

文章编号: 2095-2163(2020)10-0108-04

中图分类号: TP312

文献标志码: A

面向初学者的 C 语言在线实验系统设计与实现

林 渤, 薛 斌, 林逸滔, 胡建鹏

(上海工程技术大学 电子电气工程学院, 上海 201620)

摘 要: 针对高校选修 C 语言程序设计课程人数普遍较多, 传统授课方式缺乏效率, 指导教师无暇帮助所有学生等问题, 本文设计和实现了面向初学者的 C 语言在线实验系统。学生可以通过浏览器进行在线 C 语言实验来提高学习效率, 并可通过错误分析来自助、互助地解决程序问题, 从而降低老师指导实验的工作压力。通过易用的 web 系统界面, 学生可以随时随地通过电脑或者手机学习编程。实际应用效果表明, 本系统为 C 语言程序设计课程的在线教学提供了重要支撑。

关键词: 程序设计; 实验系统; 错误分析; 在线教学

Design and implementation of C language online experiment system for beginners

LIN Bo, XUE Bin, LIN Yitao, HU Jianpeng

(School of Electronic & Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] To solve the problems such as the large number of students taking C language programming courses in universities, the inefficiency of traditional teaching methods, and the absence of instructors to help all students, in this paper, we design and implements an online experimental system of C language for beginners. Students can conduct online c language experiment through the browser to improve the learning efficiency, and can help solve the program problems through error analysis by themselves, so as to reduce the pressure of the teacher. With an easy-to-use web interface, students can practice programming by using their computers or mobile phones anytime, anywhere. The practical application during the outbreak of COVID-19 showed that this system provided important support for online teaching of the programming course with C language.

[Key words] Programming design; Experimental system; Error query; Online teaching

0 引 言

随着社会不断发展, 对于以应用型人才培养为办学目标的众多本科院校, 如何在学生人数逐年增加、上机实验课指导老师无暇辅导所有学生的情况下提高教学质量, 已成为了 C 语言程序设计课程迫在眉睫需解决的问题^[1]。目前程序设计课程的教学培养目标多是学习程序设计的思想和方法, 通过大量实践编程练习来掌握程序设计的流程与技巧, 只有通过大量的练习才能了解和避免程序中常出现的 Bug。目前已有很多此类系统, 如 PTA 辅助教学平台, 侧重于在线考试, 交互性不强, 在教学方面无法针对于每一位同学进行错误分析^[2]; Leetcode 平台侧重于有一定基础的学生进行算法练习, 学生不能主动提交各类错误报告, 系统也无法收集学生详细的错误信息^[3]; codecode.net 平台适合于计算机专业的学生, 专业程度相对较高, 初学者不易掌握^[4]。为此, 本文设计和开发了更有针对性的在线实验系统, 面向非计算机专业的初学者用户, 简洁易用, 可

快速上手, 并提供错误分析和查询功能, 实现了学生的自助和互助; 同时进行实验在线提交并生成实验报告, 教师可以查看各种实验数据, 关注学生的学习进展。

1 系统设计

1.1 系统架构

本系统使用 B/S 架构, 采用三层结构。B/S 架构的跨平台性好, 能够做到一次开发, 处处运行。从结构上, 系统可以分为视图层、业务层和持久层三层。功能模块上可以分为系统管理模块, 在线编译执行模块, 自动评分模块, 实验报告生成提交模块, 错误采集查询模块等五大模块。系统架构图如图 1 所示。

1.2 数据库设计

系统主要实现了程序在线编译执行、学生自助评估、错误分析与查询、实验报告生成等功能。数据库的实体关系图如图 2 所示。其中, 学生表用来存放学生的信息; 教师表用来存放教师的信息; 错误分

作者简介: 林 渤(1999-), 男 本科生, 主要研究方向: 计算机科学与技术; 薛 斌(1994-), 男, 硕士研究生, 主要研究方向: 软件工程; 林逸滔(1998-), 男 本科生, 主要研究方向: 计算机科学与技术; 胡建鹏(1980-), 男, 博士, 副教授, 主要研究方向: 软件工程、系统工程。

通讯作者: 胡建鹏 Email: mr@sues.edu.cn

收稿日期: 2020-05-06

析表记录了学生提交的各种错误信息;实验小结表存放实验小结及实验报告上传信息;实验题目表用来存放每一个具体实验题目的信息,多个实验题目组合成一次实验,一个课程则可发布多次实验;测试用例表用来存放测试用例信息,一个实验题目可有多个测试用例;而实验提交表用来存放代码提交信息,用于生成实验报告。

2 特色功能介绍

(1)在线编译执行,学生自助评估。系统给学生端用户提供了在线编译执行代码的功能,支持网页端用户输入的交互功能,用户可以随时随地在浏览器中编写执行代码,还可以对编译和执行过程进行时间控制并加以提示,防止用户执行非法代码,同时对编译执行的结果进行解析,做到与 Dev-C++的编译信息一致,在此基础上增加了解析字段,帮助学生更好的理解报错信息。执行代码后,系统会给出

由输入输出结果相似度(输入输出结果与输出样例之间的编辑距离/输出样例字符串长)的评估结果,给出与测试用例的相似度。

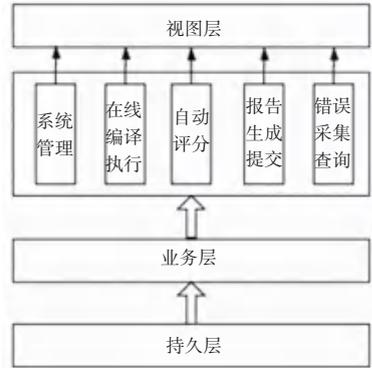


图 1 系统架构
Fig. 1 System architecture

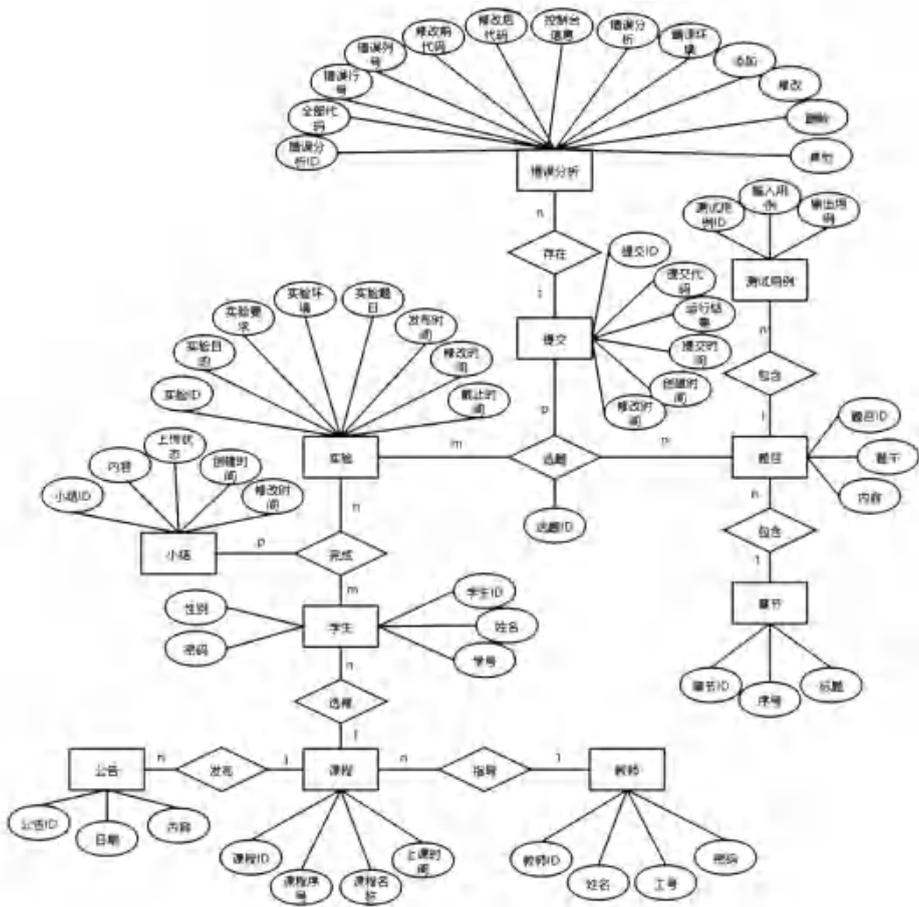


图 2 实体关系图

Fig. 2 Entity relationship diagram

(2)错误采集、错误分析及查询功能。在实验提交界面集成了错误采集功能,学生可以在调试程

序时将错误信息填入到表单中完成错误采集,包括语法错误、逻辑错误和运行时错误。系统支持学生

提交多个错误分析,并且为了优化学生填写表单时的操作,还提供了一系列的快捷按钮,如可以将第一个语法错误自动填入表单中,减少手动填写的操作。所有采集的这些错误信息经过整理后形成常见错误数据库,基于这个数据库给学生提供错误查询功能。

(3)实验提交和报告生成。在生成报告页面,学生可以预览实验报告。在第一次生成报告时,系统会自动跳转至报告底部的实验小结,学生按照要求填写实验心得,默认为PDF文件,也可打印。确认无误后可以上传至教师端,文件会自动命名为学号-姓名-班级-实验名.pdf。

(4)教师端功能。系统提供基础的登录和注销功能,教师可以对实验、题目、班级、学生、公告等信息进行管理。根据教学需要对发布的实验以及题库增删改查。根据需要及时发布公告,查看学生代码的提交情况以及实验报告的提交记录。也可将发布的一个实验复制到其他班级,减少教师需要同步发布多个实验的操作。对于新添加的班级,系统支持批量导入学生名单,只需将学生名单的xls文件上传即可完成批量导入。同时系统提供了阅读模式,可以按照实验阅读,也可以按照学号阅读,反映班级和学生的学习情况。

3 系统实现与部署

本系统采用云端部署的方式,缩短了部署、配置、调试的时间,不需要运维人员以及硬件服务器,减少了成本,满足了同时服务更多学生的需求,以达到预期的教学效果。下面介绍一些关键的技术和算法:

(1)前端WEB界面。WEB界面使用Thymeleaf模板引擎渲染。CSS库选择了支持响应式的Bootstrap,JS库选择了jQuery,能够加快前端开发的速度。在线编辑器使用CodeMirror插件,它支持代码缩进、高亮显示等功能,并且提供了十分丰富的API供用户使用。

(2)后台编译运行。代码编译执行的交互功能使用Ajax技术,可以在不重新加载整个网页的情况下刷新局部信息。代码和其他信息提交到网站之后,程序会将代码部分保存成*.c文件,输入部分保存为*.txt文件,然后调用系统命令单独打开一条线程执行编译命令。编译成功后创建一个线程池,将需要执行的任务加入到线程池中,同时执行时间控制程序,限制运行时间。运行成功后执行自动评分业务,对输入输出结果与测试用例比对得到评估结果,将所有信息返回给网页。提交实验时会将所

有信息写入数据库后返回提示信息。后台执行编译时还要注意安全问题。安全方面包括依赖环境构建、沙盒、行身份(包括GID、UID),执行时间限制包括运行时间和CPU时间,文件系统压力等一系列安全问题。

(3)执行结果自助评估。使用levenshtein distance(又叫编辑距离)算法进行相似度计算,实现执行结果自助评估。编辑距离是指两个字串之间,由一个转成另一个所需的最少编辑操作次数。编辑操作包括将一个字符替换成另一个字符,插入一个字符,删除一个字符。一般来说,编辑距离越小,两个字串的相似度越大。算法基本原理:假设使用 $d[i, j]$ 个步骤(可以使用一个二维数组保存这个值),表示将串 $s[1 \dots i]$ 转换为串 $t[1 \dots j]$ 所需要的最少步骤个数,那么在最基本的情况下,即在 i 等于0时,也就是说串 s 为空,那么对应的 $d[i, j]$ 就是增加 j 个字符,使得 s 转化为 t ,在 j 等于0时,也就是说串 t 为空,那么对应的 $d[i, 0]$ 就是减少 i 个字符,使得 s 转化为 t 。

4 应用效果

经过将近一个学期的运行,在线实验系统取得了较好的教学效果。通过对两个多月内所得到的数据进行分析,对系统收集到的所有提交的数据进行整理,筛选出其中的有效数据进行分析后发现,提交的数据中能最终运行成功的比率在逐步提高,如图3所示。

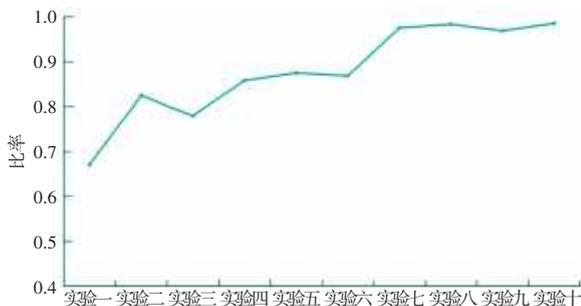


图3 提交成功率趋势图

Fig. 3 Trends in submission success rates

通过对学生们提交的错误进行分析后发现,错误类型仍大多集中在语法错误上,错误类型柱状如图4所示。而其中最主要的语法错误,在总体上也呈减少趋势,如图5所示。

通过对系统中实验问题的发布时间和收集的每一个学生用户的提交时间进行整理、分析,得出系统的平均耗时趋向,如图6所示。一般会在课上布置4个左右实验题,大多数同学是无法在半节课里全

部完成并提交实验报告的,需要他们课下继续完成,图 6 显示学生提交报告总的花费时间也在波动中逐渐减少,虽然不是完成实验的净时间,但这个趋势在一定程度上也说明学生完成实验任务的效率是越来越高的。



图 4 错误类型柱状图

Fig. 4 Error type histogram

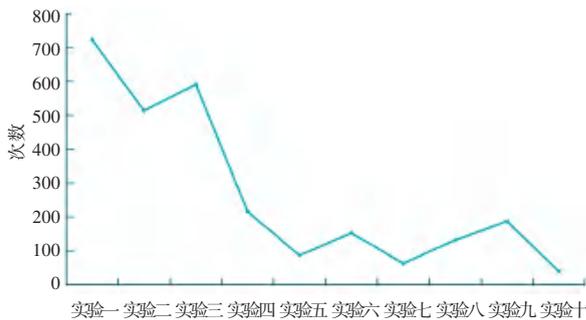


图 5 语法错误数量趋势图

Fig. 5 Trend in the number of grammatical errors

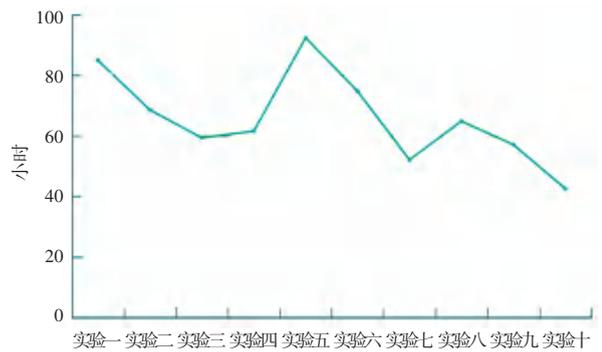


图 6 课外实验平均耗时

Fig. 6 Average time spent on extracurricular experiments

5 结束语

本文设计的 C 语言在线实验系统主要面向非电专业学生及各类程序设计的初学者,旨在将实验过程数字化、在线化,形成学生自助互助的网上学习社区。目前在线编译执行和辅助查错等模块已经完成并投入实际使用。通过使用本系统,大大提高了学生的学习效率,减轻教师的教学负担。在未来的工作中,将继续优化辅助纠错功能,增加代码静态分析纠错功能,语法错误以外的很多错误很难通过动态编译执行去判断,需要引入代码静态分析技术来实现逻辑错误和运行时错误的辅助分析。

参考文献

[1] 张曙光,刘英,周雅洁,等. C 语言程序设计[M]. 北京:人民邮电出版社,2014.
 [2] PTA. <https://pintia.cn/>
 [3] LeetCode. <https://leetcode.com/>
 [4] CodeCode. <https://www.codecode.net/>

(上接第 107 页)

不同方法采用相同的训练集和测试集。二个对比实验以及本文方法实验结果见表 4,可以看出本文从评论内容和评论用户二个视图建立基分类器协同训练的方法是有效的。

表 3 不同分类器组合识别结果

Tab. 3 Recognition results of combination of different classifiers

分类器组合 (h_1, h_2)	测评指标		
	R	P	F1
LR, SVM	0.937	0.918	0.927
SVM, LR	0.878	0.862	0.870
SVM, SVM	0.856	0.791	0.822
LR, LR	0.907	0.879	0.892

表 4 不同方法的识别结果对比

Tab. 4 Comparison of recognition results of different methods

方法	垃圾评论			总体准确率 Accuracy
	R	P	F1	
方法 1	0.911	0.950	0.930	0.976
方法 2	0.901	0.973	0.936	0.885
本文方法	0.936	0.944	0.940	0.979

4 结束语

本文针对之前微博垃圾评论识别研究时忽略了评论用户特征这一缺点,从评论内容和评论用户二个视图构建基分类器进行协同训练。实验结果表明,本文的方法是有效的,对于垃圾评论的识别取得了良好的效果。但是对于短小评论的识别效果仍然不佳,下一步将考虑提高只有几个文字的短小评论的识别效果。

参考文献

[1] 丁兆云,贾焰,周斌. 微博数据挖掘研究综述[J]. 计算机研究与发展,2014,51(4):691-706.
 [2] 黄铃,李学明. 基于 AdaBoost 的微博垃圾评论识别方法[J]. 计算机应用,2013,33(12):3563-3566.
 [3] 李志欣,兰丹媚,张灿龙,等. 基于 Co-Training 的微博垃圾评论识别方法[J]. 计算机工程,2018,44(7):212-218.
 [4] 姚子瑜,屠守中,黄民烈,等. 一种半监督的中文垃圾微博过滤方法[J]. 中文信息学报,2016,30(5):176-186.
 [5] 杜茂康,叶琪. 基于 PCA 与协同训练算法的虚假评论识别研究[J]. 计算机仿真,2019,35(2):452-457.