

文章编号：

中图分类号：G642

文献标识码：B/A

# MOOC 模式下数据库系统概论课程实验平台的建设与实践

卢 卫<sup>1,2</sup>，程一舰<sup>1,2</sup>，赵展浩<sup>1,2</sup>，杜小勇<sup>1,2</sup>

(1. 中国人民大学 数据工程与知识工程教育部重点实验室，北京 100872；

2. 中国人民大学 信息学院计算机系，北京 100872)

**摘 要：**数据库系统概论作为实践性较强的计算机专业课，如何支持在线实验，如何自动、准确、快速地量化学生对相关知识点的掌握程度，是实施 MOOC 教学的关键挑战。文章介绍数据库在线实验系统的设计与实现，并探讨如何通过课程实验平台的建设与实践来有效支撑 MOOC 教学。

**关键词：**MOOC；在线实验；数据库系统概论；自动评分；数据分析

## 0 引言

MOOC (Massive Open Online Course, 大规模开放在线课程) 的出现，给高等教育的发展带来了新机遇。优质且丰富的在线课程资源、低廉的学习成本、灵活而自主的学习方式、多样的互动交流方式等，使 MOOC 课程大受欢迎，并得到了人们的广泛参与。例如，中国人民大学数据库教学团队在中国大学 MOOC 平台上<sup>[1]</sup>开设的数据库系统概论 MOOC 课程 (包括基础篇、高级篇和新技术篇 3 部分)，自 2016 年 4 月开始选课人数已累计 9 万余人。然而，MOOC 教学仍然面临诸多挑战，如学习者的高辍学率和低完成率、专业课程的在线实验如何实施等。本文聚焦数据库系统概论的实验教学要求，分析在线实验的难点，设计在线实验平台，并开展 MOOC 和 SPOC 教学实验，验证在线实验平台的有效性。

## 1 数据库系统概论课程实验平台的建设

### 1.1 课程实验简介

数据库系统概论全面地阐述了数据库系统的基础理论、基本技术和基本方法。作为实践性较强的一门计算机专业课，数据库系统概论课程的培养目标不仅要求学生具备复杂工程问题建模 (数据库设计) 所需的理论知识 (关系数据理论)，还要求学生掌握计算机编程与软件开发相关的工程基础和专业知识 (包括关系数据库标准语言 SQL、数据库安全性与完整性、关系查询处理与查询优化、数据库恢复技术、并发控制)，最终能够应用数据库的专业知识进行复杂问题的求解。为了实现这一培养目标，除了平时的课堂教学之外，笔者还需要建设课程实验，强化学生对基本技术、基本方法的实践，同时也为教师了解学生对知识的掌握程度提供教学上的参考。

与配套的教材<sup>[2]</sup>、习题解析与实验手册<sup>[3]</sup>和 MOOC 在线课程<sup>[4,5]</sup>内容相对应，笔者从数据库的创建、维护、查询处理与优化 3 个角度出发，设计了①数据操纵语言、②数据控制语言、③数据定义语言、④存储过程与触发器的管理、⑤索引的创建、⑥事务、⑦数据一致性等 7 个实验，这 7 个实验构成了数据库系统的基本技术体系。其中，关系数据库的创建包括数据操纵语言、数据控制语言和数据定义语言 3 个实验；查询处理与优化包括存储过程与

触发器的管理、索引创建、数据一致性 3 个实验；关系数据库维护包括事务实验。

## 1.2 实验考核的内容

针对上述 7 个实验，根据课程的教学要求，笔者设计了表 1 的考核内容。

表 1: 实验的考核内容

名称	考核内容
数据操纵语言	使用 SQL 语句完成基本表中数据的插入、更新、删除；使用 SQL 语句完成对基本表的查询，包括带条件达式的查询、自连接查询、多表连接查询、嵌套查询和集合查询。
数据控制语言	使用 SQL 语句完成对基本表的查询、插入、更新、删除权限的分配、回收，及级联操作。
数据定义语言	使用 SQL 语句完成带有表级、列级完整性约束的基本表的创建、修改、删除。
存储过程与触发器	掌握数据库存储过程和触发器的设计和使用方法，能够使用 SQL 语句定义行级和表级的触发器，实现较为杂的应用逻辑。
索引	给定数据库的模式和具体的应用需求，创建和删除辅助索引，能够设计 SQL 语句来验证索引的有效性。
事务	掌握数据库事务管理的基本原理以及事务的编程方法，能够根据具体的应用需求，编写带有判断条件的存储过程，把事务的编写和存储过程的设计和使用结合在一起；当条件不成立时，回滚事务，否则，提交事务。
数据一致性检验	掌握数据库并发控制的基本原理及其应用方法，验证并发操作在不同隔离级别下的数据不一致性问题（包括丢失修改、不可重复读和读脏数据等情况），通过 SQL 语句展示不同封锁级别下的并发控制效果。

## 1.3 实验自动评价的挑战与解决方案

以往的实验课程，学生撰写实验报告，教师审阅报告内容，实验的评价主要依据实验结果的截图，这种纯手工、易主观、粗粒度的评价方式，无法满足 MOOC 教学模式下教师对自动、准确、快速地量化实验结果的要求。实验自动评价的挑战具体表现在：

**难点 1:** 解决同一个问题的 SQL 语句存在多种写法，系统无法使用基于模式匹配的方式进行结果的正确性判断。例如，在学生选课系统中，查询学号为“2017001212”学生所选修的课程名，使用如下语句均可获得正确的运行结果。

- A. `SELECT CName FROM Student, SC, Course WHERE Student.SNo = SC.SNo AND SC.CNo = Course.CNo AND Student.SNo = '2017001212'`
- B. `SELECT CName FROM SC, Course WHERE SC.CNo = Course.CNo AND SC.SNo = '2017001212'`
- C. `SELECT CName FROM Course WHERE CNo IN (SELECT CNo FROM SC WHERE SNo = '2017001212')`

**难点 2:** 不同用户操作同一个数据对象可能存在冲突或相互影响。例如，实验考核内容要求用户创建学生、课程、和学生选课 3 个关系模式。这样，不同的学生会创建同名的数据对象。除了第一个用户创建成功之外，其他用户均创建失败。很明显，相同的语句，不同的人执行结果的正确性不同，这与实验的初衷相违背。此外，基本表中记录的插入和删除操作，会影响其他用户对该基本表的操作。当要求用户插入一条记录时，由于所有用户均插入内容相同的记录，除了第一个用户操作成功之外，其他用户操作均失败。为了解决这个问题，需要设计不同用户之间操作的隔离性机制。

**难点 3:** 用户操作需要保持状态延续性。在课程实验考核内容中，不同的实验存在互相依赖关系。例如，记录的增、删、改操作，需要用户事先完成基本表的创建；数据查询要求表中存在一定的数据量。因此，在线实验平台的设计要求能够物化用户每一个阶段的操作结果。

**难点 4:** 并发控制实验需要避免系统频繁的重启问题。并发控制实验设计的目标是为了验证并发操作在不同隔离级别下数据的不一致性问题。需要注意的是，每次为数据库设置不同的隔离级别，都要重启数据库管理系统，而系统的每次重启，都会造成在线实验平台服务的不可用。因此，频繁重启系统会影响用户的使用体验，应尽可能避免。

针对实验建设过程中遇到的上述 4 个难点，笔者提出了如下的对应解决方案。

针对**难点 1**，笔者提出了基于**结果集比对**的解决方案。虽然同一个问题的 SQL 语句存在多种写法，但其执行结果是唯一的。为此，笔者事先录入正确的 SQL 语句，并对用户提交 SQL 语句与正确 SQL 语句，两者的执行结果进行比对，如果比对一致，则表明用户提交的 SQL 语句是正确的。需要注意的是，必须保证实验的执行结果不为空，否则，就有可能造成误判。此外，笔者也考虑到学生们在做题过程中可能存在的失误，也提供了一定程度的容错处理。例如，在某个题目中，要求学生编写 SQL 语句返回计算机系学生的英语成绩，并按成绩降序排序。而学生在提交 SQL 语句时，虽然结果集中能够正确包含计算机系学生的英语成绩，但成绩未按降序排序。在这种情况下，虽然学生并未提交完全正确的答案，但其总体思路明确，应该给予一定的分数。为了实现这一要求，事先录入学生可能提交的答案，并对每个答案设置一定的分数。当学生提交 SQL 语句的执行结果，跟系统事先录入某个语句的执行结果一致，则按照事先设置的分数进行评分，从而保证了系统具有一定的容错性。

针对**难点 2**，笔者提出了**基于模式 (schema) 的操作隔离机制**。数据库中模式的作用类似于文件系统中文件夹的作用。同一模式下，数据对象的命名是唯一的，而不同模式下，数据对象可以允许相同的命名。为此，为每个用户创建各自的模式，其以后针对基本表、索引、视图、存储过程与触发器的创建、修改、删除都在其对应的模式下操作。因此，不同用户之间对上述 5 类数据对象的创建、修改、删除具有隔离性。通过查询数据字典，判断用户是否成功创建了相应的数据对象，并检查该对象是否满足事先设置的创建条件，进行相应的打分。

针对**难点 3**，笔者提出了**基于事务回滚**的机制。为了保持用户操作状态的延续性，一方面，对于基本表、索引、视图、存储过程与触发器的创建、修改、删除操作，利用基于模式的操作隔离机制，保证不同用户之间操作不互相干扰；另一方面，对于基本表中记录的增、删、改、查操作，提出了基于事务回滚的的机制。具体地说，根据实验内容的考核要求，事先创建相应的基本表，并插入一定数量的记录，对于数据操纵语言实验，所有用户对记录的增、删、改、查操作，全部都在这些基本表上操作。这样设计的好处是避免每个用户重复导入数据，从而节省存储和计算资源。然而，之前提到，同一个基本表中记录的插入和删除操作，会影响其他用户对该基本表的操作。为了解决这一问题，笔者作了如下的设计：用户对基本表的插入和删除操作，只做正确性判断，但不会真正地往基本表中插入或删除实际的数据，从而保证任何用户执行操作之前和之后，数据库中的数据是一致的。为了实现这一设计，引入事务的概念，把用户的操作封装在同一个事务中，在事务中对用户的操作进行正确性的判断，在事务结尾进行回滚，使用户的操作不会对数据库造成任何影响，从而保证用户操作的隔离性。

针对**难点 4**，笔者事先构建不同隔离级别下的数据库，并在在线实验平台的客户端创建多个窗口，每两个窗口连向同一个数据库（具有相同隔离级别），从而测试在某一个隔离级别下的数据一致性检验。

## 2 课程实验的实践

根据实验的考核要求，笔者开发了在线实验平台<sup>[6]</sup>辅助 MOOC 教学。

学生必须填写相关信息并注册到某一个教学班中；相关教师通过审核后，该学生可以登录系统，进行相关知识点题目的选择和作答。平台为每个实验内容生成一个或多个功能模块，每个功能模块覆盖部分或全部的实验内容。例如，在数据操纵语言实验中，设置了 3 个功能模块，分别是单表查询、多表查询、和数据更新。在单表查询和多表查询中，分别给定了具体的应用场景，并各自设计了包含 9 道题的实验案例。要求学生能够分析用户查询要求，设

计带有条件表达式和集合统计的单表查询和带有条件的自连接查询、多表连接查询、嵌套查询和集合查询（多表查询）。当学生进行相应题目的作答并提交后，实验平台进行实时、自动打分，并将最终的判分结果返回给学生。图 1 给出了多表查询的应用场景及第一道题的题目描述。



图 1 多表查询功能模块

教师负责审批教学班学生的加入、管理以及分析学生及其答题情况。教师端的功能可以分为审核、分析、导出统计的答题或行为信息 3 个模块（在本小节中，仅介绍审核模块，分析和导出模块将在第三节中介绍）。教师负责审核本班学生注册申请请求，通过查看学生提交的实名注册信息，如果是属于本班同学，则给予通过，否则，拒绝学生的申请（参见图 2）。

29	韩梦	韩梦	1341094840@qq.com	已通过
30	2015202008	水泊禹	shui_zhiyu@163.com	已通过
31	2015201613	lij	522278689@qq.com	已通过
32	outhunder@126.com	吴雷	outhunder@126.com	已通过
33	stdu20151687曹阳阳	曹阳阳	1774176709@qq.com	已通过
34	2016100850	chenrui	645029552@qq.com	通过 拒绝

图 2 教师审批模块

### 3 课程实验的分析

通过收集学生在答题过程中的行为和提交的答案，系统给教师提供本班学生总体和个体的分析，并进行可视化展示。

#### 3.1 学生答题数据的收集

学生答题数据包括学生在答题过程中的行为和提交的答案。学生的答题数据发生在学生提交自己的实验答案之后，平台对学生提交上来的答案以及经过判断之后的判分结果进行存储。学生的行为数据发生在学生答题的过程之中，平台采集学生在每一道题目中的思考停留时间并存储。

#### 3.2 学生答题情况的分析

学生答题情况分析可分为整体分析和个体分析，如图 3 所示。

通过整体分析，教师可以查看本班学生在每个实验内容的完成情况：（1）有多少学生提交了答案，有多少学生已经开始答题、但尚未提交答案，有多少学生尚未开始答题；（2）

班级的平均成绩及方差，哪些学生成绩高于平均成绩、哪些学生低于平均成绩；（3）对于每一道题，哪些学生回答对了，哪些学生回答错了，哪些学生部分回答正确；（4）班级学生的平均完成时间，哪些学生完成时间低于平均完成时间，哪些学生高于平均完成时间。

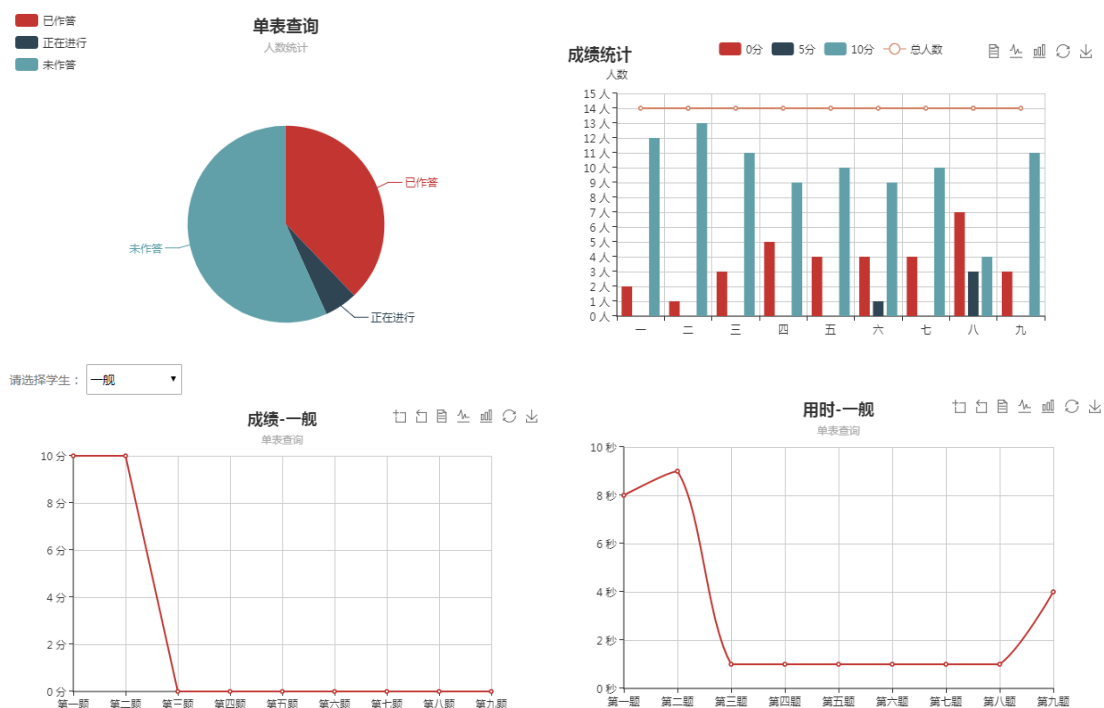


图 3 实验统计分析图

通过个体分析，教师可以查看本班每个学生的完成情况，特别是其中提交错误答案的学生个体。例如，在图 3 的成绩统计柱状图中，通过点击“0 分”红色柱子，可以查询得零分的学生姓名。在此基础上，通过查询该学生，获得该学生的答题详情，分析出该学生答错题的原因。另外，笔者也可以查询学生完成每个实验和每道题的用时。

通过整体和个体分析，教师可以细粒度地获取学生答题情况信息：哪些题学生得分普遍偏低，哪些题学生用时较多，哪些学生学习态度不端正（包括不提交作业以及每道题出错且用时明显低于平均时间），哪些学生学习成绩不甚理想。对于得分较低、用时较长的题目，要进行重点讲解，对于学习态度不端正、学习成绩不理想的学生，可以根据其错误的原因，有针对性地讲解，真正地做到因材施教。

#### 4 结语

中国人民大学数据库团队自主开发的数据库系统概论课程在线实验平台已经上线，并在数据库系统概论 MOOC 建设与应用第二届研讨会、第二届面向计算机类 MOOC 的大规模在线学习支撑工具研讨会中，向来自全国各个高校从事数据库教学的 200 多位一线教师进行了介绍，得到了大家的积极反响。平台目前已经得到了来自石家庄铁道大学、北京信息科技大学、重庆科技大学等 10 多所高校的使用申请。由于准备时间比较仓促，在实施 MOOC 教学模式时，笔者采用边探索边建设的模式，功能上将继续完善改进，也请同行们批评指正。

#### 5. 致谢

中国人民大学数据库系统概论教学小组的王珊老师、陈红老师、博士生王童童、人大金仓信息技术股份有限公司的冷健全参与了课程实验的设计。

## 参考文献

- [1] 中国大学 MOOC. 首页[EB/OL]. [2017-09-08]. <http://www.icourse163.org/>.
- [2] 王珊, 萨师煊. 数据库系统概论[M]. 五版. 北京: 高等教育出版社, 2014.
- [3] 王珊, 张俊. 数据库系统概论习题解析与实验指导[M]. 5 版. 北京: 高等教育出版社, 2015.
- [4] 王珊, 杜小勇, 陈红. 数据库系统概论基础篇[EB/OL]. [2017-09-08]. <http://www.icourse163.org/course/RUC-488001>.
- [5] 王珊, 杜小勇, 陈红. 数据库系统概论基础篇[EB/OL]. [2017-09-08]. <http://www.icourse163.org/course/RUC-1001655006>.
- [6] 卢卫, 杜小勇. 数据库系统概论在线实验平台[EB/OL]. [2017-09-08]. <http://dbm.ruc.edu.cn/rucdbcourse/>.

基金项目：国家自然科学基金青年科学基金项目（61502504）；教育部 2016 年课程教材改革研究项目“在线开放课程大数据研究项目（一期）”“MOOC 课程有效学习的分析与对策研究”；中国人民大学教师教学发展改革项目“基于数据库 MOOC 教学的在线实验平台研究”。

第一作者简介：卢卫，男，副教授，研究方向为云计算与大数据管理、空间与文本数据库管理、索引技术，[lu-wei@ruc.edu.cn](mailto:lu-wei@ruc.edu.cn)。



杜小勇（通信作者），男，教授，研究方向为数据管理技术、语义网技术、智能信息检索技术，[duyong@ruc.edu.cn](mailto:duyong@ruc.edu.cn)。



通讯作者： 杜小勇 办公电话： 010-62515259  
手机： 13501130431  
地址： 北京市海淀区中关村大街 59 号中国人民大学信息学院  
邮编： 100872

Email: duyong@ruc.edu.cn

基金项目：国家自然科学基金青年科学基金项目 61502504。

作者简介：卢卫，男，副教授，研究方向为云计算与大数据管理、空间与文本数据库管理、索引技术，lu-wei@ruc.edu.cn。

作者简介：杜小勇，男，教授，研究方向为数据管理技术、语义网技术、智能信息检索技术，duyong@ruc.edu.cn。