# 智能体系统总体设计方案（Agent = 模型 × 环境）

## 1. 最终目标

输入一份新的病人病例数据，系统能够完成从检查流程编排、意图识别、执行推进、报告生成到患者沟通的一次完整诊断流程。整个流程对应一个完整的 agent 调用（episode），产出包括报告（专业版+患者版）、影像说明和全程轨迹。

## 2. 系统组成：Environment × Model

### 2.1 Environment 组件

• State Store：保存病例状态、已执行节点、上下文变量和对话历史。  
• Rule Engine / Node Graph：定义节点依赖、回退策略、必填字段等规则。  
• Executor：读取规范 JSON 命令，执行节点动作或回退，更新 state。  
• Validator：执行一致性和合规性校验，输出错误与警告。  
• Observation Builder：根据当前状态生成输入模型的观测 JSON。  
• Output Enforcer：确保模型输出结构化且符合 JSON Schema。  
• Dialogue Manager：管理医患对话的轮次和话题边界。  
• Reward Aggregator：计算每步和终局的奖励，用于RL。  
• Logger/Tracer：记录 plan/intent/action/reward/validation 等日志。

### 2.2 Model（Qwen-0.5B 多头结构）

模型使用共享骨干，拥有四个任务头：planner\_head、intent\_head、report\_head、comms\_head。当前任务由环境通过 stage 信号和控制标记指示。

## 3. 接口契约

### 3.1 观测输入（Environment → Model）

{episode\_id, stage, step\_index, executed\_nodes, observation\_summary, plan\_hint, user\_utterance, final\_state, images, allowed\_actions}

### 3.2 模型输出（Model → Environment）

• Planner 输出 next\_node 与 issues。  
• Intent 输出4类：confirm\_next, skip\_next, rollback, clarify。  
• Report 输出专业版报告文本（含两张图文字说明）。  
• Comms 输出患者版解释或医患问答答案。

## 4. 一次完整诊断 Episode 流程

1. Planner 阶段：生成 plan。  
2. Intent 阶段：解析用户自然语言为结构化意图 JSON。  
3. Executor 执行：更新 state。  
4. Report 阶段：生成专业版报告。  
5. Comms 阶段：生成患者版解释与问答。  
6. Episode 结束，输出报告、影像说明和轨迹。

## 5. final\_state 最小字段

exam: device\_model, probe\_freq\_mhz  
gland: size\_mm(l/w/t), echogenicity, diffuse\_change  
nodule\_count, nodules[id, site, size\_mm, features, tirads]  
overall\_tirads, impression, recommendation, images

## 6. 强化学习任务与奖励

• Planner：奖励拓扑正确、减少违规与冗余、步数惩罚、终局报告成功大奖励。  
• Intent：奖励与金意图一致、执行后推进、JSON 合规。  
• Report：奖励结构完整、事实一致、合规语言、验证通过。  
• Comms：奖励行动可执行、避免术语、事实一致、安全。

## 7. 训练流程

阶段A：离线偏好/规则强化（DPO/ORPO/GRPO）。  
阶段B：小规模在线交互（端到端episode）。  
阶段C：保形蒸馏稳定策略。

## 8. 评测与验收指标

端到端闭环成功率、平均步数、Planner 一致率、Intent 准确率、Report 验证通过率、Comms 可理解度与安全评分。

## 9. MVP 里程碑

v0.1 协议冻结：确定 schema 与奖励清单，构造10条示例episode。  
v0.2 环境最小化：Executor、Validator、FSM、Enforcer。  
v0.3 训练起跑：离线RL→端到端RL，形成闭环诊断样例。