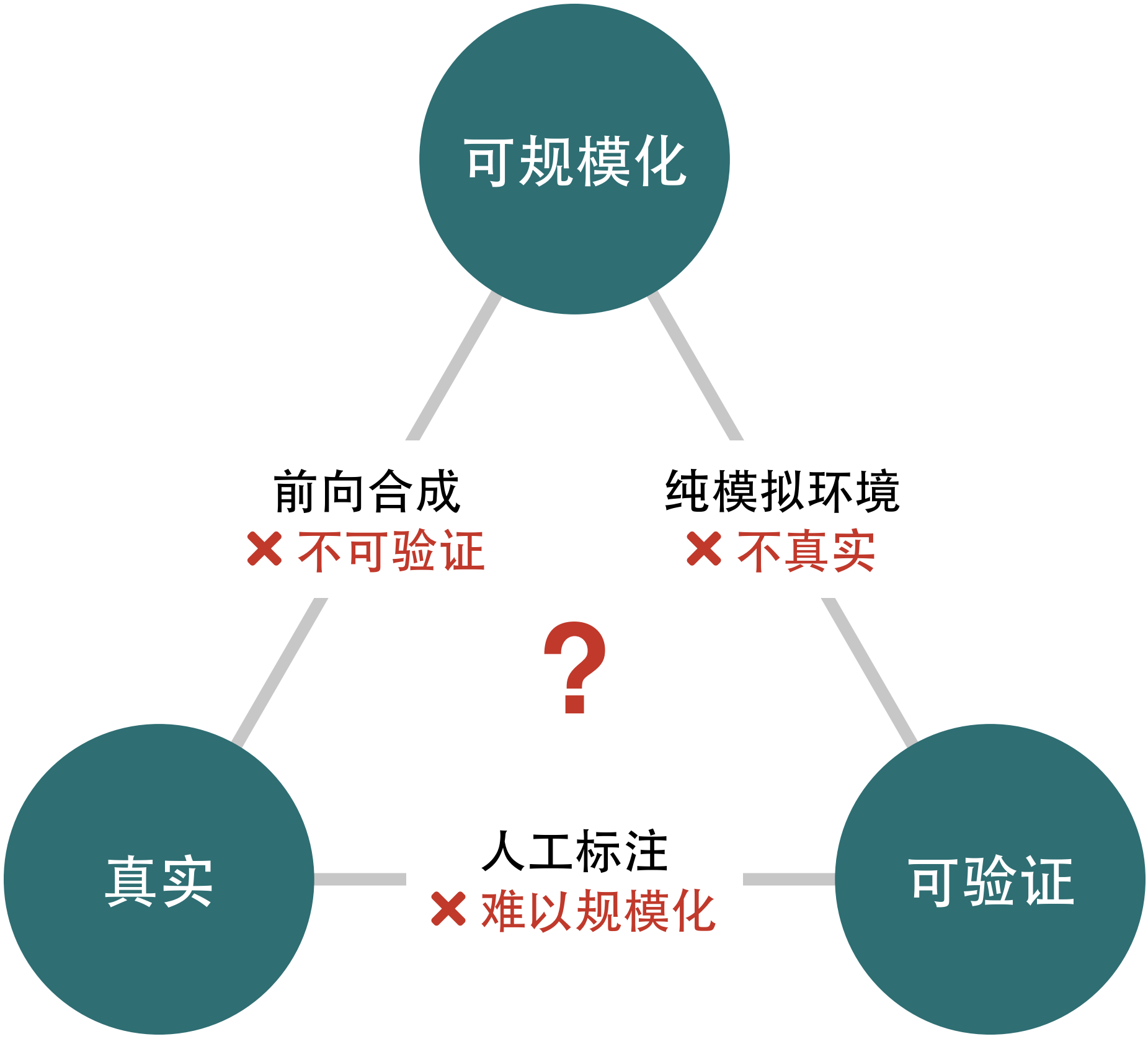
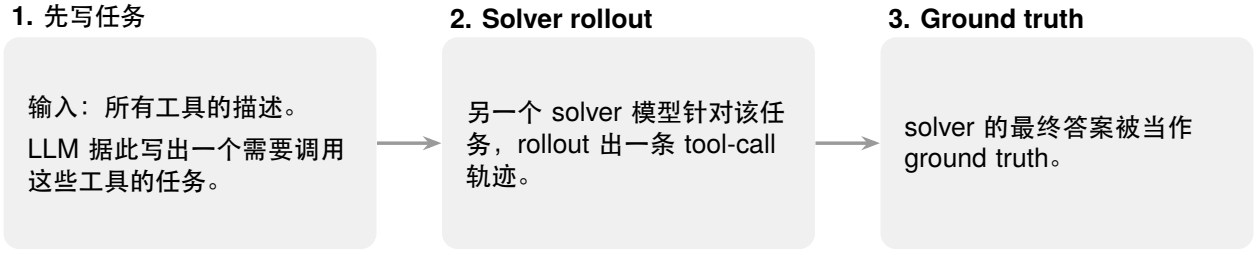
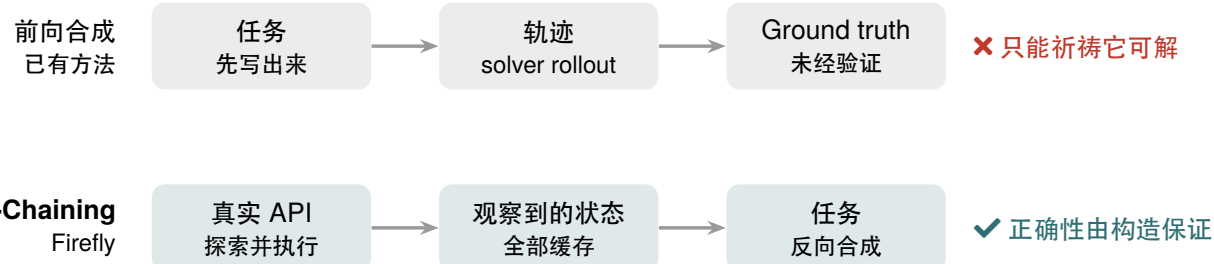
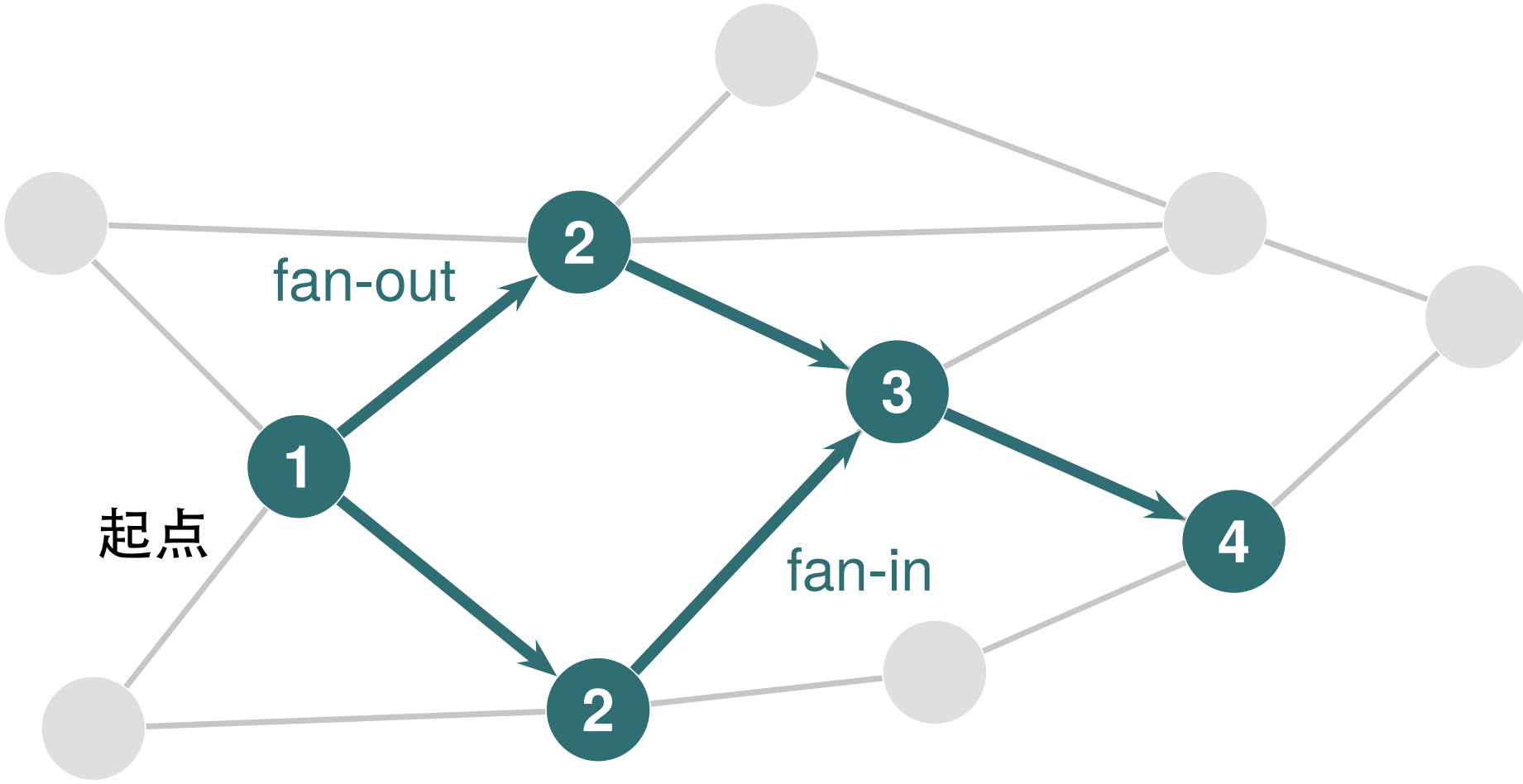

可规模化
前向合成
✕ 不可验证
纯模拟环境
✕ 不真实
?
真实
人工标注
✕ 难以规模化
可验证


1. 先写任务
输入：所有工具的描述。
LLM 据此写出一个需要调用这些工具的任务。
2. Solver rollout
另一个 solver 模型针对该任务，rollout 出一条 tool-call 轨迹。
3. Ground truth
solver 的最终答案被当作 ground truth。
✖ 任务本身可能无法完成：目标状态不可达。
✖ solver 可能做错：ground truth 随之出错。


前向合成
已有方法
任务
先写出来
轨迹
solver rollout
Ground truth
未经验证
只能祈祷它可解
Back-Chaining
Firefly
真实 API
探索并执行
观察到的状态
全部缓存
任务
反向合成
正确性由构造保证


1. 收集 MCP Server
筛选真实 server
2. 任务合成
探索 + 反向出题
3. 强化学习
在离线 Simulator 中 RL




灰色：Tool Graph（工具之间的可串联关系）
青色：一次探索得到的 tool-call DAG



**Server 会下线**

`search("hotels")`

→ `503 Unavailable`



结果不可复现

`get_weather("Boston")`

周一 → `18°C, sunny`
周二 → `9°C, rain`



本身就是随机的

`fortune_telling("Leo")`

第 1 次 → `"great luck"`
第 2 次 → `"beware Tuesday"`

探索阶段缓存的调用：`search(query="hotels", num=5, format="csv")`

## Exact Match

`query="hotels"` ✓
`num=5` ✓
`format="csv"` ✓

直接返回缓存的
真实响应

## Fuzzy Match

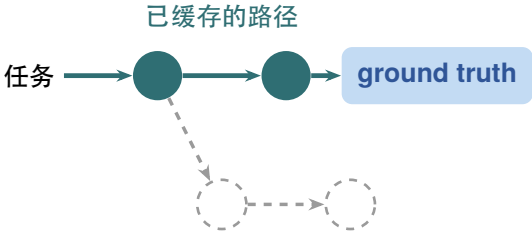
`query="hotels"` ✓
`num=5` ✓
`format="json"` ✗

LLM 基于缓存结果模拟：
把它改写成 JSON

## No Data

探索阶段
从未调用过该工具

返回错误




✔ 做对的 **Agent**：始终在路径上，每次调用都命中缓存，拿到真实响应，走到 ground truth。

✘ 做错的 **Agent**：偏离路径，可能拿不到数据。但它本来就错了，reward 反正是 0。




Firefly 测试集 pass@1（%），每题 16 次 rollout