**华中科技大学**

**基于多智能体架构的通用型AI助手技术研究任务书**

1. **项目基本信息**

参赛团队需要设计并实现一个基于多智能体架构的通用型AI助手。该助手能够按照用户自然语言指令，自动操作APP应用，完成用户交办任务。

AI助手需要利用多模态大模型、多智能体协作、自动化工具调用等技术，具备智能通信助手基础能力，如支持中国联通系列APP中设置电话智能代接、实现权益领取等；也可灵活扩展至其他日常服务，如消息回复、商品订购、出行规划、跨社交媒体分享等。

1. **项目背景**

在电信互联网深度渗透的当下，智能终端应用程序数量庞杂，且普遍存在功能入口分散、信息元素密集、交互逻辑复杂等问题。现有的个人助手工具或简单的智能应用在应对日常事务管理、信息检索、任务执行时显得力不从心，亟需更为智能和灵敏的交互技术。

基于多智能体架构的通用型AI智能体的发展和应用是当前人工智能领域的前沿方向。为了应对复杂任务处理难题，需要利用大模型的多模态信息处理能力、上下文理解能力，以及多智能体协作能力、工具调度能力，自主完成目标拆解、任务规划和操作执行的全流程闭环，在跨平台信息检索、复杂应用。

1. **项目内容**

内容一：电话智能代接

实现AI助手对中国联通系列APP中电话代接功能的自动化操作，支持用户通过自然语言指令设置智能代接规则（如特定号码拦截、自动回复等）。

扩展至其他通信场景，如自动记录来电信息、生成通话摘要等。

内容二：基于知识图谱的权益领取、消息回复、商品订购

在电信场景中，自动完成权益查询与领取（如“领取本月联通会员福利”）、权益比价与下单、智能消息回复（如“自动回复短信生日祝福”）、分享至其他社交平台。

内容三：基于高德地图MCP Server的出行规划功能

通过自然语言指令（如“规划一条避开拥堵的东湖骑行路线”）生成多模态出行方案，并自动同步至APP。支持动态调整（如途中用户指令“改去最近加油站”）。

1. **关键技术**

关键技术一：电话智能代接

基于LangChain，用户设置时解析用户指令转化为APP操作逻辑，通过Android无障碍服务或API模拟点击，完成代接设置。

接到来电后，首先大模型判断来电性质，据此确定不同的代接文本，再使用阿里的cosyvoice模型生成语音。

通过强化学习，在反复的电话代接与用户主动操作场景下，实现模型自学习。

关键技术二：基于知识图谱的权益领取、消息回复、商品订购

知识图谱：构建运营商权益库、商品信息库，支持语义匹配（如用户指令“最便宜的5G套餐”）。

基于LangChain的多智能体协作：“查询智能体”从图谱中检索信息；“决策智能体”生成分解的步骤；“执行智能体”操作APP完成领取或订购等具体操作。

通过强化学习，在反复的用户主动操作场景下，实现模型自学习。

关键技术三：基于高德地图MCP Server的出行规划功能

高德地图API集成：调用MCP Server获取实时路况、公交/骑行/步行路径规划。

多模态交互：结合地图数据与用户偏好（如“带小孩”推荐无障碍路线），生成图文报告。

自动化执行：通过Intent或深度链接启动导航APP并导入路线。

通过强化学习，在反复的用户主动操作场景下，实现模型自学习。

1. **目标软件**

一个集成了上述多模态大模型、多智能体协作功能的手机终端app，用户在手机上操作app，即可通过app自动触发后续跨平台的自动化工具调用。

1. **已有基础**

项目团队的各个成员作为一作发表论文2篇。团队有着丰富的任务流调度的研究经验，预期在比赛期间即11月前进行相关论文或专利的工作，以强化团队竞争力。

1. **工作计划与预期成果**

|  |  |  |  |
| --- | --- | --- | --- |
| **阶段** | **起始时间** | **结束时间** | **工作内容** |
| 阶段一 | 2025年7月 | 2025年8月 | 1、初步设计并实现基于LangChain的通用型AI助手框架，支持跨平台的自动化工具调用。  2、形成技术文档，通过初赛。 |
| 阶段二 | 2025年8月 | 2025年11月 | 1、完善多智能体协作机制。优化任务流调度和数据传输算法，降低通信开销和响应速度，提高执行准确率。  2、进行论文与专利申请。  3、进行决赛答辩。 |

1. **预期成果**：

**软件框架**：一套集成了上述多模态大模型、多智能体协作功能的手机终端app。

**技术论文**：投稿CCF B类的系统类论文。

**侧重点**：

* 研究多智能体（查询、决策、执行）的协作框架，提出任务分解与动态调度算法。
* 优化通信开销（如减少冗余信息传输）和响应时延（通过并行化或优先级调度）。

**2、关键指标**：

1）基础指标：

智能体任务流整体时延降低20%以上。任务间通信开销减少25%以上。智能体执行的准确率提升20%以上。

2）挑战目标：

智能体任务流整体时延降低30%以上。任务间通信开销减少35%以上。智能体执行的准确率提升30%以上。

1. **研发团队**

目前团队包括计算机类的3位硕士与7位本科生，分工如下：

1. 前端页面实现小组：3位本科生
   * 设计并开发手机终端App的用户界面。
   * 实现与后端模块的交互逻辑（如自然语言指令输入、任务执行结果展示）。
   * 集成多模态输出（如语音回复、图文报告、动态地图路线展示）。
2. 用户指令解析后端模块小组：2位本科生 + 1位硕士

* 开发自然语言处理核心模块，解析用户指令（如电话代接规则、权益领取请求）。
* 构建知识图谱（运营商权益库、商品信息库），支持语义匹配（如“最便宜的5G套餐”）。
* 与多智能体协作模块对接，将解析结果分解为任务流（如“查询→决策→执行”）。
* 优化上下文理解能力，支持动态调整（如用户中途修改指令）。

1. 多模型协作与自动化工具调用后端模块小组：2位本科生 + 2位硕士

* 实现多智能体协作框架（查询、决策、执行智能体的任务分解与调度）。
* 开发自动化工具调用逻辑（如通过Android无障碍服务/API模拟点击操作APP）。
* 集成第三方服务（如高德地图MCP Server、阿里语音合成）。
* 优化通信开销与响应时延（如并行化调度、冗余信息过滤）。