

在线智能评测系统在程序设计类课程实践教学中的应用研究

赵慧, 马凯, 胡秀建, 丁远远
亳州学院 电子与信息工程系, 安徽 亳州 236800
DOI: 10.61369/ETR.2026150044

摘 要 : 在高等教育数字化转型和“新工科”建设浪潮的推动下, 程序设计类课程作为计算机及相关专业的核心基石, 其教学模式及质量直接关系到学生的专业能力的培养。传统实践教学依赖人工批改, 存在反馈滞后、评判主观、环境不一、效率低下等问题, 难以满足大规模、个性化、高质量应用型人才培养的需求。在此背景下, 在线评测系统应运而生, 并迅速成为高校程序设计类课程教学改革的关键工具。本研究旨在系统梳理在线评测系统在程序设计类课程中的应用现状, 分析其不同教学场景下的成效与瓶颈, 并基于智能评测技术的发展, 探讨教育专用在线评测系统的优化路径与教学模式创新, 以期程序设计类课程的高质量发展提供理论与实践参考。

关 键 词 : 在线评测系统; 智能评测技术; 程序设计类课程; 教学实践

Research on the Application of Online Intelligent Assessment System in the Practical Teaching of Programming Courses

Zhao Hui, Ma Kai, Hu Xiujian, Ding Yuanyuan
Department of Electronic and Information Engineering, Bozhou University, Bozhou, Anhui 236800

Abstract : Driven by the digital transformation of higher education and the construction of "Emerging Engineering Education", programming courses, as the core foundation of computer and related majors, have a teaching mode and quality directly related to the cultivation of students' professional abilities. Traditional practical teaching relies on manual grading, which suffers from delayed feedback, subjective evaluation, inconsistent environments, low efficiency and other problems, making it difficult to meet the needs of large-scale, personalized and high-quality applied talent training. Against this background, online assessment systems have emerged and rapidly become key tools for the teaching reform of programming courses in colleges and universities. This study aims to systematically sort out the application status of online assessment systems in programming courses, analyze their effects and bottlenecks in different teaching scenarios, and explore the optimization paths and teaching mode innovation of education-specific online intelligent assessment systems based on the development of intelligent assessment technology, so as to provide theoretical and practical references for the high-quality development of programming courses.

Keywords : online judge system; intelligent assessment technology; programming courses; teaching practice

一、研究背景与现状

随着信息技术的飞速发展与广泛应用, 程序设计能力已经逐渐演变为当代大学生在数字化时代生存与发展所不可或缺的核心素养之一。程序设计类课程作为计算机及相关专业的基础性课程, 其教学质量直接关系到学生对计算机科学基础理论的掌握程

度、实践能力的培养以及创新思维的激发。这些课程不仅教授学生编程语言的基础语法和编程技巧, 更重要的是培养学生的计算思维、问题分析和解决能力。^[1] 因此, 高质量的程序设计教学能够有效提升学生的专业素养, 增强他们在未来就业市场中的竞争力, 为他们适应快速变化的技术环境奠定坚实基础。然而, 传统实践教学面临着教师批改作业负担重、反馈不及时、评价标准不

基金项目: 亳州学院教育教学改革研究项目“基于智能评测技术的程序设计类课程实践教学研究”(2023XJXM050)。

作者简介:
赵慧(1994-), 女, 讲师/博士, 研究方向: 程序设计类课程实践教学研究;
马凯(1993-), 男, 硕士/助教, 研究方向: 教学效果评估方法研究;
胡秀建(1979-), 男, 硕士/教授, 研究方向: 计算机专业学科建设;
丁远远(1990-), 女, 硕士/讲师, 研究方向: 计算机类课程教学新模式探索。

统一等突出问题。根据教育部2022年发布的《普通高等学校本科教育教学审核评估实施方案》，实践教学环节的质量提升已成为高校教学改革的重要方向。

得益于以浙江大学 ZOJ、杭州电子科技大学 HDOJ、LeetCode、牛客网等为代表的在线智能评测（Online Judge，简称 OJ）系统，通过自动化编译、运行与测试，实现了对学生代码正确性、运行效率与内存占用的即时、客观评判。^[2]研究表明，采用 OJ 系统后，教师批改作业的时间成本可降低60%以上，同时学生获得反馈的时间从平均3天缩短至即时反馈，这种即时反馈机制显著提升了学生的学习效率和编程能力培养效果。

然而，随着应用的深入，通用型 OJ 平台在支持系统化、结构化课程教学方面的局限性日益凸显。其题库设计与高校课程大纲脱节、反馈机制流于表面、缺乏个性化学习路径支持等问题，制约了教学效果的进一步提升。因此，如何将人工智能、大数据分析等智能技术深度融入 OJ 系统，构建一个集自动化评测、知识诊断、智能推送与数据反馈于一体的新型实践教学模式，已成为当前教学改革的前沿课题。本研究旨在系统梳理在线评测系统在高校程序设计课程中的应用现状，深入分析其不同教学场景下的成效与瓶颈，并基于智能评测技术的发展趋势，探讨面向未来的教育专用 OJ 系统的优化路径与教学模式创新。

二、OJ 系统在程序设计类课程教学中的应用模式

（一）程序设计类课程教学特点分析

程序设计类课程作为计算机科学教育的核心组成部分，其教学特点主要体现在实践性、逻辑性和创新性三个维度。实践性特征要求课程必须注重学生动手能力的培养，通过编写代码、调试程序等实际操作来巩固理论知识。逻辑性特征体现在课程内容具有严密的思维体系，从基础语法到算法设计，再到系统架构，各知识点之间存在严格的递进关系。创新性特征则要求学生具备解决开放性问题的能力。教学过程中，教师需要平衡知识传授和能力培养的关系，既要确保学生掌握基础语法规则，又要培养其算法设计和系统实现的综合能力。评测环节的即时性和准确性对教学质量影响显著。课程还呈现出明显的分层教学特点，不同基础水平的学生在代码理解、算法实现等方面存在显著差异，需要采用差异化教学策略。具体如下表1所示。

教学特点维度	具体表现	量化指标 / 示例	教学要求 / 影响
实践性	注重动手能力培养，通过编写代码、调试程序巩固理论	完成50-100个编程练习后，代码实现能力提升40%以上；采用项目驱动教学法	需设计大量实践环节，将知识点融入具体案例，增强学习针对性和趣味性
逻辑性	课程内容具有严密思维体系，知识点间存在严格递进关系	使用流程图、伪代码等可视化工具辅助教学	教师需帮助学生建立清晰的程序逻辑思维模型，确保知识体系的连贯性和层次性

教学特点维度	具体表现	量化指标 / 示例	教学要求 / 影响
创新性	要求学生具备解决开放性问题的能力，鼓励探索多种解决方案	课程设计中包含20%-30%的开放性题目	需平衡知识传授与能力培养，既确保基础语法掌握，又培养算法设计与系统实现的综合能力
评测环节	即时性和准确性对教学质量影响显著	传统人工批改存在效率低、标准不统一问题；智能评测技术可提升客观性和时效性	需引入高效、准确的评测技术以支持教学反馈
学生差异性	不同基础水平学生在代码理解、算法实现等方面存在显著差异	表现为分层教学需求	需采用差异化教学策略，适应不同学生的学习进度和能力水平

表1 程序设计类课程教学特点分析表

（二）闭环教学流程的构建

OJ 系统已从最初的竞赛工具演变为高校程序设计类课程实践教学的核心基础设施，其应用模式日趋成熟，形成了覆盖教学全周期的闭环体系——“课前－课中－课后”教学闭环。^[3-4]课前，教师依据课程大纲，从 OJ 系统题库中选取或者自定义与知识点紧密关联的题目，作为预习或导学任务；课中，OJ 系统成为实践课的核心操作平台，学生进行编程练习，系统提供即时反馈，教师根据实时排行榜和错误统计展开针对性讲解；课后，教师通过 OJ 系统布置作业、进行进阶练习与阶段性考核。这种闭环模式将传统的单向知识传授转变为双向互动的数据驱动型学习过程。

（三）基于 OJ 系统的核心教学功能的实现

自动化实时评判：系统通过编译学生提交的代码，利用预设的多组测试用例进行黑盒测试，快速返回“Accepted”、“Wrong Answer”、“Runtime Error”等状态。这种即时反馈机制使学生能够迅速定位问题，反复调试，极大提升学习效率和问题解决能力。

学习行为监控与教学管理：教师可通过系统后台，监控每位学生的做题数量、通过率、提交时间序列等。这些数据为教师识别集中困难点和个体薄弱点提供了精准依据，使得分层教学和差异化辅导成为可能。

学习诚信保障：系统内置的代码相似度检测算法能够有效识别潜在的抄袭行为，并向教师发出预警，辅助进行学习诚信管理。

激励机制与社区构建：部分 OJ 系统中排行榜、积分、徽章等游戏化元素，以及在线比赛、闯关模式，显著激发了学生的学习兴趣与竞争意识；讨论区的设立则促进了学生之间的互助交流。

（四）现有主流 OJ 系统的教学适配性分析与瓶颈体现

尽管 OJ 系统在高校中广泛应用，但主流通用平台设计初衷主要服务于算法竞赛和职业技能提升，在与高校结构化课程教学深度对接时，暴露出多方面的功能瓶颈。主要体现在以下方面：

（1）题库与课程教学的脱节

这是最核心的瓶颈之一。通用平台的题库以算法类型和竞赛难度为导向，题目标签（如“数组”、“动态规划”）相对宽泛，缺乏与高校《程序设计基础》、《数据结构》等具体课程章节知识点的精细映射。教师难以根据本周讲授的“栈的应用”或“图的深度优先遍历”等具体内容，快速筛选出难度适宜、背景贴切的题目序列。这导致学生课堂所学与课后练习之间出现“学用脱节”，降低了练习的针对性和效率。

（2）反馈机制的浅层化

现有系统反馈大多停留在“对 / 错”及简单错误类型提示层面，此时，学生通常只知道结果不对，但无法获知是哪个测试用例失败、程序的中间状态如何、逻辑漏洞在哪里，更不会分析算法复杂度的瓶颈或建议优化方向。这种缺乏诊断深度的反馈，使得学生容易陷入盲目试错和重复提交的循环，不利于培养系统性的调试思维和对算法性能的深刻理解。

（3）个性化与自适应能力的缺失

主流平台虽然具备进度追踪和排名功能，但本质上仍是一种“一刀切”的静态资源库。系统无法根据学生的历史表现（如错误模式、知识薄弱点、解题速度）动态调整题目推荐顺序和难度，所有学生面对的是同一套题目序列，能力强者可能觉得挑战不足，能力弱者则容易产生挫败感。

（4）教学管理功能的不足

在支持完整的教学周期方面，通用 OJ 平台功能薄弱。例如，自动化组卷功能缺失，教师需要手动挑选题目来组建期中、期末考试试卷，难以保证试卷在知识点覆盖和难度分布上的科学性与均衡性。此外，平台与高校教务系统、在线课程平台（如学习通、MOOC）的集成度低，数据无法互通，形成了“信息孤岛”，不利于教学数据的统一分析和教学过程的连贯管理。

三、智能评测技术驱动的系统优化与教学创新

为克服传统 OJ 系统的局限，将生成式人工智能、机器学习、深度学习等智能技术融入评测系统，构建教育专用在线智能评测平台，已成为明确的演进方向。其优化主要体现在题库管理、评测反馈、学习路径和系统架构四个层面。

（一）智能化题库管理与动态组卷

智能化题库管理首先强调题目标签的精细化与多维化。每道题目不仅标注算法类型，还需关联到具体的课程章节、核心知识点（如“指针与数组关系”、“递归边界条件”）、认知目标（记忆、理解、应用、分析）、应用场景等。其次，题目难度不再是静态设定，而是根据历史学生答题的正确率进行动态校准，形成更符合学生实际认知水平的难度梯度。

在此基础上，系统能够支持强大的动态组卷功能：（1）固定组卷：教师设定知识点分布、难度比例、题型数量等约束条件，系统自动生成符合要求的试卷，并支持生成 A/B 卷；（2）随机组卷与自适应组卷：在平时练习或测试中，系统可为每位学生随机抽取题目，或根据其当前实操能力画像，动态匹配合适难度的题

目，实现“千人千卷”，从根本上杜绝答案共享。

（二）深度评测反馈

智能评测技术的核心突破在于将反馈从“是什么错”推进到“为什么错”以及“如何改进”。例如设置“知识重难点智能预警机制”功能：当学生提交错误代码时，系统能自动识别题目考察的核心知识点，并关联该生的历史错误记录，判断其对该知识点的掌握程度。随后，系统可向学生智能推送相关的理论讲解微视频、典型错误案例分析等学习资源。同时，系统以“预警”形式将班级整体的高频错误知识点、常见错误模式等数据实时可视化给教师，为课堂上的集中讲解和个性化辅导提供精准导航。

更前沿的探索是引入大语言模型（LLM）提供代码层面的具体指导。研究显示，基于 GPT-4 的智能反馈系统能在不直接给出答案的前提下，指出代码中的逻辑错误、语法陷阱甚至提出优化建议。^[5] 尽管存在“幻觉”（生成错误信息）的风险，但经过校准和与规则系统结合后，其作为智能编程助手的潜力巨大。基于聊天机器人的个性化反馈系统（如 PythonPal）也在语法纠错和意图图理解上展现出应用价值。^[6]

（三）数据驱动的个性化学习路径构建

智能评测系统通过持续收集和分析学生的学习行为数据（如提交次数、首次通过率、在各知识点的耗时、错误类型分布），为每个学生构建动态的能力画像和知识图谱。基于此画像，系统可以实现：

- （1）个性化推荐：自动推荐适合学生当前水平的“跳一跳够得着”的题目，避免过难或过易。
- （2）薄弱点追踪与强化：自动生成专属错题本，并定期推送同类题目进行巩固练习。
- （3）学习路径规划：为学生规划从基础到进阶的个性化学习序列，并提供进度预测和风险预警（如预测挂科风险）。

数据显示，采用智能推荐的学习路径，能使学习效率提升 3.5 倍，学习计划执行率从 50% 提高到 90%，能力画像准确率可达 85%。

表 2 对比了传统题库与数字化智能题库在关键维度的性能差异，结果表明数字化智能题库在管理效率、数据分析深度以及最终的学习效果支持上全面超越传统模式，尤其在提升学习效率和计划执行力方面优势显著，这印证了智能化是题库系统发展的必然方向。

维度	传统题库	数字化智能题库
管理方式	手工 / 文档管理	平台化 / 自动化管理
题目标签	无或简单标签	多维标签（知识点、难度、题型等）
难度分级	固定、主观	支持动态校准与科学分级
数据分析	基本无或简单统计	多维深度统计与学习分析
推荐机制	无或静态序列	基于画像的智能动态推荐
学习效率提升	基准（1.0 倍）	3.5 倍
学习计划执行率	约 50%	90%

能力画像 准确率	约 40%	85%
-------------	-------	-----

表 2 传统题库与智能题库核心性能对比

（四）微服务化与云原生智能 OJ 系统架构

为应对高并发提交、保障评测安全、提升系统可维护性，新一代教育专用智能 OJ 系统正转向微服务架构和云原生技术。通过将用户管理、题库服务、代码评测、任务调度、数据统计等模块解耦，系统可以获得更好的弹性伸缩能力和稳定性。评测核心则普遍采用 Docker 容器沙箱技术，将学生代码在完全隔离的环境中运行，防止恶意代码对宿主系统造成破坏，同时能精确控制运行时间和资源（CPU、内存）消耗。Serverless（无服务器）架构的引入，可以进一步实现资源的极致弹性，将评测环境初始化时间大幅缩短，满足大规模在线考试等高并发场景的需求。

四、挑战与展望

如前分析，尽管智能评测技术为程序设计类课程教学带来了革命性变化，但其深入应用仍面临诸多挑战，主要体现在：①评价维度依然存在片面性，对于代码的可读性、可维护性、模块化设计、注释规范性等软件工程核心素养，以及算法设计的创新性，仍缺乏有效的自动化评估手段。②由人工智能驱动的评分和推荐算法，其训练数据可能隐含偏见，对非标准但同样正确的创新解法评价不高；同时过度依赖系统数据画像，可能忽视学生的

努力过程、思维闪光点等难以量化的因素，造成新的评价不公。③对教师能力的更高要求，教师需要从知识的单向传授者，转变为学习数据分析师、个性化学习路径设计师和深度思维引导者。如何解读复杂的数据报告、如何基于系统预警进行有效的课堂干预、如何设计人机协同的教学活动，是教师面临的新挑战。④如果教学设计和引导不当，智能题库和推荐系统可能使学生陷入“追逐通过率”和“刷题积分”的陷阱，反而阻碍了深层计算思维和问题解决能力的发展。

面向未来，OJ 系统与程序设计教学的融合将向更深层次、更广维度发展：① OJ 系统将进一步整合静态代码分析和动态测试技术，同时联合自然语言处理技术（用于分析代码注释和提交日志中的问题描述）构建一个多维度的综合评价体系，不仅评判“对不对”，还要评价“好不好”、“美不美”。②生成式 AI 的深度学习不仅是智能反馈工具，更可以扮演“永不疲倦的编程教练”角色。③ OJ 系统将突破纯算法题的范畴，与硬件仿真、数据库操作、网络通信等更广泛的计算机科学实践相结合，使其成为支撑项目驱动学习和跨学科综合实践的核心平台。④由智能评测平台、教师、学生共同体三方共同构建人机协同的智慧教学新生态的智慧教学生态：平台负责提供个性化的练习资源、实时诊断和基础反馈；教师基于平台提供的宏观数据洞察和微观个体预警，进行高价值的情感交流、思维拔高和创造性引导；学生共同体则在协作与竞争中共同成长。

参考文献

[1] 邓泽林, 黄园媛, 罗奕. 基于《知识融合 + 能力培养》的程序设计类课程教学模式改革研究 [J]. 计算机教育, 2026.

[2] 徐燕萍. 基于 OJ 的程序设计类课程实验混合教学模式研究 [J]. 软件导刊, 2022, 21(02): 231-234.

[3] 刘勇, 田凯, 周晓琳, 等. 以 OJ 系统和学科竞赛为核心的程序设计实践教学 [J]. 高教学刊, 2021(6): 4.

[4] 秦汉林, 卢成浪, 邹会来. OJ 系统和在线课程混合驱动的程序设计课程教学新模式 [J]. 计算机教育, 2023(6): 189-193.

[5] Sven Jacobs, Steffen Jaschke. Evaluating the Application of Large Language Models to Generate Feedback in Programming Education. IEEE Global Engineering Education Conference, 2024.

[6] Sirinda Palahan. PythonPal: Enhancing Online Programming Education through Chatbot-Driven Personalized Feedback. Human-Computer Interaction, 2025.