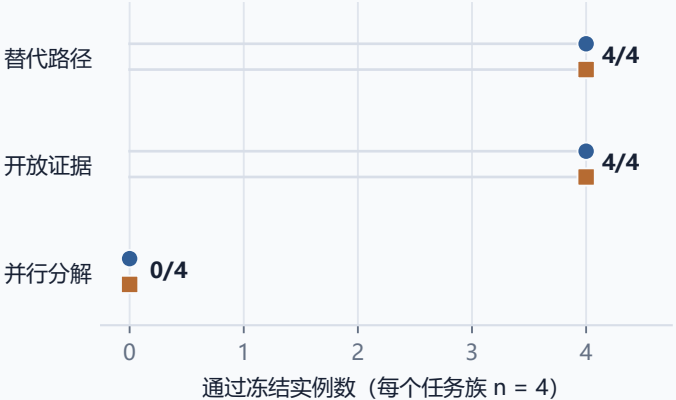


Pilot B: 成功、质量与执行成本

3 个任务族 × 4 个冻结实例 × 2 种架构 · n = 24 配对正式运行

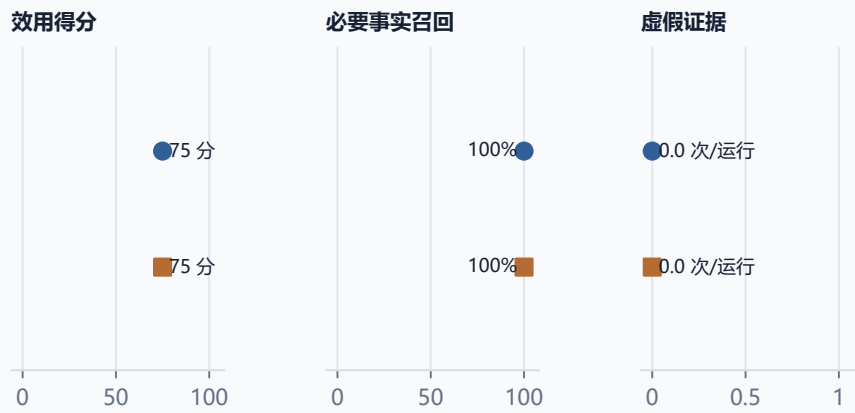
Contracted ReAct Goal-Graph Hybrid

A | 任务族成功结果 (配对)



B | 质量指标 (独立单位)

来源: reviewed pilot-b-summary.json



C | 平均执行成本比值

