等与后续证据丢失作为不同事件分别测量。

独立 Python 进程持有权威证据源。对于每个成对样本，它创建四个相同的文件系统副本，分别对应每个来源类型与动作的反事实组合，预览位为 y_0 。父进程为持久型副本打开文件描述符；易失型副本仅以路径名表示。显式测量动作后观测的事件顺序为

$$t_0 : \text{当前读取} \rightarrow a \in \{W, C\} \rightarrow O_1 \rightarrow \text{最终写入、fsync、解除链接} \rightarrow O_2.$$

在 O_1 时，等待动作读取仍有效的路径，提交动作返回序列化快照。对每个动作，持久型与易失型来源下的这些观测必须完全相同。只有此后，证据源才写入最终真值并解除链接。持久型等待保留已打开的文件描述符；易失型等待仅剩已经失效的路径名。持久型提交已关闭其文件描述符，因而两类提交历史在 O_2 使用同一快照。

生成器使用 $G \sim \text{Bernoulli}(1/2)$ ，预览翻转概率为 $q = 0.25$ 。不拟合学习读出器。在声明的保留契约下，等待仅保留引用，不保留先前预览用于最终预测。成功读取最终权威值时，预测分布为点质量；证据源缺失时，预测为 $p(G = 1) = 0.5$ ；提交的预览使用 $P(G = y_0) = 0.75$ 。若等待后也保留预览，这一固定回退预测便不再是贝叶斯最优，易失侧的严格动作反转也会消失。因此，在给定读出器下，二元 Brier 损失的总体期望为

$$L_{D,W} = 0, \quad L_{D,C} = L_{F,C} = q(1 - q) = 0.1875, \quad L_{F,W} = 0.25.$$

持久型等待与易失型提交的期望动作间隔分别为 0.1875 和 0.0625。更一般地，对任意 $0 < q < 1$ 且 $q \neq 1/2$ ，校准快照损失均满足 $0 < q(1 - q) < 0.25$ ，因此相同的严格动作反转仍成立；实验中的 $q = 0.25$ 并非仅在孤立参数点上奏效的选择。

留出测试集共 1024 对样本。三项一致性检查——当前状态、等待动作条件下的 O_1 、提交动作条件下的 O_1 ——全部落在 1024/1024。持久型等待在解除链接之后仍取回最终真值的比例是 1024/1024；易失型等待的路径查找失败比例同样为 1024/1024。实际测得的预览翻转率为 0.25098，据此按确定性规则算出的提交 Brier 损失为 0.18799。于是观测到的持久型等待间隔为 0.18799，易失型提交间隔为 0.06201。

在声明的观测与保留契约下，这给出了由操作系统机制实现的深度为 1 的碰撞与深度为 2 的分裂。打开的引用携带与后续任务有关的访问能力；仅保留快照时，这一能力会丢失。本检验不估计部署智能体中的故障频率。

附录 E 网络签名权能凭证生命周期协议

网络系统实验检验该机制能否在不同于 POSIX 文件引用的资源访问机制中成立。独立的 ThreadingHTTPServer 进程持有权威证据源。每条记录同时提供公开定位符与 HMAC-SHA256 签名权能凭证。客户端仅通过 TCP/HTTP 通信。因此，持久型与易失型的区别是协议边界上的引用权限区别，而不是是否持有打开的文件描述符。

对每对样本，先抽取源预览 y_0 ，再以概率 $q = 0.25$ 翻转预览以获得最终真值 G 。最终写入前，公开定位符与签名权能凭证返回相同值。事件顺序为

$$\text{当前状态} \rightarrow a \in \{W, C\} \rightarrow O_1 \rightarrow \text{最终写入并使公开定位符到期} \rightarrow O_2.$$

在等待下，持久型来源保留签名权能凭证，易失型来源仅保留公开定位符。在提交下，两者都产生相同的校准快照预测 $P(G = y_0) = 0.75$ ，此后不再通过活动引用获取数据。实验在最终写入前，分别测量等待与提交下持久型、易失型 O_1 的相等性。最终写入更新权威值，将公开定位符标为到期，此后对该定位符返回 HTTP 410；签名权能凭证仍有效，并返回最终值。

不拟合学习读出器。与 POSIX 协议相同，最终的等待读出器持有保留的引用，但不持有先前预览。持久型最终获取成功时，在 G 处给出点质量；易失型定位符到期时，给出 $p(G = 1) = 0.5$ ；提交的预览使用校准转移概率。因此，在这些保留规则下，总体 Brier 损失与 POSIX 实验具有相同次序：

$$L_{D,W} = 0, \quad L_C = q(1 - q) = 0.1875, \quad L_{F,W} = 0.25.$$

留出测试集包含 1024 对样本。当前状态相同、等待下 O_1 相同、提交下 O_1 相同，三项均为 1024/1024。公开定位符到期后，持久型签名权能凭证在 1024/1024 个样本中返回最终权威值；易失型公开定位符在 1024/1024 个样本中返回 HTTP 410。实际翻转率为 0.24805，因此平均提交 Brier 损失为 0.18652。观测到的持久型等待间隔为 0.18652，易失型提交间隔为 0.06348。

这是真实的客户端—服务器生命周期实验，使用不同进程，通过回环 TCP/HTTP 通信。它不说明公共互联网上此类故障的普遍程度，但消除了对 POSIX 解除链接语义的依赖，并在第二种实际引用模型下展示了相同的浅层碰撞与递归分裂机制。

E.1 保留证据与活动通路预算

预算扩展同时给两个彼此独立、先验均衡的二元目标 G, H 打分。已有记录里存着 G 的预览 X （错误概率 $1/4$ ），以及 H 的先验。主通路在持久型生命周期下会把 G 揭示出来，在易失型生命周期下则到期失效。辅助通路给出 H 的一个含噪观测 Z ，错误概率同样是 $1/4$ 。两类错误相互独立。由于每个读出器都留着 X ，主通路一旦到期，其校准预测不会发生变化。

写入将主通路或辅助通路之一分配到唯一的活动通路名额。服务器保留被选中的访问权，即使未选令牌的字节已在写入前复制，也拒绝该令牌。双名额对照保留两条通路。这是对活动授权的限制，不是对复制预览的限制，也不是一次 HTTP 请求的收费。写入后，证据源不允许改变已经提交的分配。实验控制器负责证据源的创建与最终写入；每个读出器只能访问其自身的反事实实例，不能新建证据源或更改分配。这些是受控访问实验，并非对生产服务安全性的声称。

以 R 记总贝叶斯风险，下标 *both* 表示同时保留两条通路。枚举 (G, H, X, Z) 的 16 种联合结果，得到

$$R_{D,W} = \frac{1}{4}, \quad R_{F,W} = \frac{7}{16}, \quad R_C = \frac{3}{8}, \quad R_{D,both} = \frac{3}{16}, \quad R_{F,both} = \frac{3}{8}. \quad (75)$$

诱导的机会成本为 $1/4 - 3/16 = 1/16$ ，故单名额的动作间隔为 $1/8$ 和 $1/16$ 。均衡历史对若被迫共享一个动作，最小遗憾为 $1/32$ 。以风险为状态摘要的有限状态阶梯检验了这些常数及其对深度的独立性；另一项精确联合分布计算则检验了该化简所用的贝叶斯风险。

设测试集含 n 对样本， $\mathbf{1}\{\cdot\}$ 为指示函数。令 $E_i = \mathbf{1}\{X_i \neq G_i\}$ 、 $F_i = \mathbf{1}\{Z_i \neq H_i\}$ 。观测动作间隔满足精确恒等式

$$\hat{m}_D = -\frac{1}{8} + \frac{1}{2n} \sum_{i=1}^n (E_i + F_i), \quad \hat{m}_F = \frac{3}{16} - \frac{1}{2n} \sum_{i=1}^n F_i. \quad (76)$$

取联合失效概率上界 $\alpha \in (0, 1)$ 。分别对第一个表达式中的 $2n$ 个独立错误指示变量和第二个表达式中的 n 个指示变量应用 Hoeffding 不等式^[20]，再用并集界（union bound），得到两个区间的同时覆盖概率至少为 $1 - \alpha$ ，半宽分别为

$$b_D = \sqrt{\frac{\log(4/\alpha)}{4n}}, \quad b_F = \frac{1}{2} \sqrt{\frac{\log(4/\alpha)}{2n}}. \quad (77)$$

当 $n = 512$ 时，主目标与辅助目标的错误计数分别为 107 和 121。代入后得到第 7.4 节中的动作间隔与区间。这些界针对声明的生成器产生的独立样本，而非不同部署系统之间的变化。

附录 F SQLite 快照保留协议

每个新建数据库包含 64 行，每行载荷为 1024 字节，且具有相同的整数型当前事实。数据库使用 4096 字节页面及预写日志 (WAL) 模式。通过查询该事实建立读事务。三个实验组分别保留该事务、存储 10 字节事实记录后释放事务，以及先使用 SQLite 备份接口复制完整快照再释放事务。离线副本为 98304 字节。它用于验证完整快照确实可恢复，并非与事实记录等大的压缩表示。

独立写进程逐批替换全部载荷，每批都以完全同步的方式提交。实验关闭自动检查点，并在每次提交后显式尝试同一种截断检查点操作。忙等待超时设为零，只记录检查点是否受阻。对于保留活动读事务的那一组，每次尝试都返回忙状态，WAL 帧被保留下来。跑完 8、16、32 批后，检查点尝试之后 WAL 大小的峰值依次为 725152、1450272 与 2900512 字节；其余各组均为零。写进程一直保持连接，直到未来查询与释放检验全部结束，以免连接关闭触发一次未被记录的检查点。

批量写入完成后，评估器在写进程仍保持连接的情况下挑出一行，把它的原始载荷分别与保留快照或备份中的载荷、以及当前数据库中的载荷作比较。读出器既拿不到生成随机种子，也接触不到任何能提供原始载荷的额外信息源。只含事实的记录里没有载荷；通过常规结构化查询语言 (Structured Query Language, SQL) 接口查询当前数据库状态，返回的是被替换之后的内容。上述结论只针对所声明的记录形式与数据库接口，不涉及对旧字节的取证恢复能力。实验最后关闭活动事务，重新执行一次截断检查点，结果为成功，WAL 字节数归零。每种配置都用各自新建的临时数据库，连接与子进程均显式清理。

两个随机种子的字节计数完全相同，是由固定的模式、载荷长度和写入计划决定的，不能据此推断普遍的存储规律。这项小规模实验在实际存储引擎中验证了快照保留、检查点阻塞及解除阻塞三种行为。它并不构成额外的一组持久型 / 易失型动作反转，也没有把字节数折算成任何下游决策效用。

附录 G 计算验证

补充检验依次核对递归细化、决策遗憾、后验压缩与系统协议。

有限状态检验分别为 $L = 0, \dots, 8$ 构建显式转移系统，并用精确有理数运算核对根状态的动作价值与近碰撞距离。划分细化检验包括：根状态是否恰在第 $L + 1$ 轮分裂、其他状态是否更早分裂，以及有限停止界是否成立。对 $K = 1, 2, 3$ ，穷举所有编码器和控制器，以核对均衡根状态对上的深度—容量联合遗憾规则。这些检验仅覆盖有限实例，不能替代一般性证明。

保留预览的扩展直接依据有限联合分布计算贝叶斯风险，以独立核对阈值公式。精确有理数枚举覆盖三种预览错误率、三组持久型与易失型存续概率及五种相对持有成本，共 45 个实例。HTTP 实验采用事先规定的 512 对生成样本，预览错误率为 $1/4$ ，主分析成本为 0.06，完整成本网格见表 4。读出器仅可使用保留的预览和实际验证响应，评估器另行使用目标真值评分。持有成本在选择保留通路时计入，无论后续响应如何都需支付，不能通过放弃可选的最终查询来免除。