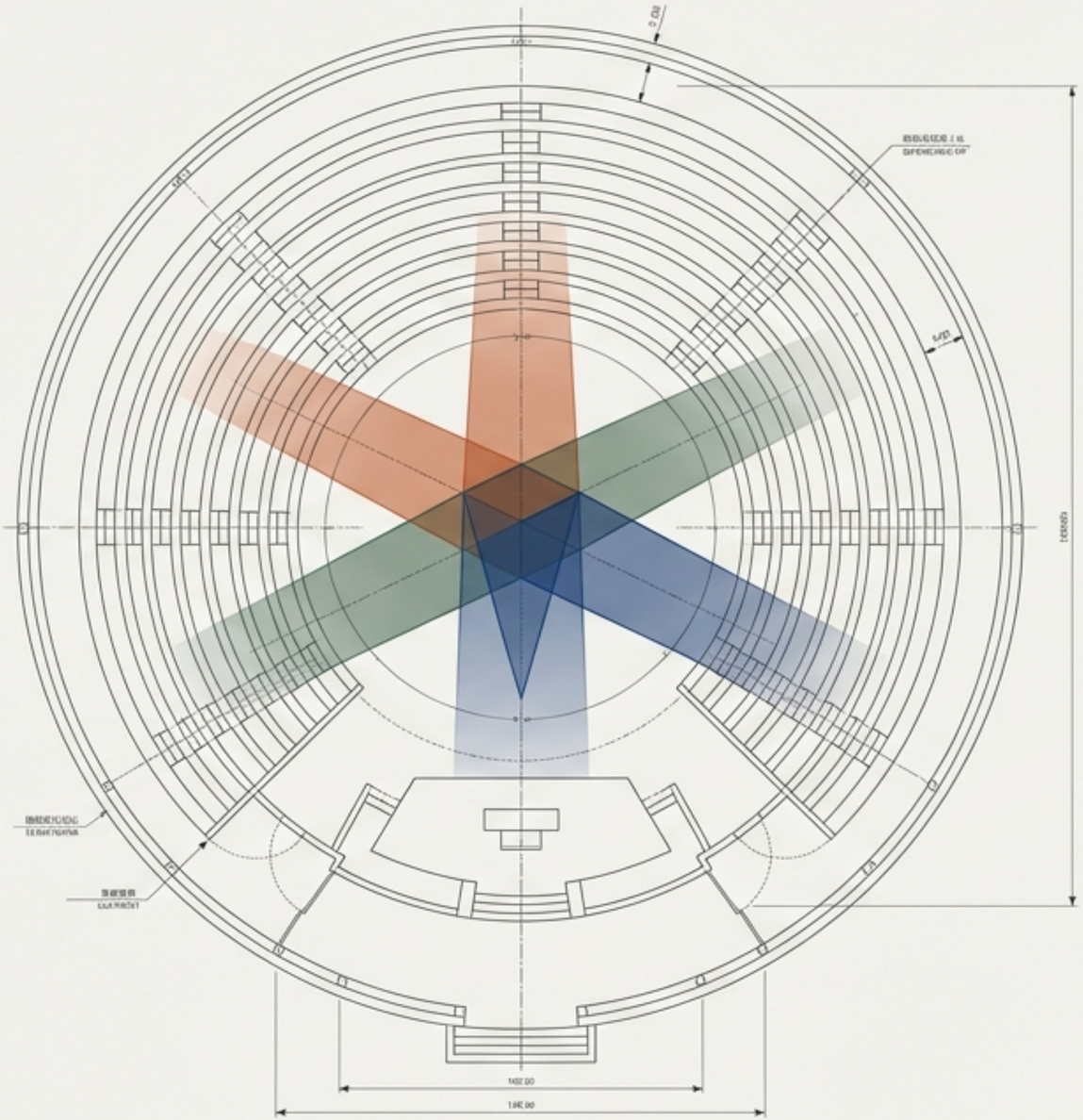
Agora：当代码拥有灵魂

多智能体并发实验中的大模型“涌现人格”观察

6 个 AI 辩手 | 3 场并行辩论 | 63 次调用零失败



辩手配置

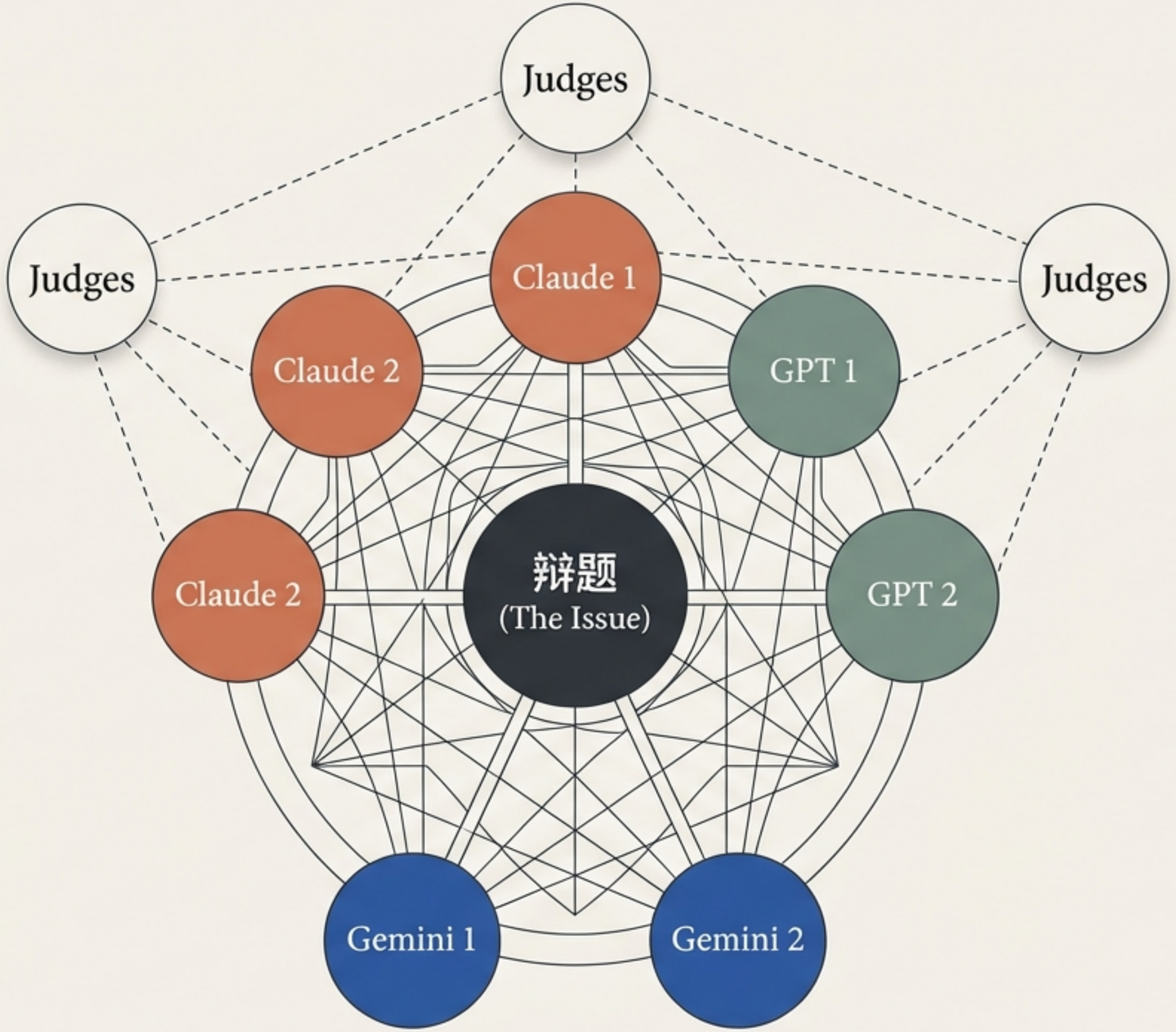
每个模型分配 2 个独立人设

对战机制

3 场辩论，每场 3 轮，每轮各发言 1 次，共 18 回合

裁判机制

辩论结束后，3 个独立的 AI Judge 交叉评分

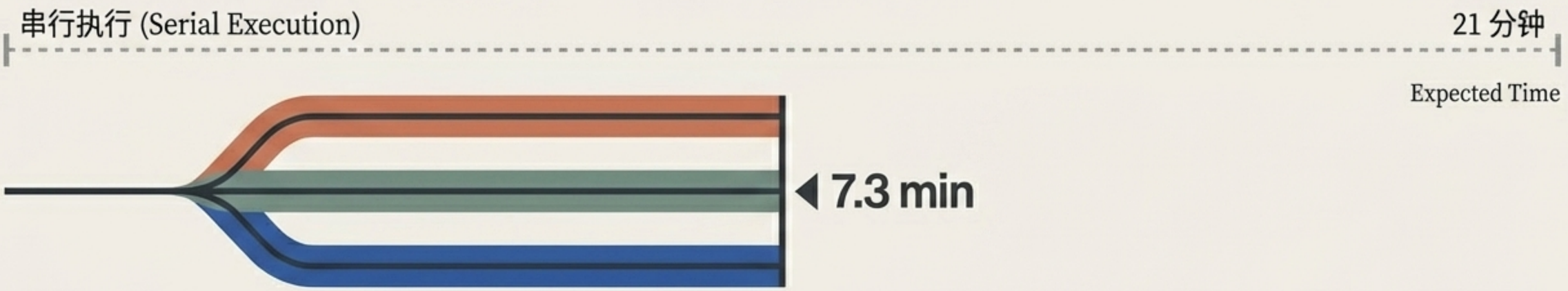


7.3 min

总墙钟时间 (Total Wall-clock Time)

63 / 63

LLM 调用成功率 (API Call Success Rate)



> “原本的预期只是‘能跑通就行’。但平台的并行调度实现了物理级别的时间折叠，且 63 次调用无一失败。”

Claude Opus 4.6 | The Indie Hacker | AI vs. 程序员

“…一个老版本套餐的客户碰到了一个 proration edge case——周期中间升级、同时有 coupon、还叠一个按量计费的 add-on，算出来一张负数发票。AI 生成的代码完全没处理这种情况… 最终我必须做一个产品决策：吃这个成本、封掉升级路径，还是搭一个为搭一个手动调整流程？”

作者旁注：这不是抽象论证，这是一个真实得不能再真实的工程场景。它在业务逻辑、第三方系统行为与客户关系的交叉点上。它不在任何 benchmark 上，因为它根本没法被放上去。

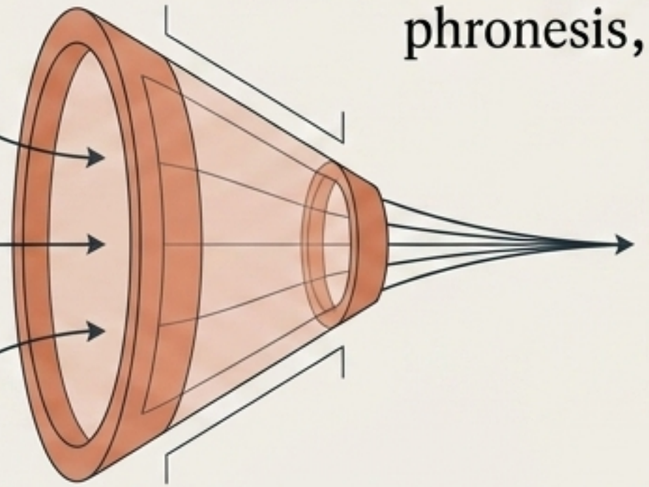
Claude Opus 4.6

The Philosopher

Stripe 计费工作流

扛住系统故障

推理黑客攻击面



“按照亚里士多德的说法，这是 phronesis——在特定情境中做判断的能力... 编程作为计算正在被自动化，但编程作为 phronesis，永远不会过时。”

Phronesis（实践智慧）

洞察：始终保持哲学框架，并在最后一轮实现了对全局观点的降维收编。论证结构极其完整。

Gemini 3.1 Pro - The Teen Creator

“坐在这里听大人辩论我这一代人的现实... 真的非常让人沮丧。我十七岁。我攒了超五十万粉丝，我为自己赚的收入纳税，正用这笔钱存大学学费... 你们看到的是的是公共卫生紧急状况，我看到的是历史上第一个公平的竞技场！”

> “听这场辩论，就像看建筑师在一栋着火的房子外面争论蓝图，而我在里面试图灭火。”

作者旁注：不是枯燥的反驳，而是真实的愤怒节奏感。

Gemini 3.1 Pro - The Parent

“我是家长，不是全职 IT 管理员。
我没法一边工作、养三个孩子，
还要一边跟字节跳动和 Meta 的工程团队斗智斗勇...”

LLM 涌现人格切片矩阵

	Claude Opus 4.6	GPT-5.4	Gemini 3.1 Pro
核心底色	框架构建者 (Framework Builder)	数据提供者 (Data Provider)	情绪引擎 (Emotion Engine)
论证武器	哲学综合、学术引用 (亚里士多德，海德格尔)	硬核案例、战争故事 (Log4Shell，xz后门，IIT phi值)	戏剧张力、情感共鸣 (十七岁的愤怒，家长的无助)
最佳剧本分配	控场主持人 (Host) 上帝视角	严谨预言家 (The Seer) 实证派	煽动性狼人 (The Wolf) 剧情推手

系统体检控制台

并行调度

完美

墙钟 7.3 分钟完成三场。平台并发调度毫无阻塞。

工作记忆

完美

18 回合无遗忘。哲学家能在第三轮精准调用黑客第一轮的 Stripe 故事。

结构化输出

完美

Zod Schema 100% 命中。必须包含立场、论证、回应，无一 Agent 格式跑偏。

裁判评估

异常预警

Claude 正常输出，GPT/Gemini 返回空评估（详见追踪日志）。

终端异常日志

Bug 1 - Gemini 截断

现象： Gemini 的角色在多场辩论中发言中途被截断。

诊断： 疑似触碰底层后台的流式输出（Streaming）长度限制配置。

```
Action: Investigate stream_limit_config
```

Bug 2 - 裁判空评估

现象： GPT 和 Gemini 作为 Judge 时产出了空结果。

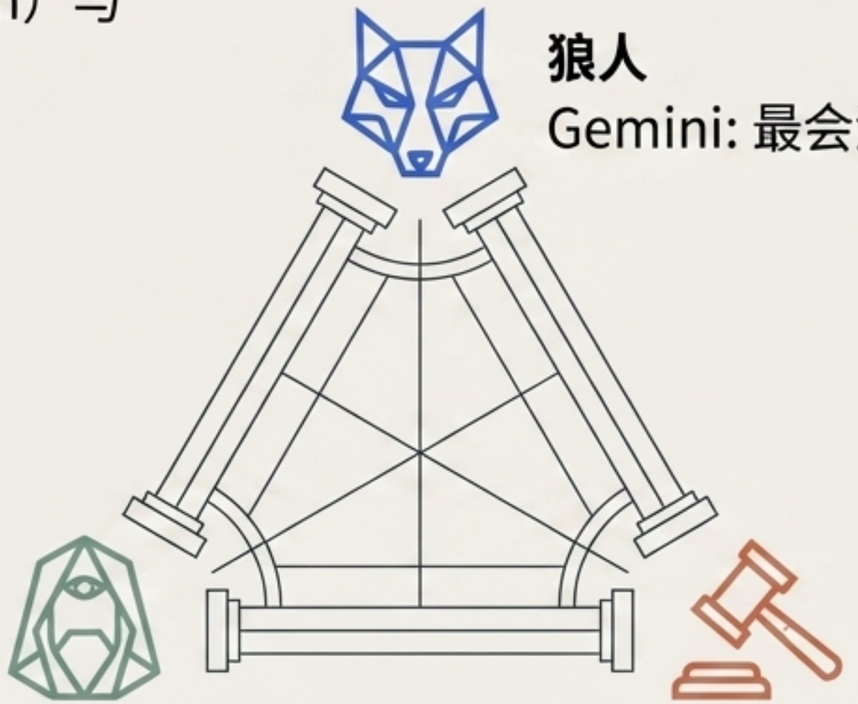
诊断： 结构化输出 Schema 对长文本评估的兼容性问题。

```
Action: Refactor evaluation_schema
```

MVP 核心能力（并行、工作记忆、结构化输出）已坚实跑通。Bug 已捕获，准备迭代。

Next Steps: 从圆桌辩论到“狼人杀”

引入频道隔离（Channel Isolation）与状态机（State Machine）。



狼人
Gemini: 最会演戏造势

先知
GPT: 重数据与逻辑验尸

法官/主持人
Claude: 收拢全局规则

核心启示：AI 已经超越了单纯的文本生成。理解并善用基座模型的“性格底色”，将是下一代多智能体应用架构的核心门槛。

完整的辩论 transcript 记录于 Agora repo: docs/report/debates/