

论记忆量 · 已相当可观

记忆量很大 $\neq$ 更聪明

历史太多，反而 —

选错按钮

走过期流程

套不适用规则



从第一性原理看

记忆的价值，只有一个衡量标准

✗ 误区 ~~记忆 = 容量问题：存得越多~~ → 智能自然出现

↓ 翻转

✓ 价值的唯一标准 未来需要时，能不能——

①

找到正确的信息

Retrieve the right info

+

②

转化成正确行动

Turn it into the right action

📦 存了多少 = 次要

人类记忆分类 认知科学

系统映射

语言 Agent 对应 CoALA 映射



工作记忆

此刻正在用

上下文窗口



Context · Scratchpad



情节记忆

过去某次经历

具体经历



过去发生了什么



语义记忆

从多次抽象

抽象规律



通常什么成立



程序记忆

已经能执行

可执行技能



具体应该怎么做

≈ 非 =

借用映射 · 不是画等号

人类记忆分类 → 四类记忆

此刻

过去

规律

技能

①



工作记忆

此刻

现在手上是什么

定义

当前正在使用的信息

例子

正在填哪张表、上一步工具返回了什么

②



情节记忆

过去

过去发生了什么

定义

某次具体经历

例子

上周退款时网站报了什么错、用哪个入口成功

③



语义记忆

规律

通常什么规律成立

定义

多次经历抽象出的知识

例子

提交退款前核对订单状态 / 退款期限 / 最终金额

④



程序记忆

技能

具体应该怎样做

定义

已经能够执行的技能

例子

定位订单 → 检查资格 → 填写原因 → 核对金额 → 验证状态变化



 向量相似度：只对“看起来像”负责

找到相似文本

向量相似度 · 只是“看起来像”

≠

提炼有效规律

抽象 + 验证 · 真正能指导行动

保存经历

> 形成知识



② 失败反馈 → 一段文字反思 → 放回上下文

×

任务失败信号



✎ 文字反思便签

"我过早提交了表单，没检查金额"



↺

放回上下文
提醒改变下次行为



被忽视的真因

🔒 权限不足

⚠️ 危险假设：反思必须正确

🔒 真因 = 权限不足

Agent 却错记成



👉 “多点一次提交按钮”

写入长期记忆



🧠 错误规则进长期记忆

↳ 越用越自信地错

⚖️ 同一次踩坑的两种记法：情节记忆 vs 语义经验

情节记忆



“上次在某个页面点错了第二个按钮”

📍 只对那一页有效

◆ 具体但狭窄 · 记下来就行

VS

语义经验



“执行不可逆操作前，先读页面状态，再二次确认”

📍 可能跨网站 / 用户迁移

◆ 抽象但需验证 · 必须经过验证

🔍 向量库返回的「第一条相似记忆」——

能直接拿来当 Agent 的行动指令吗？

能 ✓

弹幕扣 111

不能 ✕

弹幕扣 000

⌚ 扣完别走，答案就在后面几分钟里 🗣️

🔄 记忆生命周期 · 四个动作（上）写入 · 巩固



②

巩固

CONSOLIDATE · 新经历 × 已有记忆

与已有记忆比较：

重复

补充

冲突

替代

10 → 1

十条相似 → 压缩成
一条更稳定的规律



证据来源



适用范围

不永远保留十条原始记录 —— 压缩，但保留可追溯性



🔄 记忆生命周期 · 四个动作（下）检索 · 遗忘



④ 遗忘

FORGET · 记忆也要能退役

现实世界会变化 —— 记忆必须能**退役**

 **过期**
到期不再采信

 **降权**
调低置信度

 **替代**
被更新记录取代

 **撤销**
证伪后作废

永久保存 $\neq$ **目标** —— 知道何时不再相信一条经验，才是适应动态世界的能力

每条长期记忆 = 一张带元数据的 记忆卡

何时创建 · 何时复查 · 与哪条旧记忆冲突

memory_card.json · mem#124

长期记忆 · long-term

01 内容
content

: 提交退款前需核对最终金额

02 来源任务
source

: task#417 · 退款流程

03 验证者
verified_by

: 后台状态校验 ✓

04 适用范围
scope

: 用户=企业版 / 网站=refund.x / 工具 v3

05 置信度
confidence

: 0.82

06 创建时间
created

: 2026-07-10

07 复查时间
review_at

: 2026-10-10

08 状态
status

: 待复核 ⚠

09 冲突
conflicts_with

: mem#88 ("管理员才能提交")

⚡ 待消解 → S12

解决条件差异，两个矛盾句子 → 更精确的规则

旧规则 · 管理员提交 ⚡ 新轨迹 · 员工也成功 未解决

系统需要检查 · 解决条件差异

① 权限政策变了吗？

管理员限制被放宽了？

② 两类订单适用不同规则？

企业版 vs 标准版，条件不同？

③ 成功经后台状态验证？

那次提交真的入库成功了吗？

✓ 更精确的规则

企业版订单 → 管理员提交

标准版订单 → 员工可提交 (经后台验证)



检索 → 行动 · 必须过闸，不能直连



检索命中

向量库最相似

× 不直接执行



候选证据

只是证据 · 非命令

→ 与环境权衡



关卡 · 必须通过

① 与当前环境匹配

② 高风险：显示来源 / 独立验证

✓ 通过验证



行动

已验证执行



否则：向量库第一条相似文本 = 未经审查的「最高规则」

🔗 记忆系统体检 · 只看 ~~召回率~~ ✖ 不够 → 至少看 四项指标

① 相关性
Relevance



取回即相关的占比

高 ↑ 越高越好

取回经验与当前任务是否相关

② 过时调用比例
Stale-Call Rate



过时经验被误调用

低 ↓ 越低越好

过时经验被调用的占比

③ 成功率改善
Success Lift



▲ 较无记忆 +34%

高 ↑ 越高越好

记忆是否改善最终任务成功率

④ 成本可控
Cost Bounded



历史增长仍平稳

稳 → 保持可控

随历史增长检索成本/占用是否可控

🎯 记忆服务的目标是 未来行动，不是在搜索测试里找到一段 相似文本



📁 只存不管 · 落灰静止

⚙️ 持续治理 · 有进有出

记忆不是 Agent 外接的一块硬盘，而是一套持续运转的信息治理机制



本集 · 第一性结论

原始经历 $\neq$ 知识



只有经过

① 验证

② 选择

③ 压缩

④ 检索

⑤ 遗忘

经历才可能成为未来能力的一部分



带着这套框架，我们继续 →

即使 Agent 总结出「以后该怎么做」，
在真实环境里 仍然可能失败.....

▶ 下一集见分晓



经验笔记

软 · 会飘 · 以后该怎么做

可执行技能

坚实 · 可组合 · 即插即用

▶ 下集预告

EP03

持续学习系列 · 第 3 集

把「做过」变成「会做」

经验笔记 → 可执行技能

🙌 感谢看到这里

👍 点赞

🌐 投币

💬 到评论区聊聊你的例子

你踩过最坑的一条「错误记忆」是哪条？

👉 我会一条条回

▶ 我们下集见

版权声明

唐国梁 Tommy 原创出品 | 熵智未来（深圳）科技有限公司版权所有。保留所有权利。

本课程及配套资料，包括但不限于课程视频、PDF 文档、课件、代码、图表、示例项目、知识库文档及其他学习材料，仅供购买本课程的用户本人学习使用。未经版权所有者书面许可，任何个人或机构不得以任何形式对本课程及配套资料进行录制、复制、传播、上传、公开分享、分发、转售、改编、翻译、汇编，或用于商业培训、企业内训、商业化产品、知识库建设、数据集构建、人工智能模型训练、微调及其他商业用途。如发现侵犯本课程及配套资料相关权益的行为，权利人有权依法要求停止侵害、删除下架、赔偿损失，并追究相应责任。

本课程及配套资料中涉及的第三方开源软件、模型、代码、数据集、商标、图片及其他素材，仅在相关许可证允许范围内进行教学展示，其相关权利归原权利人所有。

侵权举报 / 授权合作: legal@tgltommyai.com

微信公众号: 唐国梁TGLTommy | 官方网站: TGLTommy.com (国内访问需科学上网)