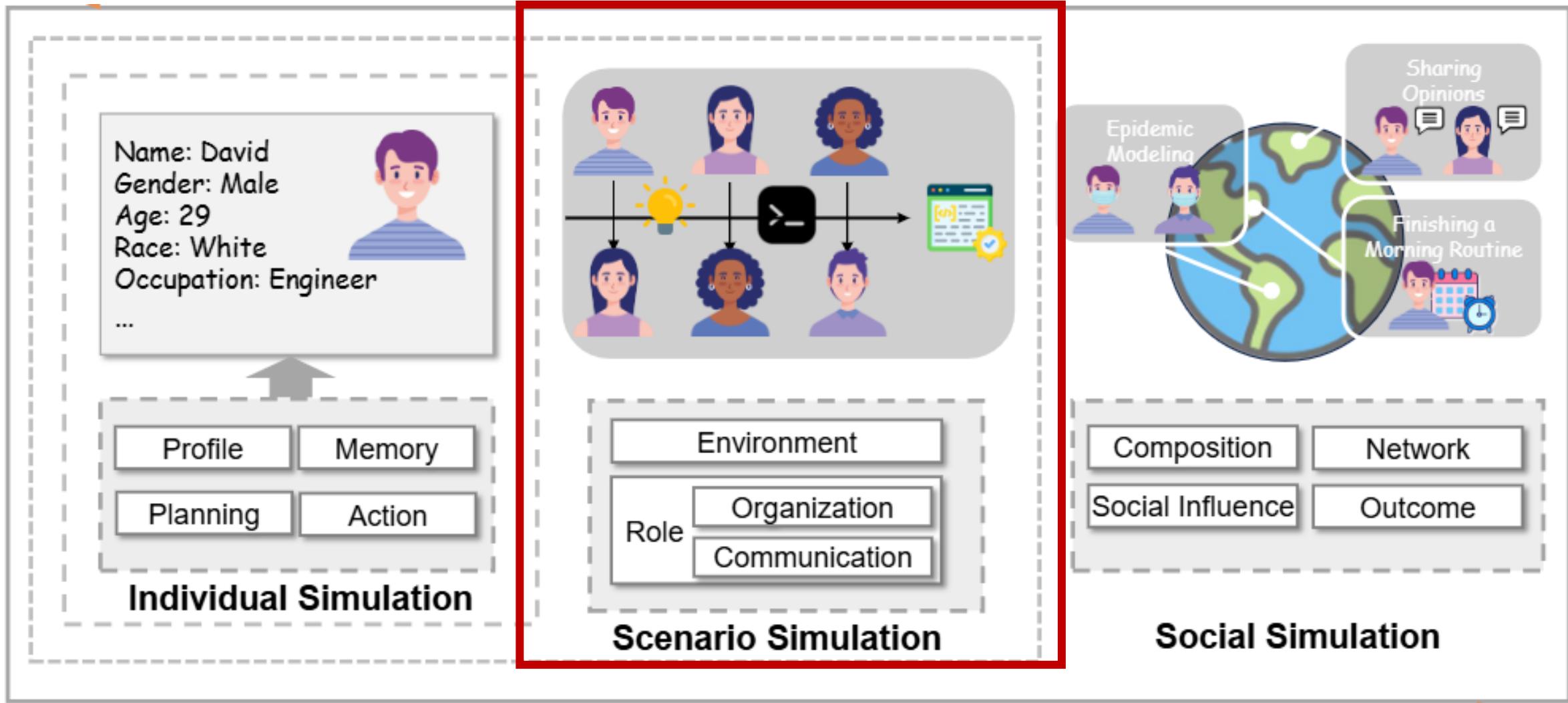


智能体驱动的社会行为研究



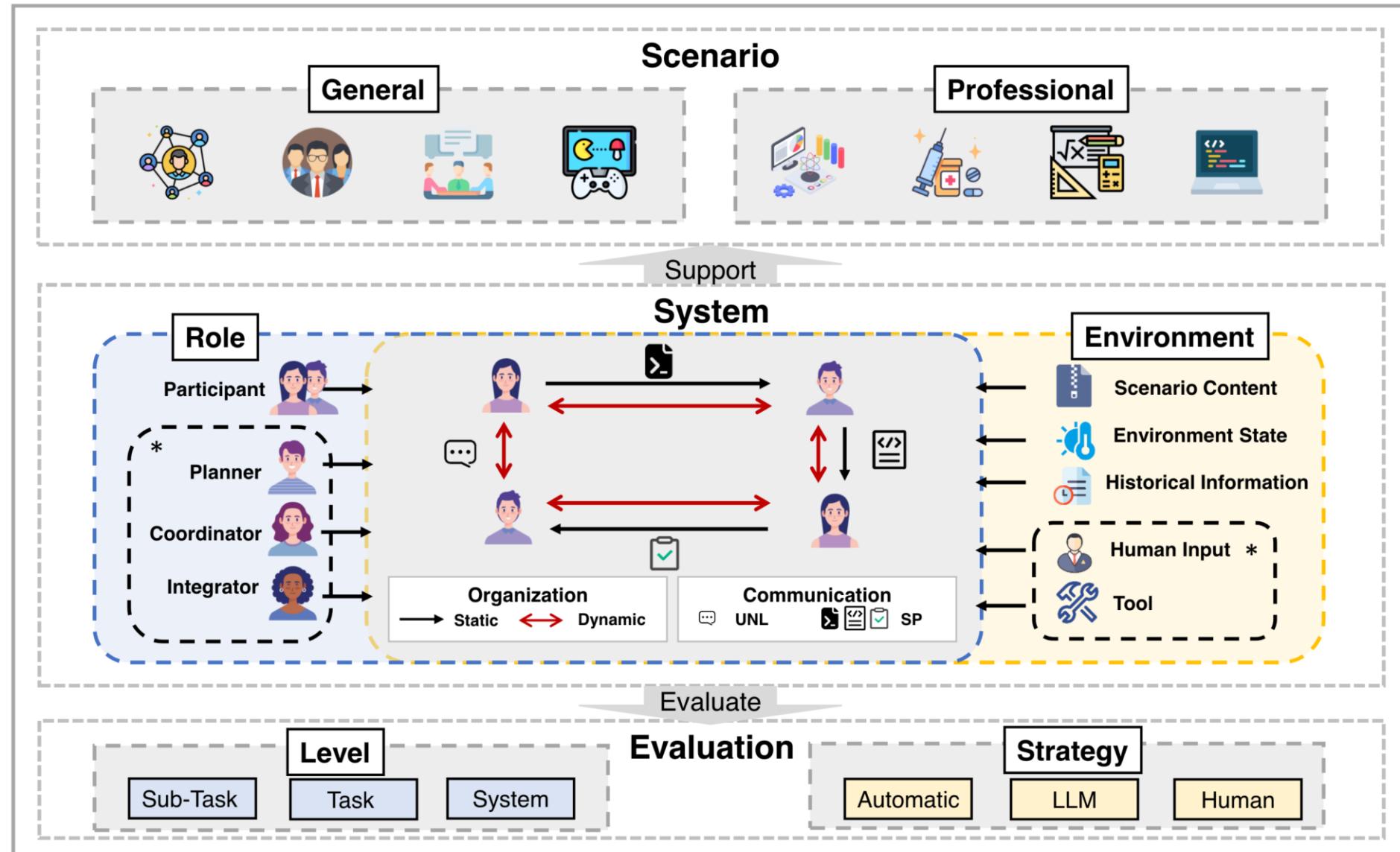
场景模拟



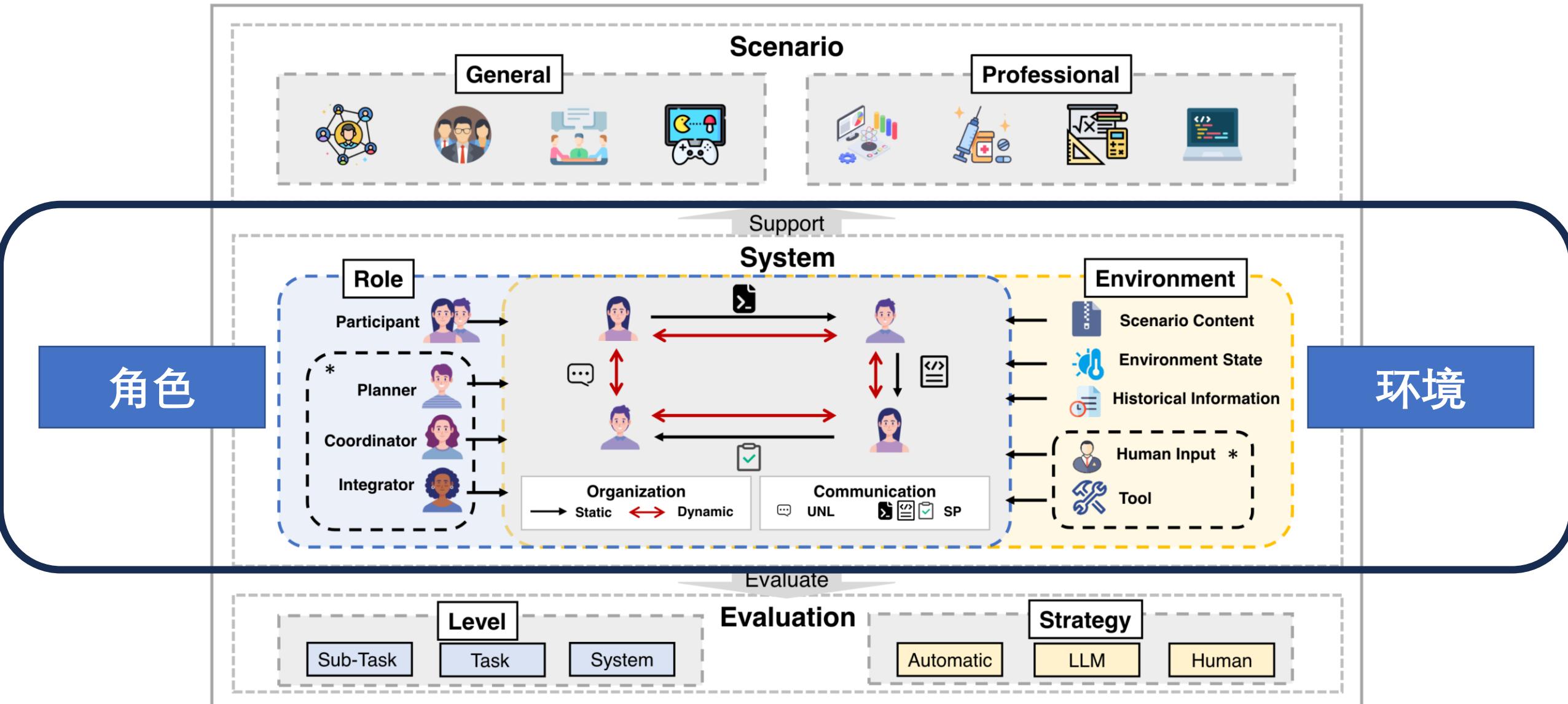
- 个体之间并非独立存在，个体与个体、环境交流形成场景
- 任务驱动：个体参与协作，完成场景中的特定任务
- 社交驱动：个体参与交互，实现交互中的社交目标

- 个体之间并非独立存在，个体与个体、环境交流形成场景
- 任务驱动：个体参与协作，完成场景中的特定任务
- 社交驱动：个体参与交互，实现交互中的社交目标
- 大模型智能体能否像人类一样协作，完成某个具体场景中的任务/目标，实现群体智能？
 - 正确回答某个问题
 - 完成某个产品开发
 - 完成某个案件审判
 - ...

场景模拟-整体框架



场景模拟-整体框架





场景模拟的研究要点

- 如何构建场景模拟系统? (组成要素)

模拟场景选择

场景模拟构建

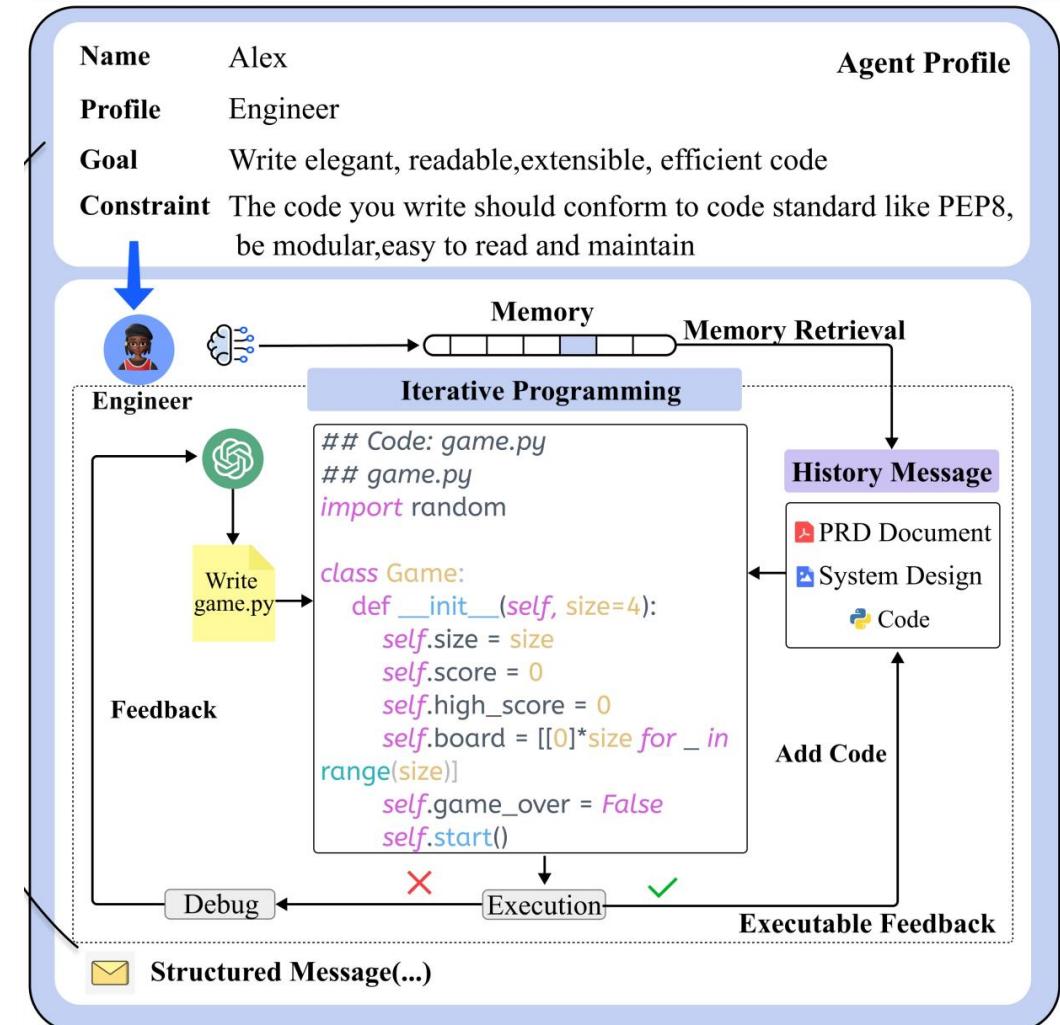


场景模拟的环境组成

- 场景设定：场景的基本信息，初始化智能体的目标
- 环境状态：场景执行过程中环境提供的信息，影响智能体的决策行为
- 历史信息：过去与任务相关的记录，确保智能体的行为一致性
- 外部工具：提供与场景模拟任务相关的专业功能

环境-场景设定

- 事件 (Event): 需要解决的主要任务、目标、焦点、背景描述
 - 法律案件
 - 讨论主题
- 配置 (Profile): 智能体相关的个性化信息(特定场景)
 - 兴趣, 目标, 角色



环境-环境状态



- 观察 (Observe): (主动获取) 环境和智能体当前状态的变化
 - 环境状态
 - 智能体状态
- 反馈 (Feedback) : (被动接收) 智能体在执行操作后收到的响应

■ 环境反馈

Task: Your task is to melt ice cream.; **Time:** 14; **Score:** 35; **Action history:** <extra_id_0> Action 5 (+5): open fridge --> You opened fridge. In it, you see an ice cream... [...] <extra_id_9> Action 14 (+0): move metal pot to stove --> You move the metal pot to the stove. </s> **Current environment:** This room is kitchen. You see: a fridge (closed) | a sink | an oven (closed, turned off) | a stove (turned off; on it: a metal pot containing ice cream ...). | [...] | **Inventory:** an orange, ... </s> **Visited:** workshop, hallway, kitchen </s> **What should be the next action?**

 **Swift:** (T5-large w/ imitation learning)

→ **Next action:**
activate stove



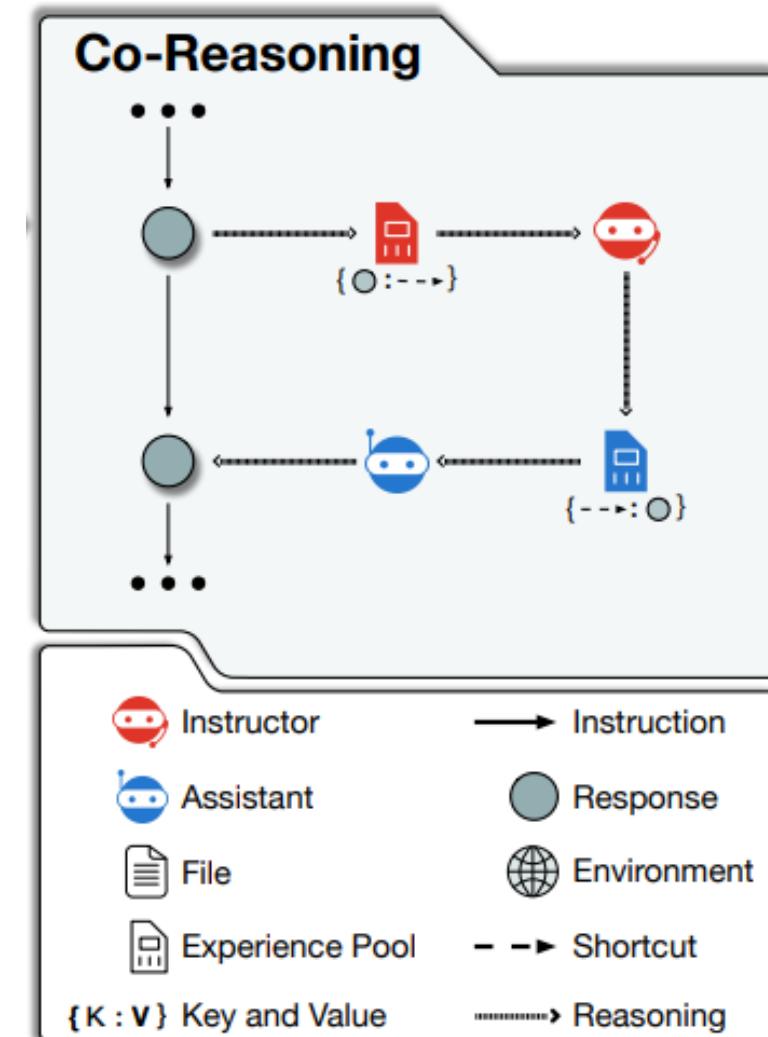
→ **Obs. 15:** The stove appears broken and can't be activated.

✗ **Time to switch**

 **Sage:** Prompting LLMs (GPT-4) for **Planning** and **Grounding** the Next Subgoals

环境-历史信息

- 直接拼接 (Direct Integration): 将原始历史数据拼接到当前输入
- 精炼 (Refinement): 根据历史数据迭代更新和增强响应
- 摘要 (Summarization): 从历史数据中提炼出重要的见解
- 记忆机制 (Memory Mechanisms): 将历史信息集成到智能体的记忆模块



环境-外部工具

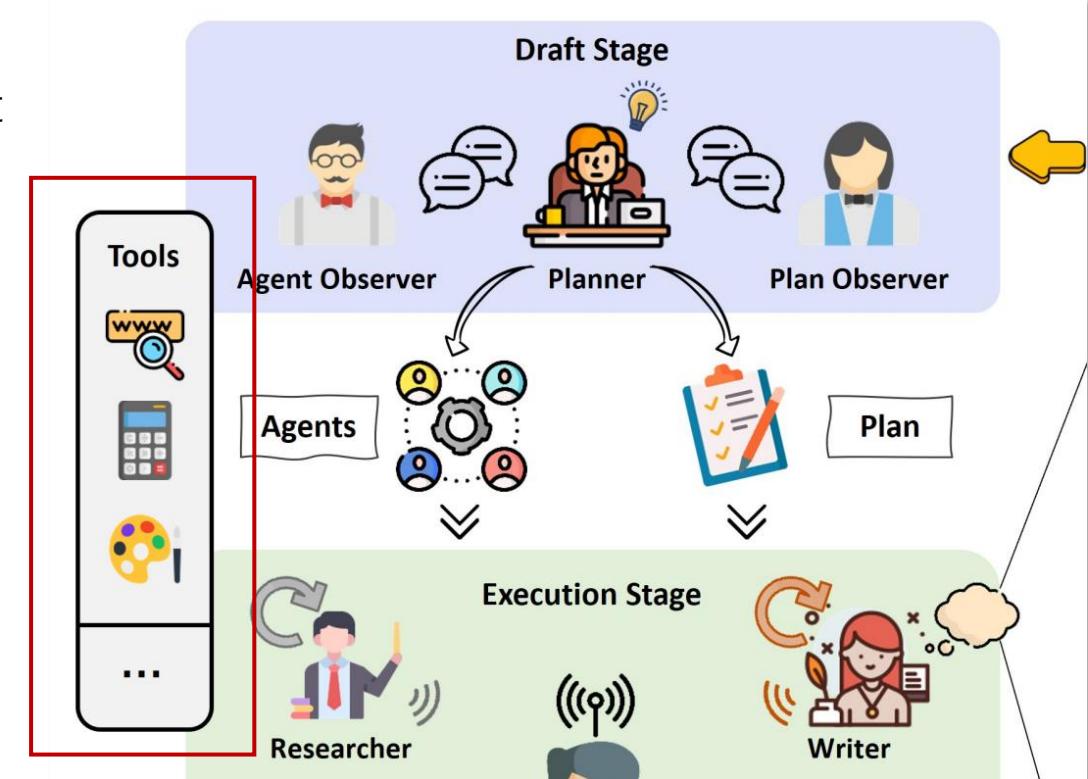


■ 编程语言 (Programming Languages)

- Python, SQL (ChatDev [Qian et al., 2024b], MetaGPT [Hong et al., 2023], AutoTQA [Zhu et al., 2023])

■ 任务相关工具 (Task-related Tools)

- 计算器, 预定义工具和 API (AutoAgents [Chen et al., 2023a], OpenAgents [Xie et al., 2023b])



AutoAgent [Chen et al., 2023]

场景模拟的角色组成

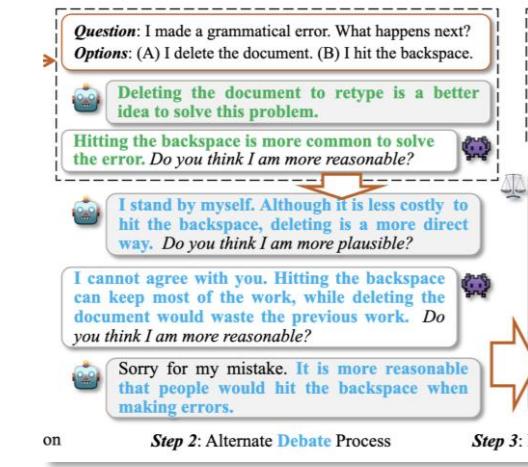


- 参与者 (Participant) : 场景模拟的关键成员，实际参与任务的角色
- 计划者 (Planner) : 定义目标, 分析用户需求和优化执行计划
- 协调者 (Coordinators) : 负责管理和协调智能体之间的协作
- 集成者 (Integrator) : 最终决策和总结
- 人类 (Human) : 人类在环路

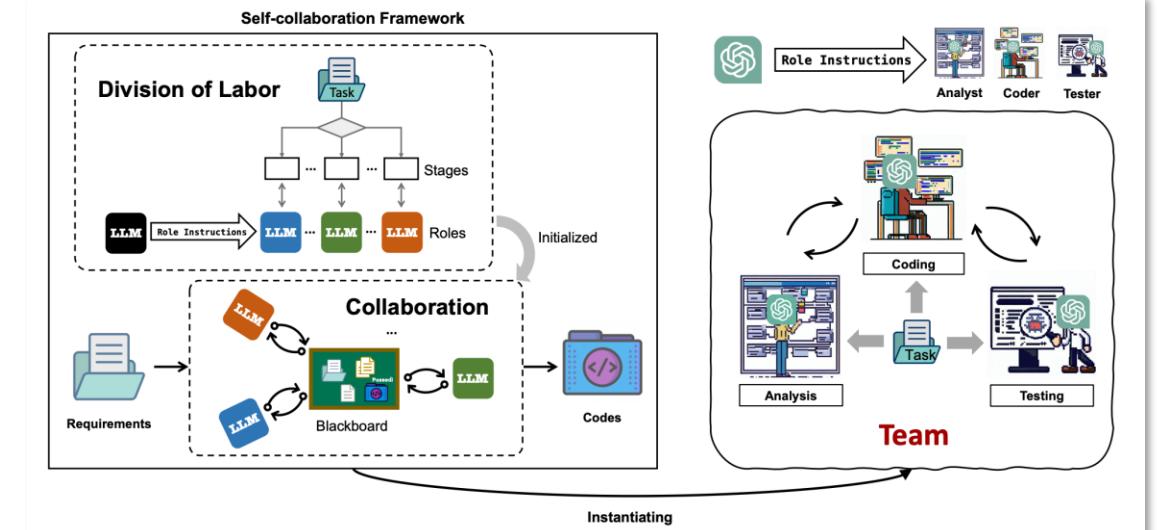
角色-参与者

- 对话导向 (Dialogue-oriented)
 - 信息交换, 反馈和任务指导
 - 问答, 游戏 (Blind judgement [Hamilton, 2023], Dera [Nair et al., 2023], FORD [Xiong et al., 2023])

- 执行导向 (Execution-oriented)
 - 任务执行和操作
 - 软件开发, 医学 (Self-collaboration [Dong et al., 2023], ICL-AIF [Fu et al., 2023], DR-CoT [Wu et al., 2023a])



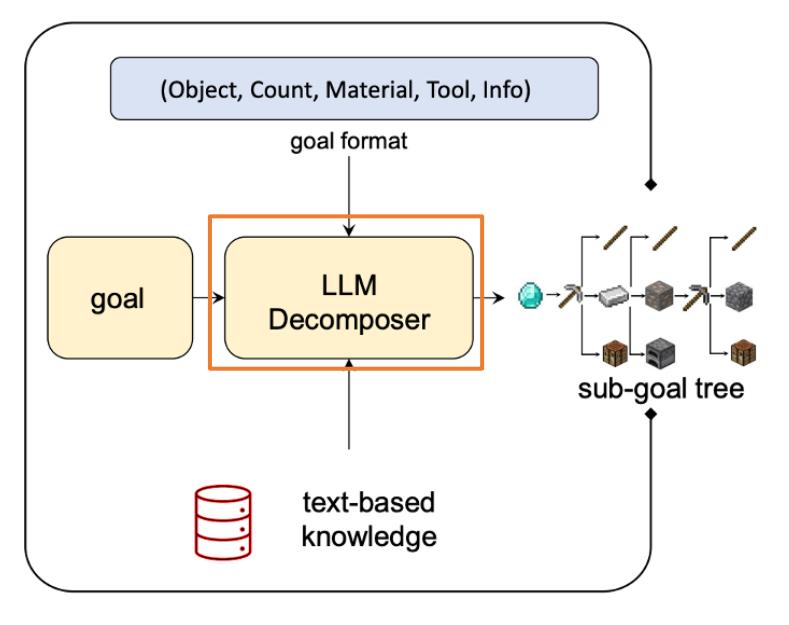
FORD [Xiong et al., 2023]



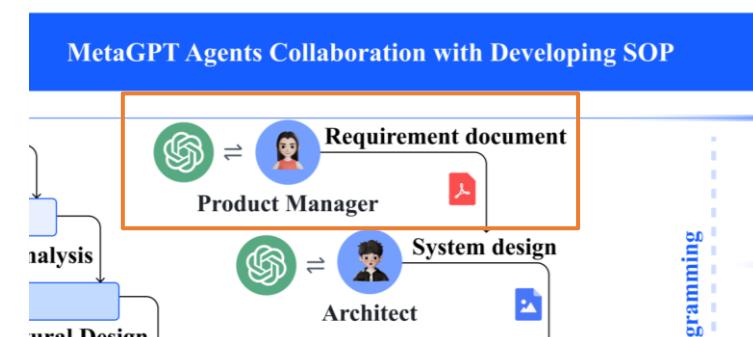
Self-collaboration [Dong et al., 2023]

角色-计划者

- 职能：定义目标，分析用户需求和优化执行计划
- 示例：分析师 (Self-collaboration [Dong et al., 2023]), 分解者 (GITM [Zhu et al., 2023]), 产品经理(MetaGPT [Hong et al., 2023])



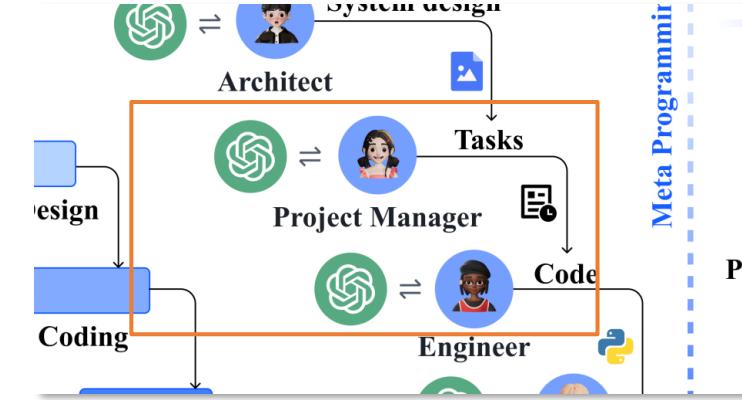
GITM [Zhu et al., 2023]



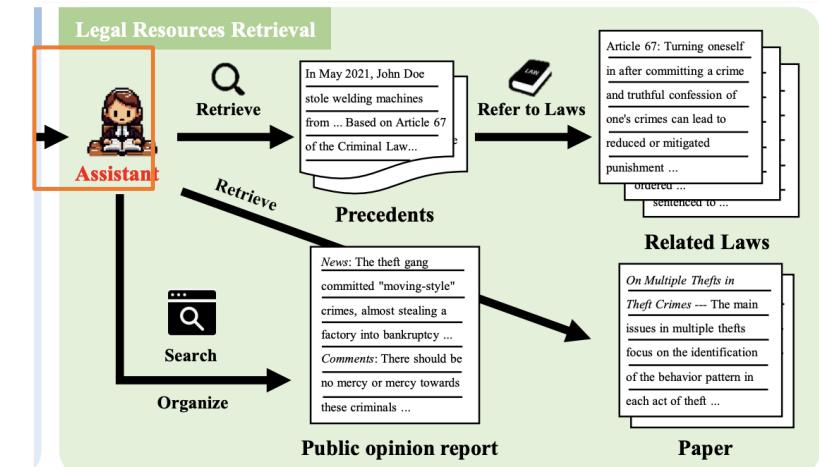
MetaGPT [Hong et al., 2023]

角色-协调者

- 职能：负责管理和协调智能体之间的协作
- 示例：项目经理 (MetaGPT [Hong et al., 2023]), 法官助理 (SimuCourt [He et al., 2024]), 秘书 (CosmoAgent [Jin et al., 2024])



MetaGPT [Hong et al., 2023]

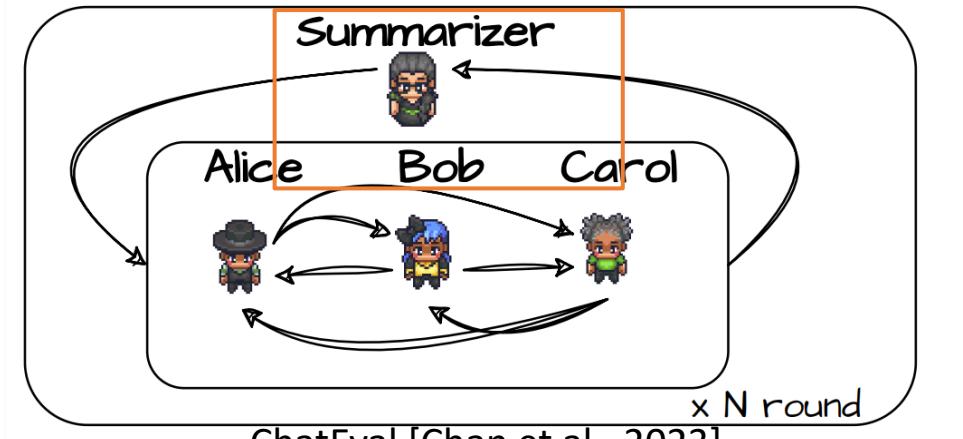


Simucourt [He et al., 2024]

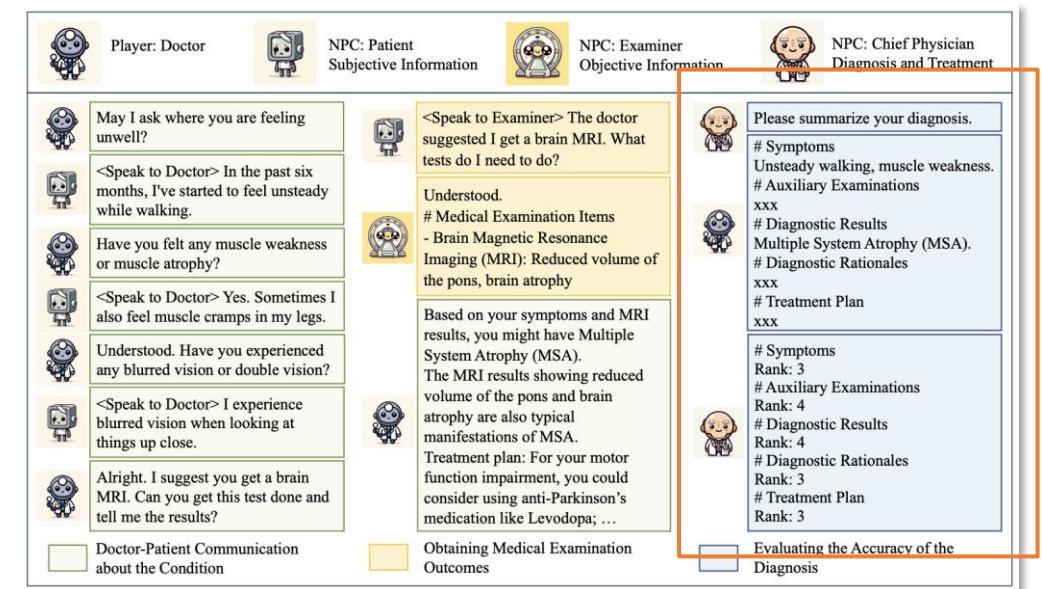
角色-集成者

- 职能：最终决策和总结

- 示例：总结者 (ChatEval [Chan et al., 2023]), 决策者 (Dera [Nair et al., 2023]), 主任医师 (AI Hospital [Fan et al., 2024])



ChatEval [Chan et al., 2023]



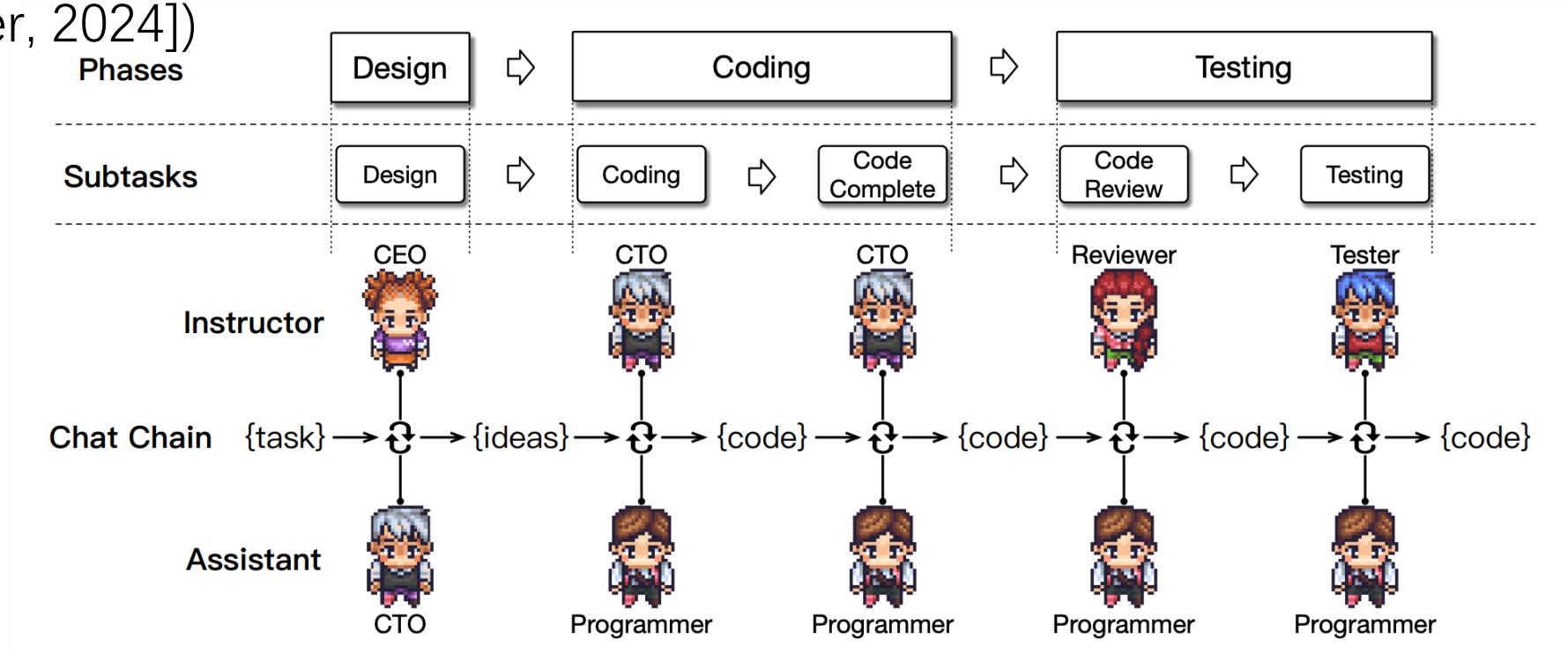
AI Hospital [Fan et al., 2024]

角色 - 组织形式



■ 静态模式 (static mode): 基于任务性质预定义的组织形式

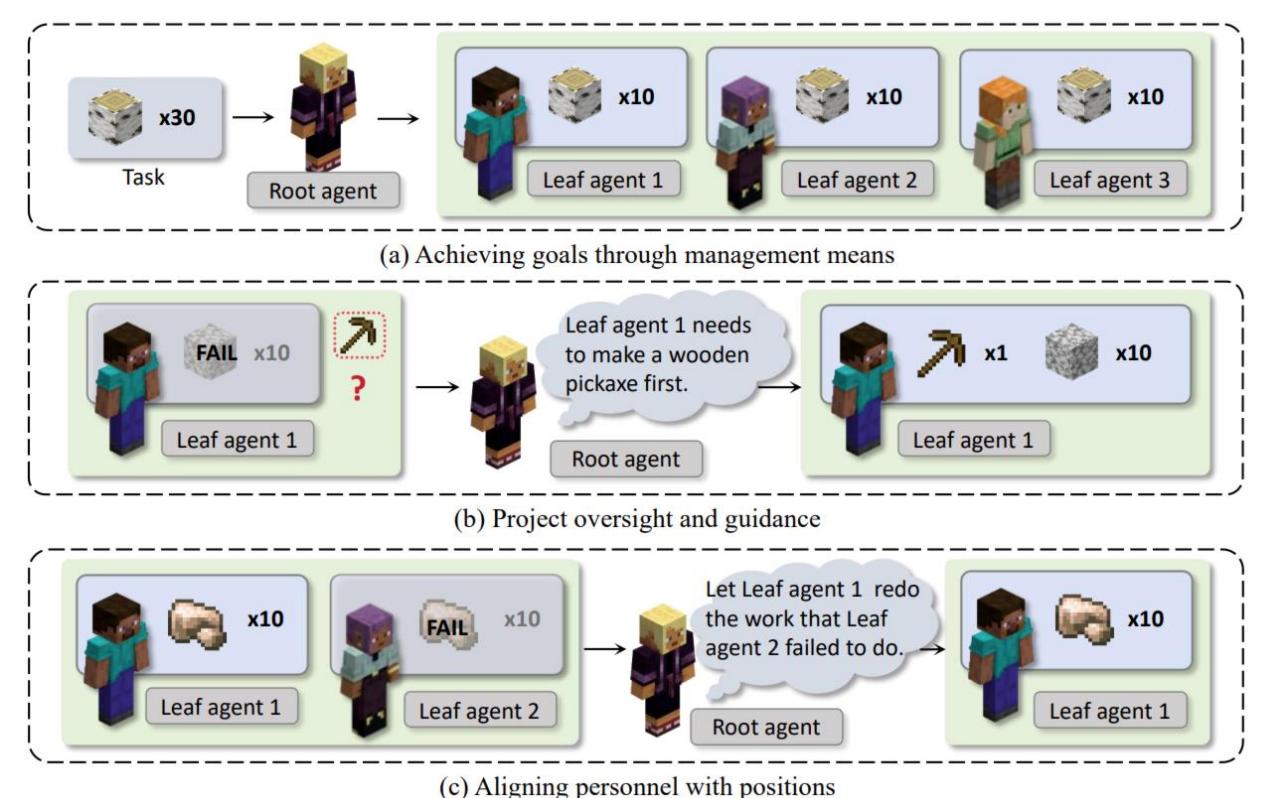
- 标准化软件开发 (ChatDev [Qian et al., 2024b], MetaGPT [Hong et al., 2023])
- 司法流程 (Simucourt [He et al., 2024], Simulating The U.S. Senate [Baker and Azher, 2024])



角色 - 组织形式

- 静态模式 (static mode): 基于任务性质预定义的组织形式
- 动态模式(dynamic mode): 基于动态性, 启发式的组织形式
 - 灵活创建/招募 (OpenAgents [Xie et al., 2023b], AutoAgents[Chen et al., 2023a], DoG[Ma et al., 2024])
 - 自组织 (S-Agents [Chen et al., 2024d])

S-Agents [Chen et al., 2024]

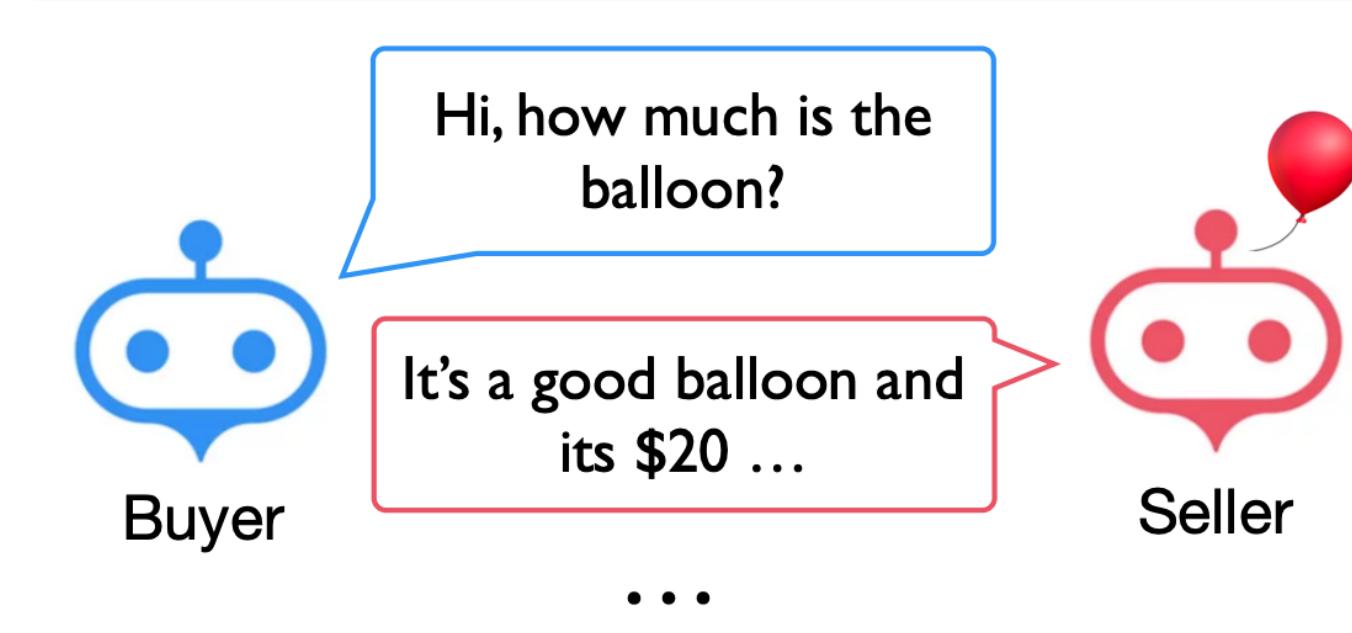


角色 - 沟通方式



■ 非结构化自然语言 (unstructured natural language) :

- 通过自由形式的交流实现灵活且即时的信息交换 (Dera[Nair et al., 2023], Camel[Li et al., 2023b], ICL-AIF[Fu et al., 2023])



角色 - 沟通方式

- 非结构化自然语言 (unstructured natural language) :

■ 结构化语言 (structured language)

- JSON、逻辑表达式 (AutoForm[Chen et al., 2024])
- 代码 (MetaGPT[Hong et al., 2023], ChatDev[Qian et al., 2024b])
- 医学报告 (MedAgents[Tang et

AutoForm [Chen et al., 2024]



A, B and C wear different colored hats: red, blue, and green.
A doesn't wear red. B wears green, what are their hat colors?

Natural Language

Since A doesn't wear red and B wears green, we can know
that A must wears blue. Then, C must wears red.



Logical Expression

$B=\text{green} \wedge A=\neg\text{red} \Rightarrow A=\text{blue}$, $A=\text{blue} \wedge B=\text{green} \Rightarrow C=\text{red}$



Code

```
colors = ["Red", "Blue", "Green"]
friends = {"Alice": None, "Bob": None, "Carol": None}
...
```

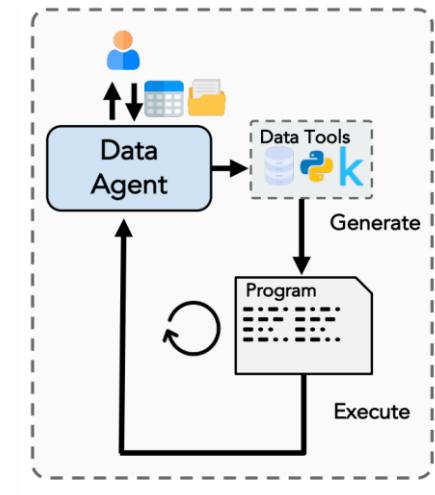


人类在环路



- 人类任务完成 (Human Task Accomplishment)

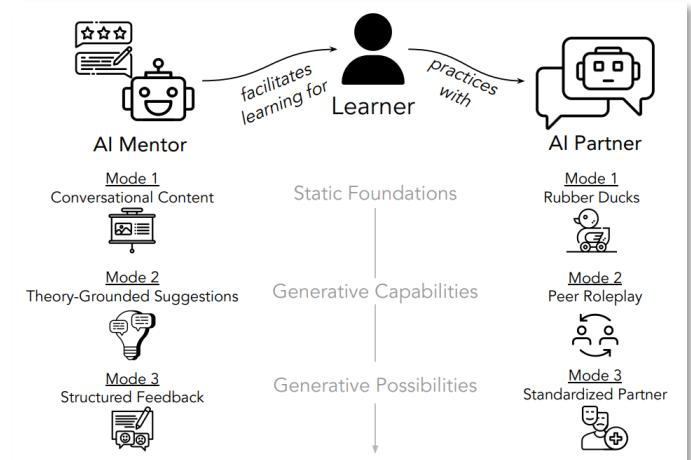
- (辅助任务完成) 个体提供有针对性的输入或补充信息以实现特定



OpenAgents[Xie et al., 2023]

- 人类技能发展 (Human Skill Development)

- (促进个体能力提升) 通过与智能体的交互体验来增强个体在特定领域内的能力



APAM[Yang et al., 2024]



场景模拟的研究要点

- 如何构建场景模拟系统? (组成要素)
- 如何应用场景模拟到现实任务? (场景分类)

一般性

职业

模拟场景选择

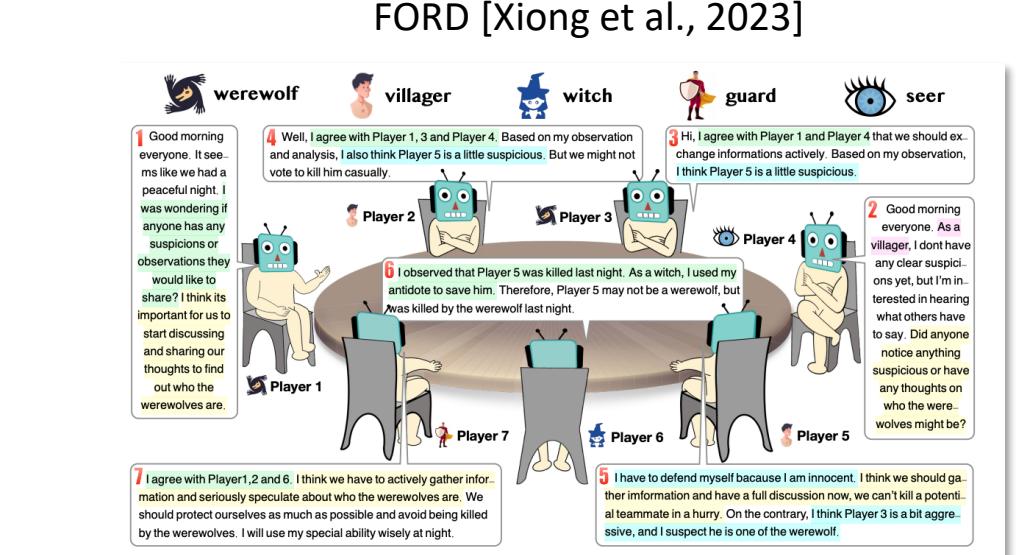
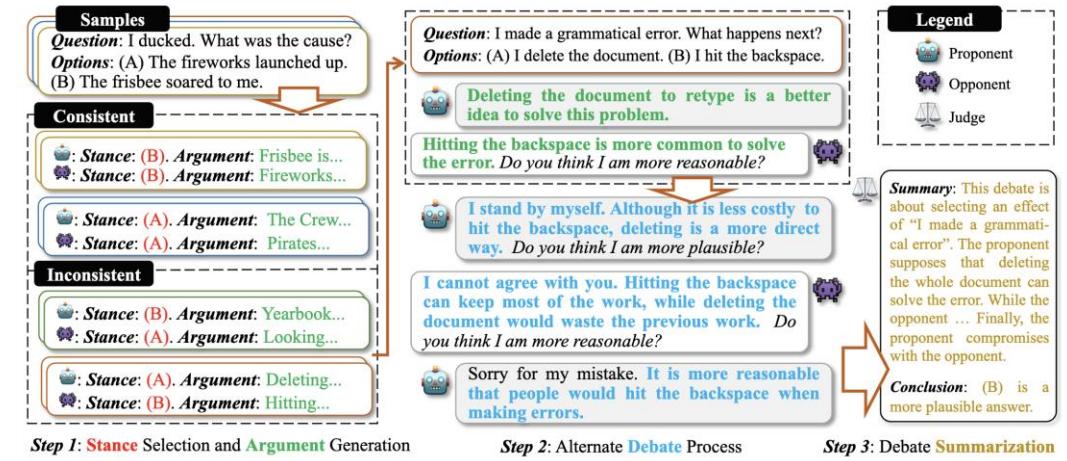
场景模拟构建

场景模拟-一般性场景



■ 问题回答(QA)

- 协作流程、战略推理和集成游戏
(FORD [Xiong et al., 2023], MAD [Liang et al., 2023b], DoG [Ma et al., 2024])
- 狼人杀、阿瓦隆等桌(ReCon [Wang et al., 2023e], Player* [Zhu et al., 2024b], WWQA [Du and Zhang, 2024])



Werewolf [Xu et al., 2023]

场景模拟-一般性场景

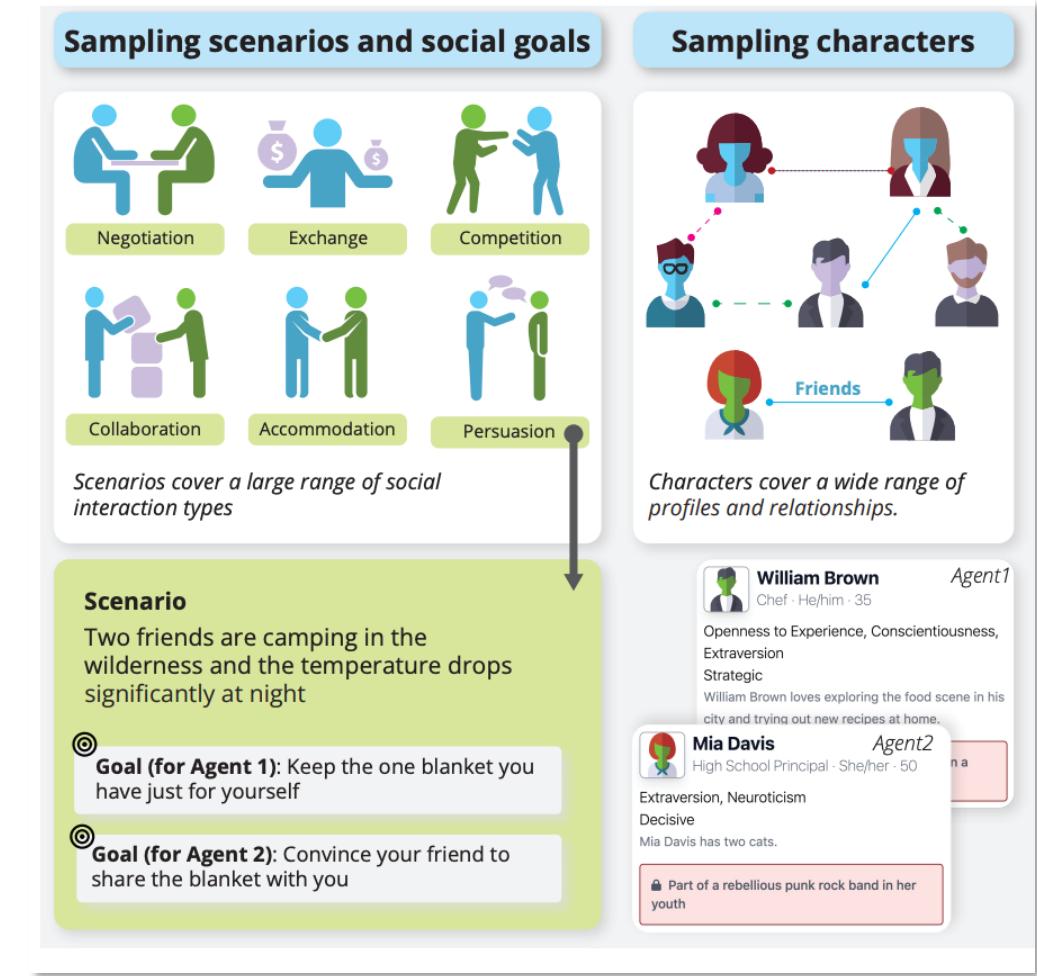


■ 问题回答(QA)

- 协作流程、战略推理和集成游戏
- 狼人杀、阿瓦隆等桌

■ 社交 (Social)

- 简单社交场景中的任务完成，比如说服朋友分享、邀请他人聚会、提供情感支持 (Sotopia [Zhou et al., 2023], AgentSense [Mou et al., 2024a], APAM [Yang et al., 2024])

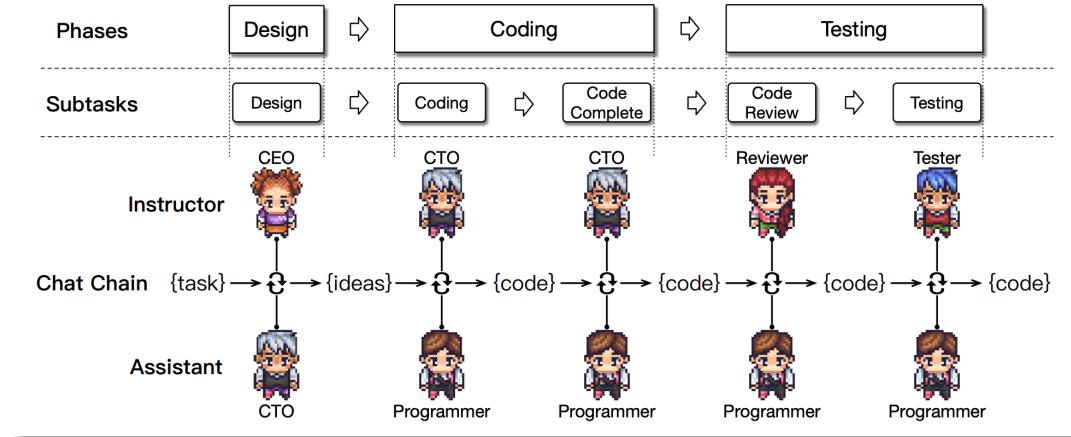


Sotopia [Zhou et al., 2023]

场景模拟-职业场景

■ 软件开发

- 软件开发, 生命周期管理 (ChatDev [Qian et al., 2024b], MetaGPT [Hong et al., 2023], Experiential Co-Learning [Qian et al., 2023b])

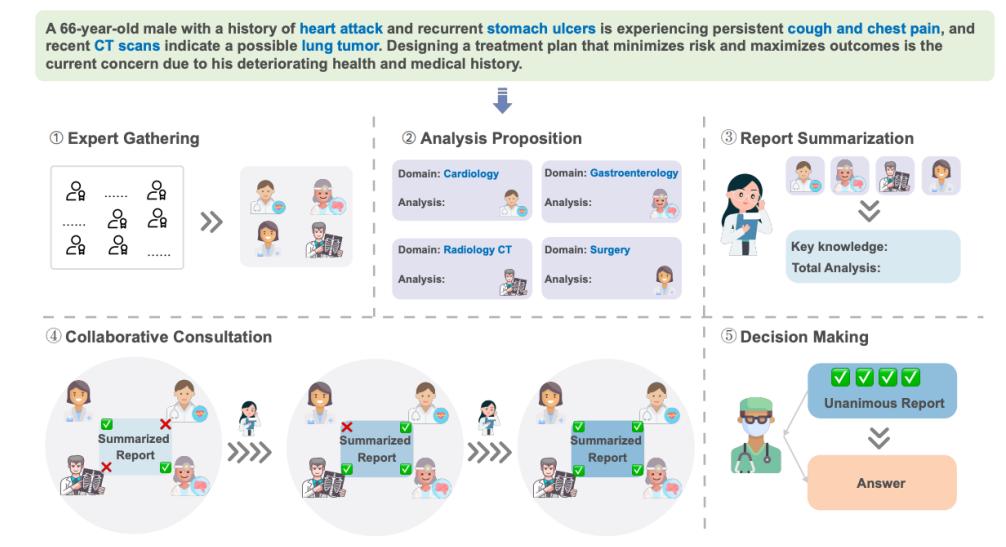


Chatdev [Qian et al., 2024]

场景模拟-职业场景



- 软件开发
 - 软件开发, 生命周期
- 基础与应用科学
 - 医学, 数学, 数据科学, 内容分析
(DR-CoT [Wu et al., 2023a], MedAgents [Tang et al., 2023], Agent Hospital [Li et al., 2024c])

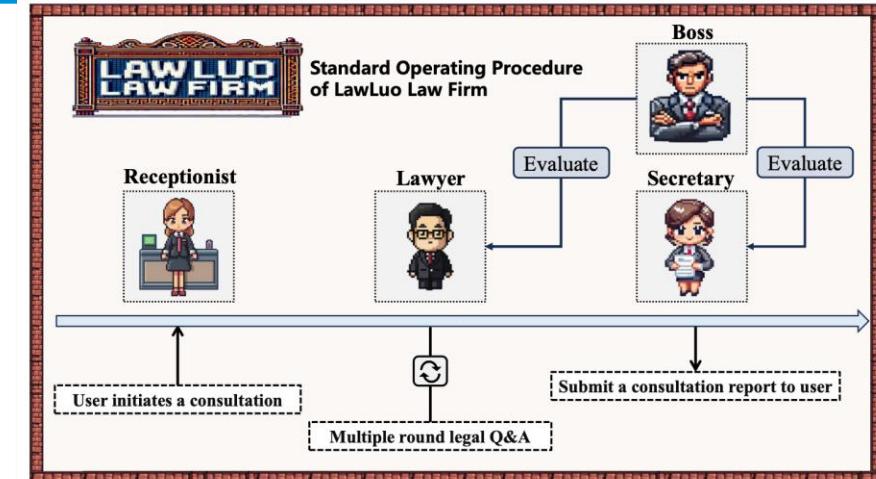


MedAgents [Tang et al., 2023]

场景模拟-职业场景



- 软件开发
 - 软件开发, 生命周期
- 基础与应用科学
 - 医学, 数学, 数据科学, 内容分析
- 其他行业
 - 司法 (LawLuo [Sun et al., 2024a], Blind Judgement [Hamilton, 2023], Simucourt [He et al., 2024])
 - 经济 (TradingGPT [Li et al., 2023c], Information Bazaar [Weiss et al., 2024])
 - 教育 (MAIC [Yu et al., 2024a], Mathvc [Yue et al., 2024])



LawLuo [Sun et al., 2024]

The screenshot shows the MAIC platform interface. On the left, the 'Current Slide' section displays a slide titled 'From Single-Agent to Multi-Agent' with content about multi-agent systems and applications like social simulation and software development. A 'Manager' agent is interacting with the slide, asking 'Who speaks next? ...'. Below this is the 'Class Roles' section, which includes icons for Teacher, Assistant, Classmate(s), and User. On the right, the 'Dialogue History' section shows a conversation between a 'Teacher' (blue icon), a 'Classmate' (orange icon), and a 'User' (grey icon). The Teacher asks about LLM-empowered multi-agent systems. The Classmate expresses excitement about the future and mentions emergent behaviors. The User asks if there are patterns similar to human groups. An 'Assistant' (yellow icon) responds with 'Good question! ...'. At the bottom, there is a 'Course Management' section and an 'Interactions in MAIC' section, both with placeholder text for interaction.

MAIC [Yu et al., 2024]

场景模拟的研究要点



- 如何构建场景模拟系统? (组成要素)
- 如何应用场景模拟到现实任务? (场景分类)
- 如何评估场景模拟? (评估方法)

模拟场景选择

场景模拟构建

场景模拟评测



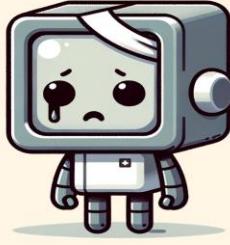
多智能体医疗场景赋能

- AI-Hospital: 多智能体交互式问诊框架与评测
- SFMSS: 多智能体场景仿真的医疗导诊模型
- SMART: 长短序列结合的多智能体联调框架



多智能体协同的交互式问诊



| | | | |
|---|--|---|---|
|  |  |  |  |
| 病人 | 检查员 | 实习医生 | 主任医生 |

- 病人： 配合医生问询，提供自身信息（如症状、病史等）
- 实习医生： 主动询问患者的症状、病史等关键信息，还原完整病史
- 检查员： 掌握病人全部的客观信息，忠实向病人反馈检查结果
- 主任医生： 掌握病人全部信息（包括主观信息和客观信息），评估实习医生问诊结果

AI-Hospital: 基于多智能体的交互式问诊平台

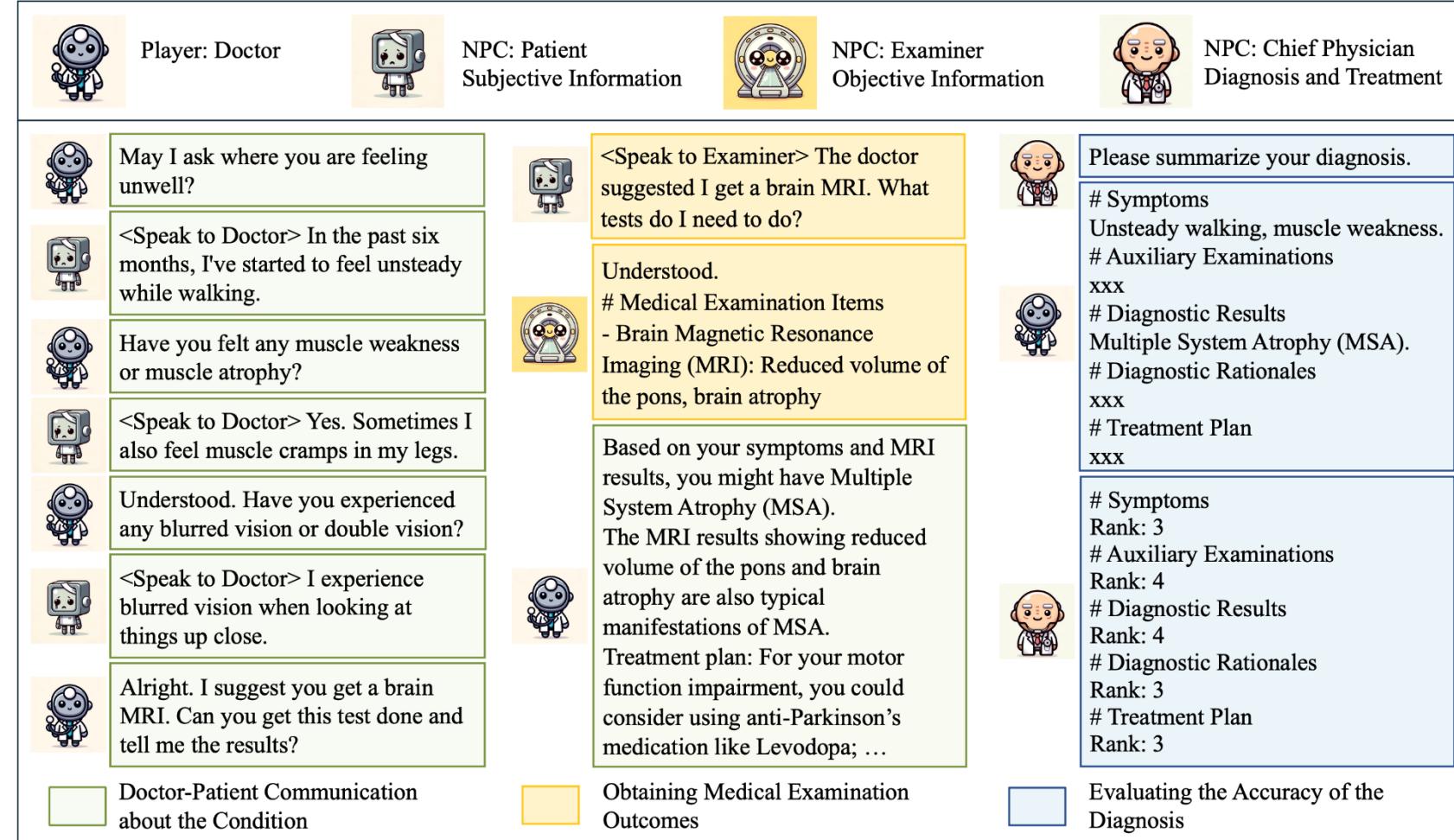


- ▶ 主任医生: GPT-4
- ▶ 病人: GPT-3.5
- ▶ 检查员: GPT-3.5
- ▶ 实习医生:

百川智能
BAICHUAN AI

文心一言

ChatGLM



大模型是否有能力完成不同角色的工作？

数据集建构



Table 1: Departments distribution.

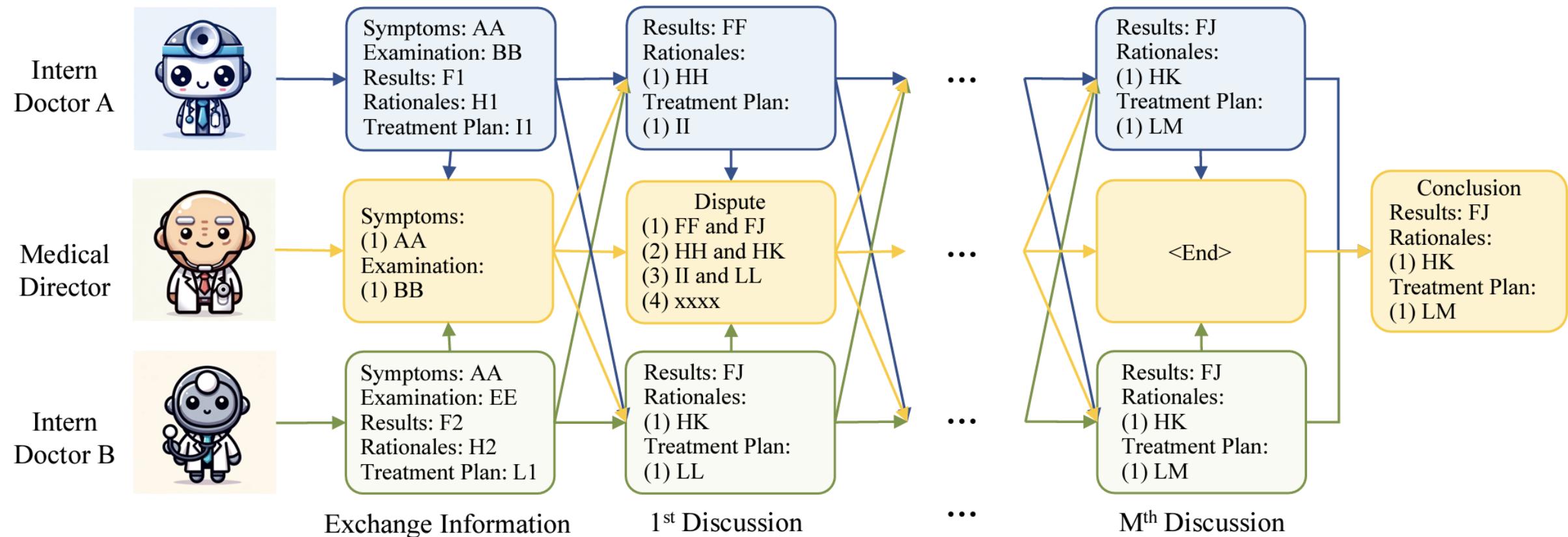
| Department | # |
|---------------------------|-----|
| Surgery | 180 |
| Internal Medicine | 153 |
| Obstetrics and Gynecology | 94 |
| Pediatrics | 29 |
| Otorhinolaryngology | 23 |
| Others | 27 |

实习医生: 无信息
病人: 主观信息 (症状、病史等)
检查员: 客观信息 (检查结果)
主任医师: 全部信息

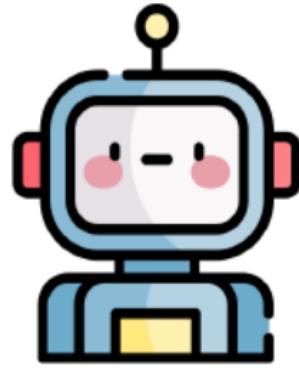
探索高质量的医疗服务方式（争议解决机制）



- ▶ 多实习医生，多轮交互讨论的设定
- ▶ 主治医生与实习医生进行多轮讨论，修正争议点，完成最终的评估



医疗大模型在场景落地遇到的挑战



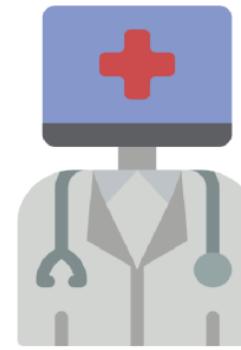
通用语言模型



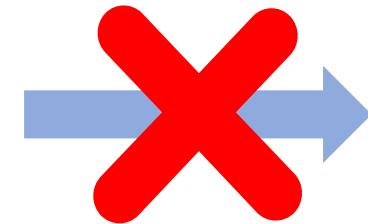
指令微调

单轮问答

医学知识
课本/科普



医学语言模型



多轮对话

场景复杂

多样患者

微调样本与真实场景的错配



患者：

一名 12 岁的女孩因发热、呼吸粗促和..她病情最可能的原因是什么？

LLM：

她病情最可能的原因是肺炎链球菌。

患者：

我发烧一天了，身体疼，喉咙也疼。我应该挂什么科室啊？

护士：

您好！您还有其他症状吗，比如咳嗽或咳痰？您最近有没有遇到过有类似症状的人？

患者：

我就想知道该去哪个科室，**问这些干啥！赶紧告诉我应该去哪！**还有我这次**能用医保吧？**



真实医疗场景的落地需求



▶ 真实场景的多样性和复杂性

真实的患者模拟

- 不同的**性格，背景**的患者，在不同**情绪**，不同**场景**下表现出多样化的**语言风格和行为偏好**
- 患者存在“不够理想”的行为逻辑：岔开话题/回避问题/情绪爆发

▶ 服务流控制

动作空间
监督信息

- 目标驱动的多轮对话，完成目标需要做什么->决策制定->进行回复
- 完成医疗目标需要的一系列**决策**和对应的能力（主动问询/信息收集/情感支持）

仿真环境-数据支持&场景准备



真实世界
人口统计分布

- 性别年龄
- 教育水平
- 收入情况
- ...



真实世界
人格特质样本

- 外倾性
- 友善性
- 神经质
- ...



公立医院
门诊记录

主诉:
发热两天, 胸腹部疼痛,
偶尔咳嗽

现病史:
患者持续发热, 最高达
39°C, 否认发热人员接
触史,

分诊科室:
发热门诊

初步诊断:
肺炎合并胸膜炎



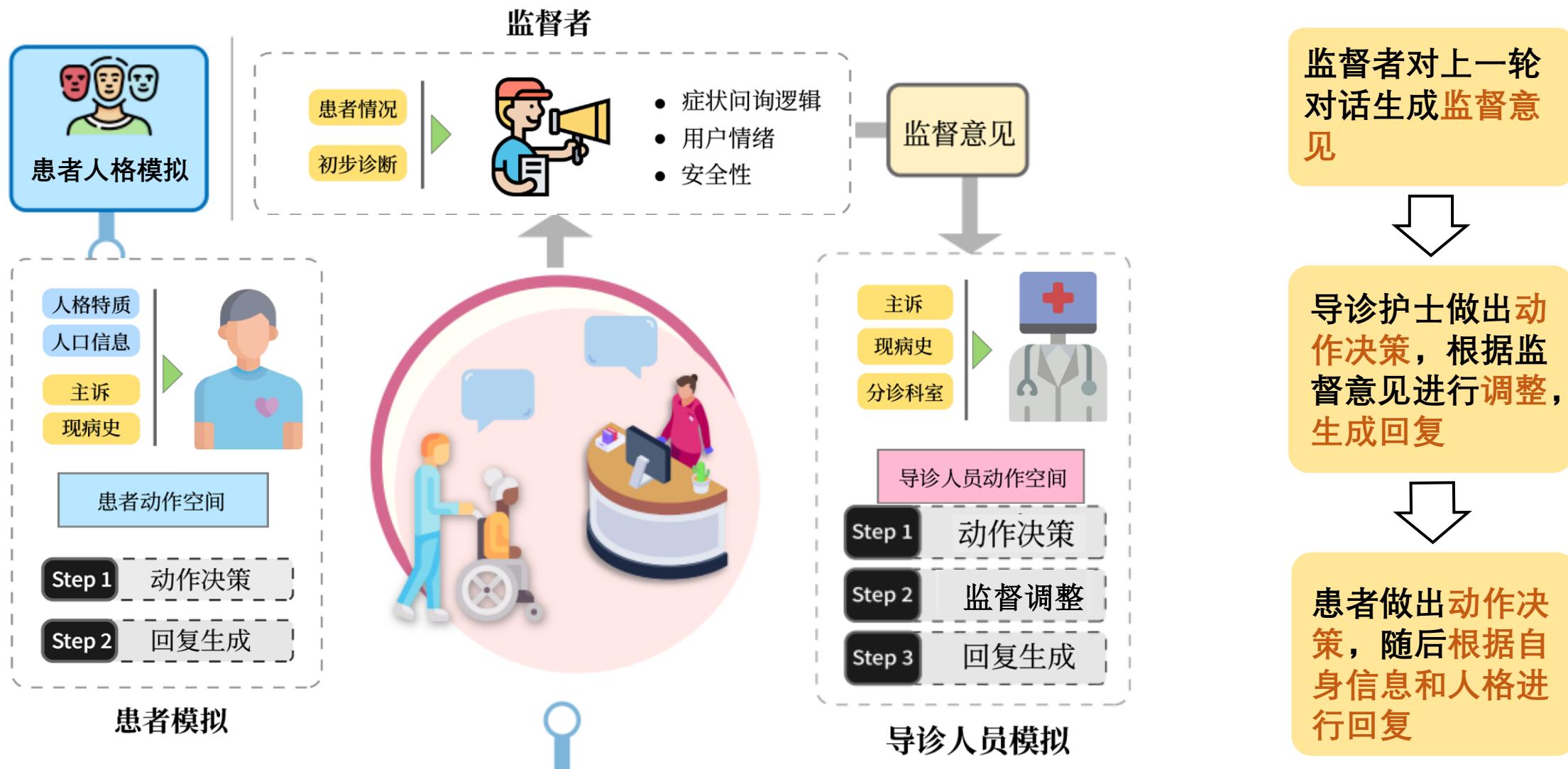
患者画像

场景设定

对医学知识的了解程度, 对症状的描述表达能力, 对导诊人员提
问的理解能力, 数字和逻辑能力, 沟通的积极性等方面。

患者来到医院时的场景和内心活动, 根据患者症状的严重程
度表述出患者不同的心理状态

仿真环境 - 对话模拟





实验结果 - 总体结果

- ▶ 使用SFMSS构造2000条仿真样本，在Qwen2-7B上训练得到 **SFMSS-Nurse**
 - Role-playing baseline：直接对GPT-4o角色扮演prompt代替导诊护士模拟，传统方法构造数据集进行训练。

| Method | Model | Accuracy | Overall Score | Info score | Average Turn Number | Average Turn Length |
|-----------------|------------------------------|--------------|---------------|-------------|---------------------|---------------------|
| Directly Prompt | GPT-4o | 0.717 | 3.83 | 2.16 | 3.54 | 207.98 |
| | Qwen2-7B | 0.634 | 3.65 | 2.28 | 4.22 | 336.40 |
| | Llama-8B | 0.401 | 3.24 | <u>2.65</u> | 4.44 | 678.14 |
| | HuatuoGPT2-13B | 0.501 | 3.25 | <u>2.17</u> | 3.57 | 258.38 |
| Fine-tuned | Role-playing baseline | <u>0.786</u> | <u>3.92</u> | 2.20 | <u>3.37</u> | <u>202.55</u> |
| | SFMSS-Nurse | 0.822 | 4.01 | 3.01 | 3.22 | 139.54 |

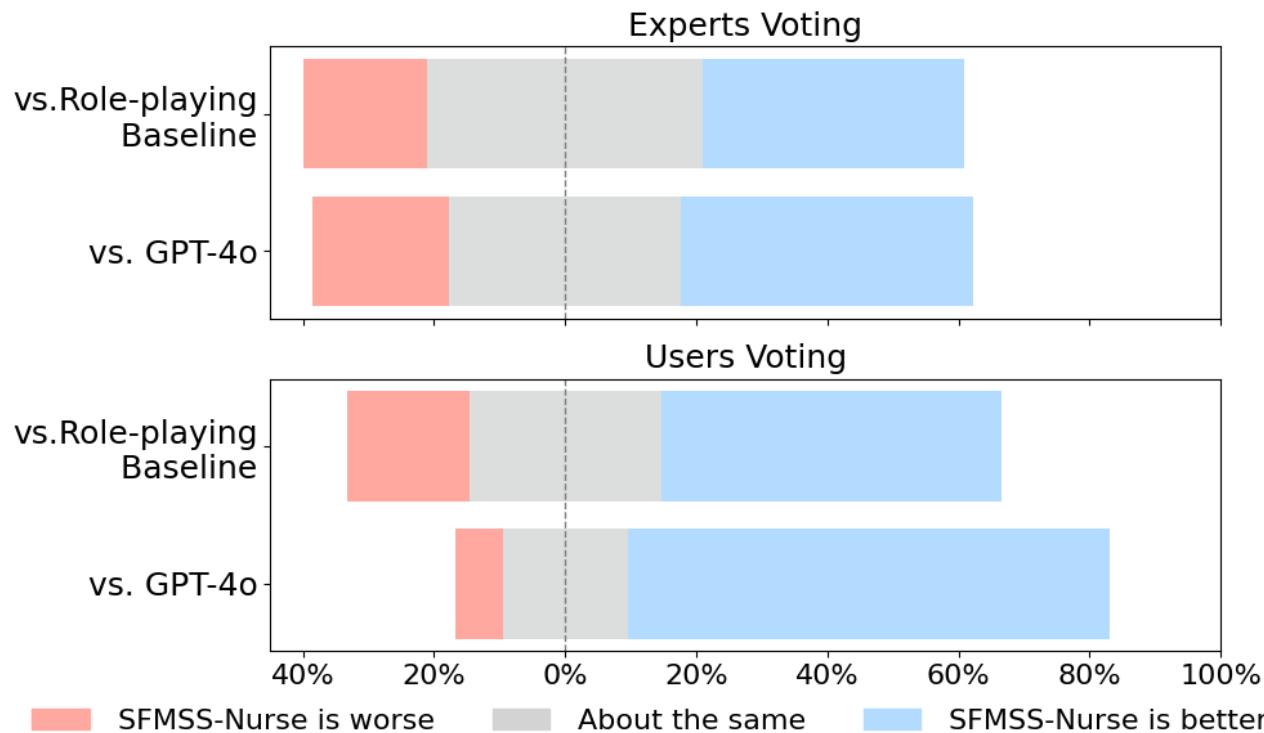
分诊更准确，更高效，信息收集能力更优，更好的整体表现

实验结果 - 用户&专家评估



▶ 招募15名用户和15名临床专家对进行匿名投票。

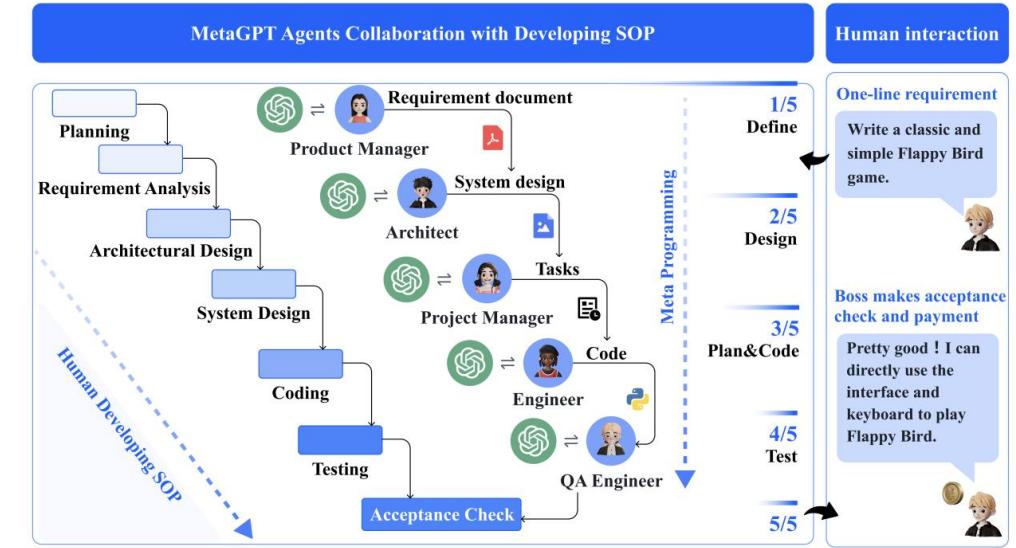
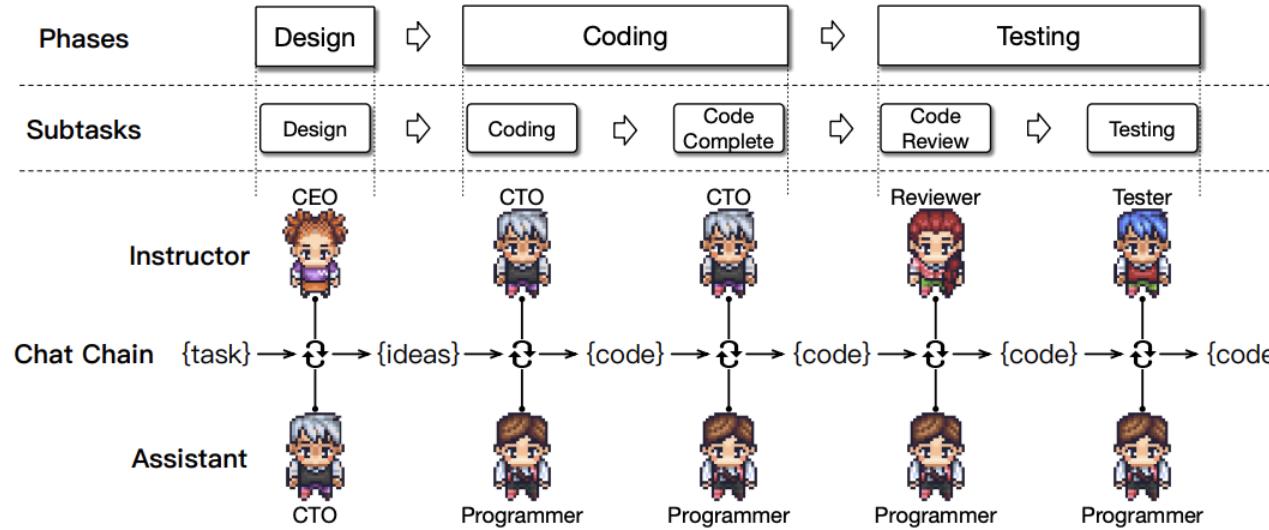
- 用户扮演患者与不同的模型进行对话，两两对比选择胜出的模型
- 专家阅读不同模型作为导诊护士与模拟患者生成的对话样本，并进行两两比较，选出胜出的模型



更加符合人类偏好，能够
处理真实情况

更加符合临床需求，医学
逻辑与真正的医生更对齐。

面向任务完成的多智能体协同框架



- 大部分多智能体框架都是**非训练方式**完成协调。
- 引导多个智能体在复杂的轨迹上进行协作是一项长期存在的挑战。

[1] *ChatDev: Communicative Agents for Software Development*, ACL 2024

[2] *MetaGPT: Meta Programming for A Multi-Agent Collaborative Framework*, ICLR2024

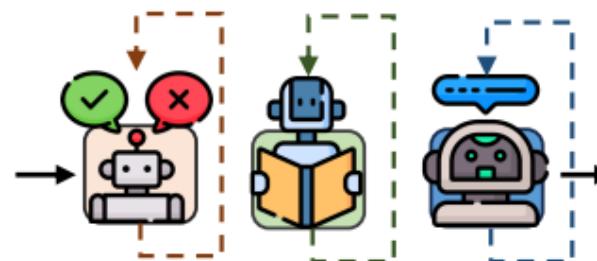
多智能体协同框架的训练方法

模块化操作

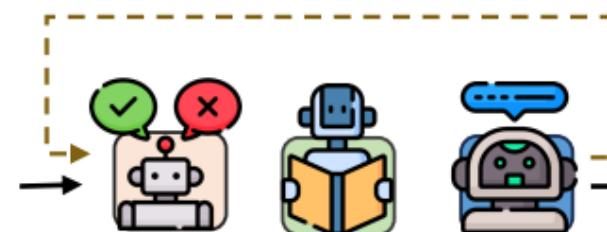
- 模块独立训练： 独立训练支撑某个角色的智能体，训练简便
- 串联智能体推理： 流水线化各个智能体，模块协作能力欠训练

端到端操作

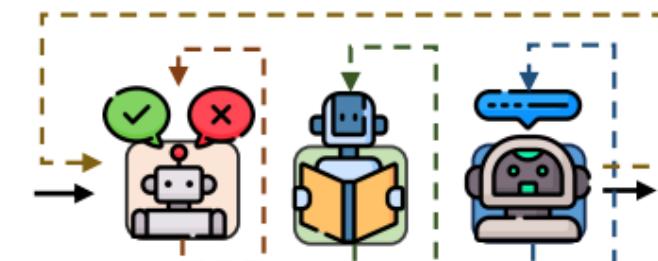
- 端到端的轨迹训练： 使用完整流程的轨迹训练多智能体协同过程，整体训练
- 串联智能体推理： 存在单一智能体能力不足的问题



(a) Modular Optimization



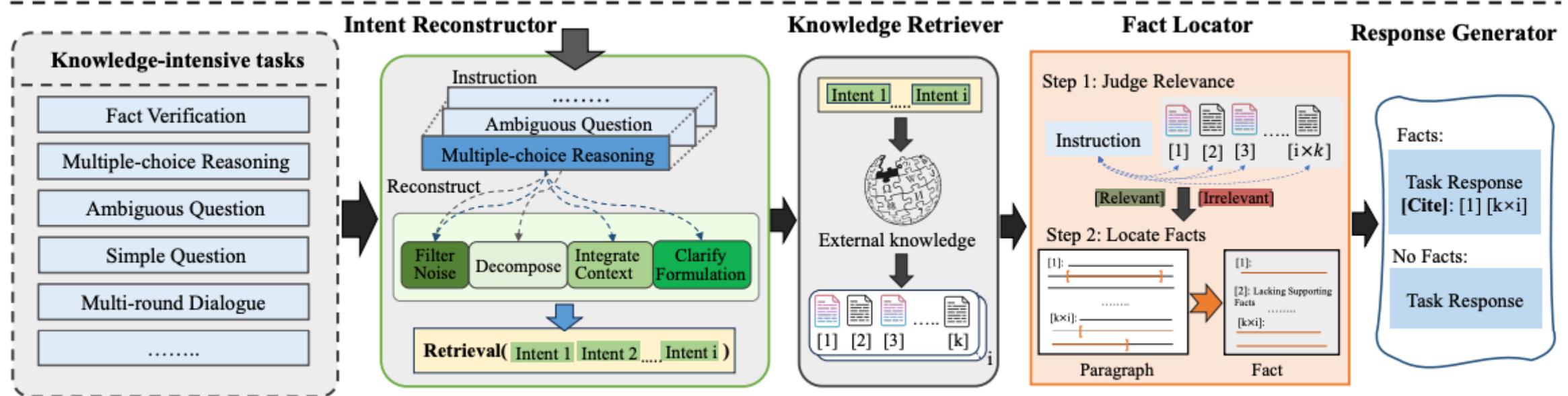
(b) End-to-end Optimization



(c) Ours

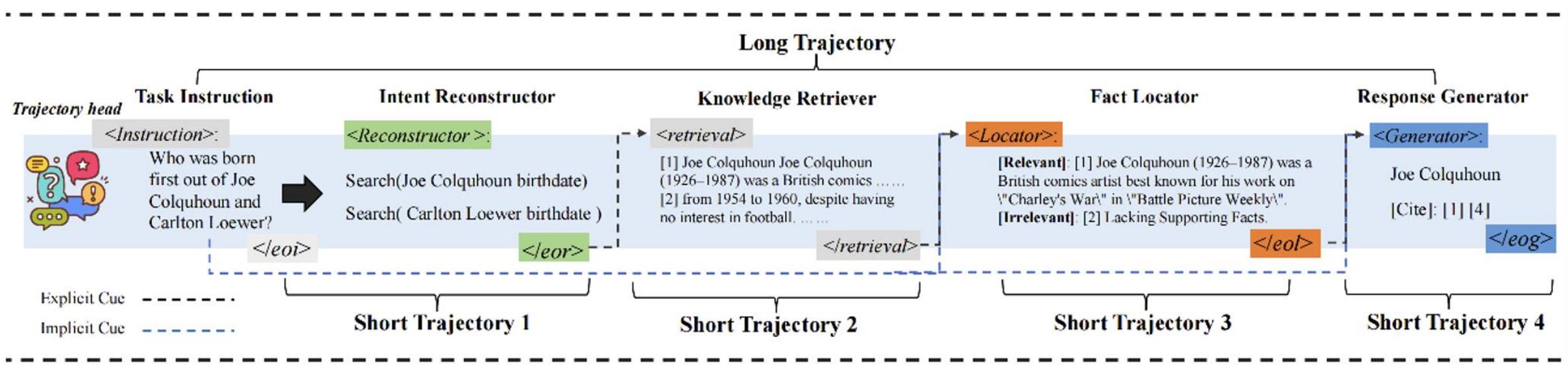
面向知识复杂场景的多智能体框架 - SMART

- 意图重构智能体
- 知识检索智能体
- 事实定位智能体
- 回复生成智能体



长短序列结合的多智能体协同训练

- 短轨迹学习：强化单一智能体的任务完成能力。
- 长轨迹学习：通过学习轨迹骨架，确保多个智能体之间的协同作用。



对比实验

➤ 相比于知识内化方法

- 获取长尾知识 (PopQA)
回复更准确。

➤ 相比于知识增强方法

- SMART的收益不仅来自
于多智能体框架，还来自
长-短轨迹学习机制。

| Task Metric | Health Acc | ARC-C Acc | PopQA Acc | Squad1 Acc | Str_EM | ASQA R-L | Mauve |
|--------------------------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| <i>General-purpose LLM</i> | | | | | | | |
| Aplca2 7B* | 44.78 | 36.43 | 25.58 | 11.50 | 14.42 | 28.72 | 51.24 |
| Mistral-Instruct 7B | 65.45 | 57.84 | 22.37 | 14.97 | 20.80 | 32.20 | 33.47 |
| Llama-2-Chat 7B | 47.95 | 47.95 | 25.44 | 14.13 | 16.79 | 32.35 | 24.21 |
| Vicuna-v1.5 13B | 63.01 | 57.59 | 17.94 | 15.25 | 31.95 | 22.99 | 68.41 |
| Llama-2-Chat 13B | 62.20 | 48.72 | 21.22 | 15.97 | 19.97 | 30.37 | 40.23 |
| ChatGPT | 76.08 | 77.3 | 29.30 | 22.90 | 39.94 | 35.73 | 44.63 |
| <i>Knowledge enhancement methods</i> | | | | | | | |
| Aplca2 7B* | 26.44 | 35.15 | 33.38 | 21.41 | 23.59 | 27.21 | 50.09 |
| REPLUG 7B* | 41.72 | 47.26 | 37.24 | 24.23 | 26.54 | 33.25 | 54.03 |
| VANILLA 7B* | 29.52 | 42.74 | 37.52 | 25.92 | 32.25 | 34.93 | 39.54 |
| RAIT 7B* | 52.98 | 62.10 | 38.02 | 23.86 | 25.68 | 15.99 | 12.35 |
| INTERACT 7B* | 65.45 | 48.12 | <u>41.31</u> | 31.52 | <u>34.54</u> | 35.51 | 43.45 |
| SelfRag 7B | 68.99 | 65.52 | 40.67 | 22.39 | 28.68 | 34.11 | 83.00 |
| MMAgent 3*7B* | 70.82 | 63.99 | <u>36.88</u> | <u>23.79</u> | 33.04 | <u>36.49</u> | <u>88.98</u> |
| SMART (OURS) | <u>73.18</u> | <u>65.58</u> | 42.60 | <u>27.80</u> | 41.16 | <u>40.66</u> | 91.47 |

Table 2: Comparison results against general-purpose LLM and knowledge enhancement methods. * denotes the method we reproduce based on the same base. * denotes re-implemented methods based on the same initial model. The **bold** numbers represent the best results and the underlined numbers represent the second.

场景模拟发展趋势

