

# 基于 AI 智能体编排系统的实验难点挖掘\*

袁平海<sup>1\*\*\*</sup> 黄冲<sup>2</sup>

陈付龙<sup>1</sup>

1. 安徽师范大学计算机与信息学院, 芜湖  
241000

2. 安恒信息技术股份有限公司, 杭州  
310051

**摘要** 在教学过程中, 从学生作业和实验报告中获取有效反馈是提升教学质量的关键环节。传统依赖教师手动分析大量实验报告的方法不仅耗时费力, 也难以满足大规模教学的需求。为解决这一问题, 本文提出了一种基于 AI 智能体的编排系统, 自动化分析学生提交的作业与实验报告, 智能识别其中的实验难点。文章系统阐述了该编排系统的架构设计与实现方法, 并展示了其在实际教学场景中的应用。实践结果表明, 该系统在实验难点挖掘方面具有显著优势, 为智能教育技术的发展提供了全新的解决思路与实践范式。

**关键字** 智能教育, AI 智能体, 编排系统, 实验难点挖掘, 教学反馈自动化

## Automated Mining of Experimental Teaching Challenges Using AI-Agent Orchestration

YUAN Pinghai<sup>1</sup> HUANG Chong<sup>2</sup>

CHEN Fulong<sup>1</sup>

<sup>1</sup>School of Computer and Information  
Anhui Normal University,  
Wuhu 241000, China,

<sup>2</sup>DBAPP Security Co., Ltd  
Hangzhou 310051, China

**Abstract**—Obtaining effective feedback from student assignments and experimental reports is crucial for improving teaching quality. Traditional methods, which rely on teachers to manually analyze large volumes of reports, are not only time-consuming and effort-intensive but also struggle to meet the demands of large-scale education. To address this issue, this paper proposes an AI-agent-based orchestration system that automatically analyzes student-submitted assignments and experimental reports to intelligently identify teaching challenges. The paper systematically elaborates on the architecture design and implementation of this orchestration system and demonstrates its application in real-world teaching scenarios. Experimental results show that the system exhibits significant advantages in mining experimental challenges, providing a novel solution and practical paradigm for the development of intelligent education technologies.

**Keywords**—intelligent education, AI agent, orchestration system, teaching challenge mining, automated feedback

## 1 引言

在软件工程教学中, 实验环节承载着不可替代的核心教学功能<sup>[1,2,3,4]</sup>。以《程序设计》等软件工程类课程为例, 其教学目标强调理论知识的理解与工程能力的协同发展, 需通过“实践→理论→再实践”的迭代式学习路径, 引导学生完成知识的内化与能力的转化。实验教学不仅是将抽象理论具体化为可操作工程实践的关键桥梁, 更是落实“新工科”建设与“强基计划”等国家战略对高素质复合型人才培养的重要过程<sup>[5,6]</sup>。实验教学面临的核心挑战在于: 如何精准识别学生在实验过程中所遭遇的认知障碍与技能短板, 进而构建动态反馈机制, 实现教学难点的靶向干预。该问题的

有效解决, 直接关系到实验教学质量的整体提升。

从教学实践来看, 通过系统性分析学生实验报告以识别共性难点, 是一种行之有效的教学改进手段。这一方法能够帮助教师全面掌握学生在学习实践各关键环节(包括设计、编码、调试、测试与部署)中所面临的典型问题, 从而揭示教学过程中的薄弱环节, 为课程优化与教学策略调整提供决策依据<sup>[6,7,8]</sup>。然而, 传统的人工分析方式存在显著局限。其一, 面对大量实验报告, 人工处理不仅耗时耗力, 且难以保证全面性; 其二, 实验难点的识别本质上是一个高度依赖专业判断的过程, 对教师的专业能力与经验水平提出了较高要求。因此, 如何借助快速发展的人工智能技术, 实现实验难点的自动化挖掘与智能分析, 已成为当前教育技术领域亟待突破的重要课题。

直接采用“豆包”、“千问”或DeepSeek等通用AI智能体(AI-Agent, 简称智能体)进行实验难点抽取面临显著的技术局限性<sup>[9,10,11,12,13,18,19]</sup>。由于难点识别本质

\* \* 基金资助: 安徽省新时代育人质量工程项目(研究生教育)课程思政示范课程“高级计算机体系结构”(2024szsfk033), 安徽省高等学校省级质量工程项目虚拟教研室项目“计算机系统虚拟教研室”(2023xnjys008)。

\*\* 通讯作者: 袁平海 yph@ahnu.edu.cn。

上是一个复杂教学分析任务，需要同时考虑具体实验目标和内容、软硬件实验环境，以及学生报告的个性化等多维度因素。因此，必须与智能体进行多轮交互才能完成任务<sup>[14, 15]</sup>。另一方面，不同智能体的功能特性和访问配额存在明显差异，有必要利用好不同智能体的优势。通过编排多个智能体协同完成复杂任务，成为解决实验难点抽取问题的可行方案。目前，多智能体协作模式已在部分工业场景中得到初步验证<sup>[16, 17]</sup>。

本文旨在构建一个面向实验难点挖掘的AI智能体编排系统。该系统通过集成多个智能体，在满足各个智能体约束条件的前提下，充分发挥其各自优势，完成任务的智能分配与高效调度，从而协同完成实验难点的识别与提取。除支持“豆包”、“千问”等公开可用的通用智能体外，系统自身还内置了丰富的知识库，可用于对任务进行高层次语义理解与结构化拆解，支持任务的精细化分解与重组。此外，系统设计了灵活的扩展机制，提供多个功能增强模块以及用户自定义的模块扩展接口。利用扩展模块，用户可以在任务执行过程中对智能体的交互进行行为调整，并对输出结果进行过滤和修正，从而提升整体处理的准确性。

本文的主要贡献在于提出了一个面向教学场景的AI智能体编排框架，并基于该框架实现了一个系统原型。同时，本文以实验难点抽取任务为示例，详细阐述了系统的应用流程。

## 2 编排系统架构设计

编排系统架构设计旨在构建一个融合记忆管理、工具调用、任务规划与智能协作能力的结构化框架，以实现多智能体的高效决策与协同任务执行。该架构以动态记忆系统为基础，涵盖短期记忆与长期记忆模块，用于存储和管理任务上下文与历史交互信息。同时，系统集成很多功能组件（如计算引擎、信息检索、代码解析器等），为智能体提供丰富的能力支撑。

### 2.1 编排系统架构

本编排系统采用分层化架构设计，自上层应用至底层智能体，构建了一条完整的服务支撑链条，以实现高效、灵活的AI任务协同执行。如图1所示，在系统架构的底层，智能体负责具体任务的执行，包括图像识别、内容抽取、实验报告分析等操作。位于其上的智能体适配层用于对接来自不同平台的智能体实例（如豆包、千问、DeepSeek等），不仅保障了跨平台的兼容性，同时通过统一的接口规范，实现了异构智能体间的协同调用与统一调度。

智能体调度器是编排系统的核心，作为系统中枢，负责任务分配、资源协调与执行流程管理。另外，编排系统支持功能扩展，运行调度起外部模块进行交互。

其中，系统预定义功能扩展模块，提供开箱即用的基础能力组件；用户自定义功能扩展模块，支持个性化功能的集成与开发，满足多样化的教学需求。为便于实际应用的快速集成与定制，编排系统还通过标准化API接口提供服务接入能力，支持灵活的系统配置与功能扩展。

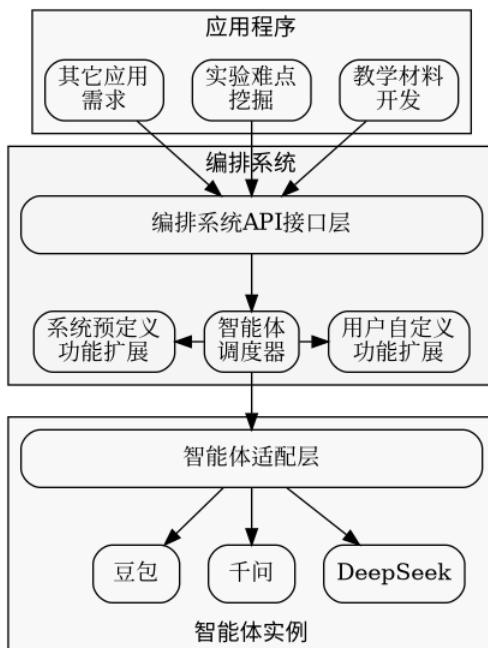


图 1 编排系统架构

最上层的应用程序层根据具体教学任务（如“实验难点挖掘”）构建定制化应用，充分体现出系统对多样化教学场景的适配能力。该分层架构通过接口抽象与模块化设计，既保障了核心调度功能的稳定性与高可用性，又通过灵活的扩展机制满足多场景、可定制的教学智能化需求，构成了一套可扩展、可演化的智能体协同工作解决方案。

### 2.2 智能体调度器

调度器为智能体提供了丰富的辅助功能，支撑其在复杂任务中的高效运行。如图2所示，围绕智能体运行的核心模块包括记忆(Memory)、规划(Planning)、行动(Action)与工具(Tools)，共同构成了智能体的功能支撑体系。

其中，记忆模块为智能体提供信息存储功能，分为短期记忆与长期记忆两个层次。短期记忆用于暂存当前任务中的临时数据，长期记忆则积累跨任务的知识与经验，智能体可随时读写，以支撑任务上下文的保持与知识迁移。规划模块紧密集成于智能体之中，通过反思、思维链(Chain of Thought)、自我批评与子目标分解等机制，引导智能体制定科学的任务执行策

略，为其行为提供明确的方向与逻辑支持。行动模块则负责将规划策略转化为实际操作指令，是智能体对外部任务环境产生实际影响的执行出口。此外，工具

模块为智能体提供包括日历、计算器、代码解析器、搜索引擎等在内的外部辅助能力，进一步增强其对复杂任务的适应性与处理能力。

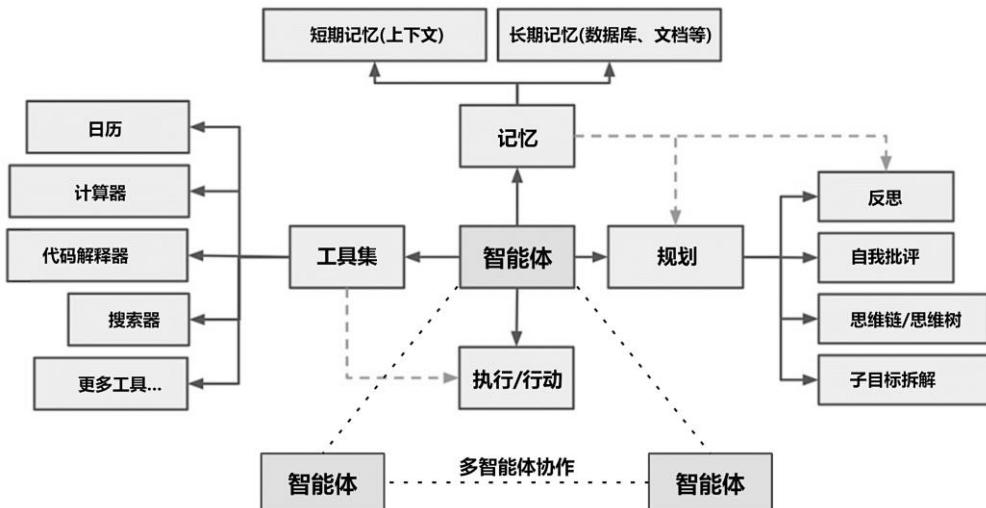


图 2 智能体调度器



图 3 交互式编排环境

在任务执行过程中，系统通过动态的反思机制和目标拆解策略，不断优化智能体的行动路径，提升任务完成的准确性与执行效率。同时，系统架构支持多智能体之间的协同交互，借助能力互补与信息共享，激发系统层面的“群体智能”效应。整体设计既保障了智能体个体的自主性，又强调系统整体的协同性，具备良好的适应性、可扩展性与演化能力，为面向复杂教学任务的AI应用提供了通用、稳健的技术支撑。

### 2.3 编排流程

系统提供了一个交互式的编排环境，用户可根据具体的应用需求自由配置任务处理流程与所需插件。智能体的输入内容通常为文件或文件集合，输出则是使用插件自定义的，整个处理流程由系统内部的智能体协同完成，确保任务的有序推进与高效执行。

如图3所示，在编排流程中，系统首先调用插件对文档内容进行提取与预处理，随后通过数据去重和结构化拆分，将原始报告转化为多个待分析的数据集合。接下来，系统智能体对学生报告中的设计、编码、调试、测试、部署等环节内容逐一展开分析与判断，

析，提炼出最核心的教学难点与指导策略，形成具有针对性的教学引导方案。

最终，系统将教学引导方案与详细分析报告通过工程化手段进行整合打包，生成统一的成果文件并输出给用户，即得到教学反馈数据信息，为教学辅助与个性化指导提供高质量的智能支持。

|    | A  | B    | C    | D    | E    | F    | G    | H    | I    | J    |
|----|----|------|------|------|------|------|------|------|------|------|
| 1  | 张千 | 子因曲  | 何羽音  | 郭东颖  | 何羽阳  | 吉工立  | 赵子翔  | 张苗   | 刘昱宣  |      |
| 2  | 李千 | 1    | 0.06 | 0.1  | 0.06 | 0.07 | 0.06 | 0.07 | 0.08 | 0.07 |
| 3  | 王思 | 0.03 | 1    | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| 4  | 何羽 | 0.07 | 0.06 | 1    | 0.06 | 0.12 | 0.06 | 0.12 | 0.07 | 0.12 |
| 5  | 郭东 | 0.03 | 0.03 | 0.03 | 1    | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
| 6  | 刘昱 | 0.05 | 0.04 | 0.19 | 0.04 | 1    | 0.04 | 0.98 | 0.06 | 0.96 |
| 7  | 王玉 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 | 1    | 0.03 | 0.03 | 0.03 |
| 8  | 赵子 | 0.05 | 0.04 | 0.18 | 0.04 | 0.96 | 0.04 | 1    | 0.06 | 0.92 |
| 9  | 张茂 | 0.05 | 0.04 | 0.07 | 0.04 | 0.05 | 0.04 | 0.05 | 1    | 0.05 |
| 10 | 刘昱 | 0.05 | 0.04 | 0.19 | 0.04 | 1    | 0.04 | 0.98 | 0.06 | 1    |

识别各维度存在的问题，并将分析结果汇总至教学引导智能体。教学引导智能体进一步对问题进行深度剖析。

给出用户，即得到教学反馈数据信息，为教学辅助与个性化指导提供高质量的智能支持。

图4 学生作业重复率检查结果

我的创建 / 运行调试

教学评估智能体-V1

智能体编辑中

保存    删除    返回

Agent对话

```

    "input": "请参考实验 提取其中的学习难点。",
    "guideFile": "file:缓存手册.pdf<fc75273e-a3bd-485d-aaea-46693b",
    "reportFiles": "file:学b6dd-4c50-9ca5-adf85417cc32>",
  }
  
```

缓冲区溢出漏洞    学生实验报告合集.zip

教学评估智能体-V1 6:37

根据指导手册和学生实验环境分类呈现如下：

根据指导手册和学生报告集合，排名前20的难点：

**一、基础概念理解难点**

- 1、栈帧结构与返回地址存储机制的具象化理解
- 2、缓冲区溢出与其它内存漏洞（如use-after-free）
- 3、控制流劫持中EIP/RIP寄存器修改的底层机制

**二、漏洞利用实操难点**

- 1、精确计算填充字符串长度以覆盖返回地址/函数指针
- 2、构造包含恶意地址的输入字符串（需处理字节顺序）
- 3、SUID权限提升攻击中环境变量的正确处理
- 4、ROP链构造时gadget地址的动态定位

**三、防御机制验证难点**

- 1、Stack Canary校验失败的现场保护机制分析
- 2、FS段寄存器在canary保护中的特殊作用
- 3、ASLR环境下函数地址的动态预测方法
- 4、PIE（位置无关可执行文件）与ASLR的协同工作

**四、实验环境相关难点**

- 1、编译选项（-fno-stack-protector等）对漏洞修复的影响
- 2、32位与64位系统下溢出攻击的差异处理
- 3、gdb调试器中观察栈内存变化的技巧

上一步    发布

图5 实验难点自动挖掘

### 3 教学实践

本文以《程序设计实验》课程为例，探讨编排系统在实验难点分析中的应用效果。

#### 3.1 应用需求分析

在数据处理阶段，系统首先采用文本相似度算法对学生实验报告进行预处理，过滤可能存在的抄袭内容，确保后续分析基于具有代表性的原创数据。对于学生在报告中明确描述的问题，系统可直接提取相关的关键信息；而对于未显式表达的问题，则借助代码

静态分析、执行日志分析以及大模型推理等手段，实现隐性问题的智能挖掘与识别。

在问题分析环节，系统会对学生报告中呈现或推断出的问题进行多维量化，涵盖问题类型（如概念理解、代码编写、实验操作等）、难度等级（简单、中等、复杂）以及涉及的知识点。对于未明确陈述的问题，系统通过对作业文本与代码的深度分析，挖掘潜在学习障碍，并将其纳入结构化建模过程。接下来，所有识别出的问题都会被编码为结构化特征向量，涵盖错误类型、难度层级、教学上下文等多个维度。该量化与建模过程不仅提高了问题分析的客观性与可计算性，还确保结果与教学目标、课程节奏高度契合，从而为后续的教学反馈与策略制定提供坚实的数据支撑。

### 3.2 自定义任务插件

虽然系统集成了大量插件模块，但针对特定任务需求，还需要自定义自己的插件。例如，用于学生报告预处理的“实验报告去重插件”是关键一环。抄袭现象在学生作业中较为普遍，若不加以剔除，可能导致模型在训练过程中出现偏倚，影响分析结果的准确性和公正性。因此，需要使用额外插件进行去重。

该插件基于 Python 实现，通过计算文本相似度，识别并剔除高相关性的抄袭报告。如图 4 所示，插件运行后会生成一个二维相关性矩阵，矩阵中的每个数值代表两份报告之间的相似程度，数值越高表示抄袭可能性越大。值得注意的是，该矩阵并非对称的，这是由于部分学生在抄袭过程中仅复制了他人报告的一部分内容。从图中可以看到，第“10”行和第“F”交叉处的内容为“1”，表明两位同学实验报告完全一样，这种报告需要踢除。

### 3.3 实验难点自动抽取

基于实验参考手册，智能体编排系统能够对学生提交的实验报告进行自动分析，从中识别并提取实验过程中的关键难点。系统首先解析实验报告内容，结合实验各个环节，梳理出可能存在的主要问题与瓶颈，如图 5 中间栏所示，通过与智能体编排系统进行交互，系统可以按模块化方式列出了典型的实验难点。

此外，如图 5 右侧所示，系统不仅列出每一项难点，还针对不同问题提供了相应的建议和扩展内容，涵盖知识补充、实验技巧提示、代码优化思路等，辅助学生进一步理解和掌握相关知识点。这种面向难点的自动提取与反馈机制，有助于教师关注并讲解原本未意识到的难点，提升了教学针对性与学生学习效果。

### 3.4 成效评估

本文所提出的 AI 智能体编排系统在“网络空间安全”专业 62 名学生的实验教学实践中发挥了显著的辅

助作用。如表 1 所示，从多个关键教学维度来看，该系统在效率与效果上均明显优于传统的人工处理方式。

首先，在时间效率方面，教师在人工审阅实验报告时，通常需要通读全文，并通过横向对比多份报告以归纳共性问题，平均每份报告耗时约 5 分钟。以一个班级为例，完整处理一次实验的全部作业需时约 5 小时。而通过本系统，在一台普通配置的 PC 上（如 i5-8250U，8GB 内存），可在 10 分钟内完成同等任务，效率提升高达 30 倍，大幅减轻了教师的重复劳动负担。

在实验难点识别能力方面，人工方式往往只能总结出少量显性问题，且容易遗漏隐性细节。而 AI 系统借助静态代码分析、语义理解等能力，可提取出更多具有代表性的实验难点，尤其是在变量命名、边界处理、模块设计等技术细节层面表现突出。在作弊检测方面，系统集成了文本相似度计算与图片相似度对比算法，可有效且准确地识别可疑报告。因此，系统在发现难点的同时，也显著提升了作业评估的公平性。

此外，在错误类型的覆盖广度方面，人工评阅通常侧重语法和基本语义问题，而 AI 系统能够全面覆盖包括语义逻辑、操作流程、环境配置等多个维度。这得益于系统中多智能体协同机制的支持，如语法检查模块、文本语义分析模块与环境一致性检测模块等实现了任务的专业化分工与协作。

综上所述，该 AI 智能体编排系统在提升教学效率、识别教学难点、保障评估公正性等方面表现出色，为智能化教学改革提供了可复制、可推广的实践路径，也为大规模教学反馈机制的建立奠定了坚实基础。

表 1 AI 智能体编排系统与人工处理的对比

| 评估维度   | 人工处理方式        | AI 编排系统             | 提升效果        | 附加说明                           |
|--------|---------------|---------------------|-------------|--------------------------------|
| 时间效率   | 5 小时/次实验      | 10 分钟/次实验           | 效率提升 30 倍   | 普通 PC: i5-8250U/8GB RAM        |
| 难点识别   | 预估约 6-8 个/次实验 | 20-40 个/次实验         | 发现率提升 3-5 倍 | 系统通过代码静态分析+文本语义理解，捕捉人工易忽略的隐性错误 |
| 作弊检测   | 依赖经验检出率~20%   | 自动识别 100% 检出        | 准确率显著提升     | 结合文本相似度和图片相似度进行检测              |
| 错误类型覆盖 | 关注语法语义错误      | 涵盖更多错误类型包括实验环境与操作步骤 | 分析维度显著增加    | 多智能体分工：语法检查，报告文本分析，环境配置检测      |

## 4 结束语

根据学生作业表现构建有效的反馈系统，是提升教学质量的关键环节。传统依赖人工分析的方法不仅耗时费力，而且难以适应大规模教学环境的需求。本文提出了一种基于 AI 智能体编排系统的自动化解决方案，设计实现了一个面向教学场景的智能编排系统。

该系统以“实验难点挖掘”为核心任务，构建了模块化、可扩展的编排架构，集成多种插件与知识库，支持任务流程的灵活配置与动态调度。教学实践表明，系统能够高效处理大量学生实验报告，自动识别并归纳出典型实验难点，生成针对性反馈建议，显著提升了教学干预的及时性与有效性。本研究不仅为智能教育技术在教学中的落地提供了有力支撑，也为构建面向教学的智能反馈系统探索出了一条切实可行路径。

## 参 考 文 献

- [1] 徐利锋,丁维龙. 面向对象 Java 编程课程教学设计策略[J]. 创新创业理论研究与实践, 2024, 7(11): 28-34+71.
- [2] 王娟. 基于 C++语言的计算机软件编程研究[J]. 信息记录材料, 2025, 26(05): 157-159+240.
- [3] 崔松健. 操作系统进程互斥与同步教学难点突破[J]. 电脑知识与技术, 2016, (05): 118-119+122.
- [4] 周林娥,史芳行. 数据结构课程混合式教学模式的探索与实践[J]. 信息与电脑(理论版), 2024, 36(20): 236-238.
- [5] 秦婧. 机器学习实验课教学设计研究. 计算机技术与教育学报, 2024,12(3): 1-5.
- [6] 高亚兰,吴孝银,徐旭,等. 基于 OBE 的工程教育模式在《计算机组成原理》课程中的应用与实践[J]. 电脑知识与技术, 2022, 18(21): 134-135.
- [7] 白丽瑞,张钢. 基于计算思维的软件工程导论课程教学实践[J]. 集成电路应用, 2023, 40(07): 94-95.
- [8] 张皓. 计算机软件开发设计难点及解决方式研究[J]. 计算机产品与流通, 2019, (04): 27.
- [9] 施敏,杨海军. 生成式人工智能的算法伦理难点分析与探索[J]. 大数据, 2025, 11(01): 167-174.
- [10] 李擎,崔家瑞,杨旭,等. AI 赋能工科专业实践教学管理的探索,[J]. 高等工程教育研究, 2025.
- [11] 李雪,范青刚,王忠, 等. AI 赋能的程序设计类课程项目化教学模式探索[J]. 计算机教育, 2025, (05): 33-38.
- [12] 王悦. AI 大模型赋能教辅图书智能化升级--以教辅智能体为例[J]. 编辑学刊, 2025, (02): 116-120.
- [13] 李飞飞, 张三, 王五, 等. 大模型检索增强生成技术研究进展[J]. 计算机学报, 2024, 47(5): 1089-1107.
- [14] 王冬青,陈自力,邵文豪,等. 从“负能”到“赋能”: 基于 LLMs 的思维链提示设计与教研 AI 智能体构建--以课堂教学智能分析为例[J]. 中国电化教育, 2025, (03): 111-117+125.
- [15] 董越,刘鹏,陈阵,等. 基于 LLM 的多智能体在核心网质差分析中的研究[J]. 网络安全和信息化, 2025, (01): 34-36.
- [16] 董之南,张勤学,胡进,等. 面向大模型多智能体系统的多维评估方法[J]. 指挥控制与仿真, 2025, 47(02): 121-131.
- [17] 王兵书, 李静, 雍珊珊, 等. 基于提示工程的程序设计探索与实践. 计算机技术与教育学报, 2024,12(3):166-172
- [18] Telnov Yury-Filippovich,Diveev Rustem-Ilyasovitch. Application of Multi-agent Technologies for Transforming Network Engineering Education[C], International Conference on Information Technologies in Engineering Education (Inforino): IEEE, 2024: 1-7.
- [19] Lu Chenglang,Yang Xuliang,Wang Beibei. Research on Higher Education Policy Effect Prediction Based on Multi-Agent Simulation Technology[C], International Conference on Consumer Electronics and Computer Engineering (ICCECE): IEEE, 2025: 428-432.