**广东技术师范学院硕士学位论文**

**基于遗传算法的试卷生成系统研究与设计**

**Research and design of test paper generation system based on genetic algorithm**

**研究生姓名**  **赖复响**

**导师姓名职称**  彭 平 **教授**

**学科专业名称**  **系统理论**

**研究方向**  **计算机应用**

**培养单位**  **计算机科学学院**

**广东技术师范学院硕士学位论文**

**基于遗传算法的试卷生成系统研究与设计**

**Research and design of test paper generation system based on genetic algorithm**

**研究生姓名**  **赖复响**

**导师姓名职称**  彭 平 **教授**

**学科专业名称**  **系统理论**

**研究方向**  **计算机应用**

**培养单位**  **计算机科学学院**

**独创性声明及论文使用授权说明**

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 **广东技术师范学院** 或其他教育机构的学位或证书而使用过

的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

本人完全了解 广东技术师范学院有关保留、使用学位论文的规定，即:学校有权保留送交论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。同意 **广东技术师范学院** 将本人的学位论文提交清华大学中国学术

期刊（光盘版）电子杂志社全文出版和编入 CNKI《中国知识资源总库》，传播学位论文的全部或部分内容。

□公开 □保密（ 年 月） (保密的学位论文在解密 后应遵守此协议)

学位论文作者签名： 导师签名：

签 字 日 期 ： 签字日期：

摘 **要**

目前来说，考试依然是考察学生的知识水平和能力水平主要的测量方式，要做到科学地、客观地和公正地设计出一份好试题，就必须综合运用相关学科的知识，如：教育学、心理测量学、教育统计学以及计算机科学等。但是目前的现状是，题库系统无统一的标准，从而导致相关的试题资源达不到有效的共享，试卷生成的方法还不太完善，在教学过程中应有的作用得不到有效的发挥，这样看来，对现有的试题资源进行有效管理、研究出更令人满意的试卷生成系统具有极为重要的意义。

文章首先分析了现有组卷系统的发展状况，然后，根据中职学校的实际需要，对该系统的开发结构和方案进行了设计，建立了具有试题的管理、试卷生成、试卷的提交等主要功能的基于遗传算法的试卷生成系统。本文的第二部分详细的介绍了试卷生成的相关算法。基于试卷的质量需求，详细的分析了多个试卷生成的条件，进而利用基于遗传算法的智能搜索能力强和多目标参数优化的特点，研究试卷生成的方法。在对现有试卷生成模型的特点和不足分析的基础上，充分考虑满足试卷结构、试卷总分、试题曝光程度的要求，研究出了新的试卷生成模型，即将难易程度比例的分布以及覆盖的程度作为该系统最终的要求；然后将有条件约束问题转化为与之相对应的无约束条件问题，以此保持了该群体的多样性。本文的第四部分布置了系统的开发环境，并对其中的技术难点进行了分析，包括试题的导入、重复性检测以及系统的安全性；最终对该课题进行了总结，并展望分析了课题的后续工作。

通过运行该系统，对其通用性、安全性和可扩展性进行了验证，在较好的满足试卷生成要求的复杂情况下，实现了试题质量的控制、教学效率的提高、以及教考的分离。

**关键词：试卷生成； 遗传算法； 试卷生成模型； 罚函数**

**Abstract**

Examination for candidates' knowledge and ability level is the main measurement method, it needs to make comprehensive use of educational measurement, psychometric, educational statistics and computer science and other disciplines of knowledge, scientific, objective and fair design examination questions. The existing database management system are not unified standard format, test resource cannot be shared effectively, automatic test function is not perfect, can not effectively in the process of teaching play its due role, therefore the existing database resources for effective management, to study the better test method has important significance.

This paper firstly analyzes the development status of item bank system, and then based on the examination of the actual needs of the design of the system development and the plan structure, was constructed based on Web genetic algorithm was applied to test paper generation system, achieve test management, test generation, element management and examination paper. One of the fifth parts of the automatic generating test paper algorithm are introduced in detail. From the papers quality needs, in a plurality of test conditions on the basis of the analysis, use genetic algorithm to multi-objective parameter optimization and intelligent search capability and strong features, for method of test papers. Through the analysis of the existing test paper model characteristics and shortcomings, to meet the test structure, test scores, item exposure degree, designed the difficulty distribution ratio and coverage as the goal, time and test reliability as the restriction condition test model; and by adopting the method of penalty function will have the constrained problem into the unconstrained problem for infeasible solution, increase the penalty, so as to maintain the population diversity. The sixth part of the system development environment for the layout, and the key technology of the system difficult points: examination of the batch, test repeatability of detection and the security of the system were designed and realized; finally sums up the subject, and the subject of the follow-up analysis.

Through the operation of the system, the system has a certain universality, scalability and security, better able to meet the complex requirements, contribute to the realization of examination paper quality control, improve teaching efficiency, carrying out the separation of teaching and testing.

**Keywords: Automatic test; Test model; Genetic algorithm; Penalty function**

目 录

[摘要](#_Toc686313474) 2

[摘](#_Toc686313475)[要](#_Toc686313475) 2

**[Abstract](#_Toc686313476)** 3

[目 录](#_Toc686313477) 3

[第一章 绪论](#_Toc686313478) 4

**[1.1](#_Toc686313479)** [课题研究的背景](#_Toc686313479) 4

**[1.2](#_Toc686313480)** [背景与现状](#_Toc686313480) 5

**[1.3](#_Toc686313481)** [论文进行的工作和所要达到的目标](#_Toc686313481) 5

**[1.4](#_Toc686313482)** [论文整体结构和撰写大纲](#_Toc686313482) 9

[第二章 试卷生成算法研究](#_Toc686313483) 10

**[2.1](#_Toc686313484)** [组卷问题分析](#_Toc686313484) 10

**[2.1.1](#_Toc686313485)** [测量理论](#_Toc686313485) 10

**[2.1.2](#_Toc686313486)** [试题属性](#_Toc686313486) 10

**[2.2](#_Toc686313487)** [组卷算法的发展概述](#_Toc686313487) 12

**[2.2.1](#_Toc686313488)** [回溯试探法](#_Toc686313488) 12

**[2.2.2](#_Toc686313489)** [随机选取法](#_Toc686313489) 12

**[2.2.3](#_Toc686313490)** [基于专家系统组卷法](#_Toc686313490) 12

**[2.2.4](#_Toc686313491)** [遗传算法组卷法](#_Toc686313491) 13

**[2.3](#_Toc686313492)** [遗传算法概述](#_Toc686313492) 13

**[2.4](#_Toc686313493)** [基于遗传算法的组卷方法](#_Toc686313493) 13

[2.4.1 组卷模型](#_Toc686313494) 13

**[2.4.2](#_Toc686313495)** [约束条件的处理](#_Toc686313495) 15

[2.4.3 组织试卷过程中遗传算法的相关设计](#_Toc686313496) 15

**[2.4.4](#_Toc686313497)** [实验结果与分析](#_Toc686313497) 18

**[2.5](#_Toc686313498)** [本章小结](#_Toc686313498) 20

[第三章 系统的需求剖析和设计](#_Toc686313499) 20

**[3.1](#_Toc686313500)** [系统需求剖析](#_Toc686313500) 20

**[3.1.1](#_Toc686313501)** [业务流程](#_Toc686313501) 20

**[3.1.2](#_Toc686313502)** [系统对功能以及性能的需求](#_Toc686313502) 21

**[3.1.3](#_Toc686313503)** [系统角色分析](#_Toc686313503) 21

**[3.1.4](#_Toc686313504)** [系统工作流程](#_Toc686313504) 21

**[3.2](#_Toc686313505)** [系统的总体设计](#_Toc686313505) 22

**[3.2.1](#_Toc686313506)** [设计原则](#_Toc686313506) 22

**[3.2.2](#_Toc686313507)** [系统软件的体系结构](#_Toc686313507) 22

**[3.2.3](#_Toc686313508)** [网络拓扑结构](#_Toc686313508) 22

**[3.2.4](#_Toc686313509)** [功能模块设计](#_Toc686313509) 23

**[3.2.5](#_Toc686313510)** [数据库设计](#_Toc686313510) 23

[3.3 本章小结](#_Toc686313511) 23

[第四章 系统实现及难点解决方案](#_Toc686313512) 24

**[4.1](#_Toc686313513)** [开发环境](#_Toc686313513) 24

**[4.1.1](#_Toc686313514)****[NET](#_Toc686313514)**[开发工具—](#_Toc686313514)**[C#](#_Toc686313514)** 24

**[4.1.2](#_Toc686313515)****[ASP. NET](#_Toc686313515)**[简介](#_Toc686313515) 24

**[4.1.3](#_Toc686313516)****[ADO. NET](#_Toc686313516)**[简介](#_Toc686313516) 25

**[4.1.4](#_Toc686313517)****[SQL Server](#_Toc686313517)**[简介](#_Toc686313517) 25

**[4.2](#_Toc686313518)** [系统实现](#_Toc686313518) 26

**[4.2.1](#_Toc686313519)** [系统登录](#_Toc686313519) 26

**[4.2.2](#_Toc686313520)** [试题管理模块](#_Toc686313520) 26

**[4.2.3](#_Toc686313521)** [系统设置模块](#_Toc686313521) 26

**[4.2.4](#_Toc686313522)** [试卷管理模块](#_Toc686313522) 27

**[4.2.5](#_Toc686313523)** [系统的安全性实现](#_Toc686313523) 27

**[4.3](#_Toc686313524)** [本章小结](#_Toc686313524) 28

[第五章 总结与展望](#_Toc686313525) 28

[参考文献](#_Toc686313526) 28

[在学期间发表论文情况](#_Toc686313527) 31

# 第一章 绪论

## **1.1** 课题研究的背景

自二十世纪以来的教育发展中，其重大成就之一就是在相关教学领域中应用了计算机技术，并且该项技术的使用也标志着教育现代化的开端。其中计算机应用于现在教育领域的一个主要分支就是计算机辅助教育，随着其应用范围日益扩展，已经蔓延到家庭教育、学校教育、社会教育的各个方面，朝着信息化的方向快速发展一门新学科现代教育技术就逐渐的形成了。计算机辅助教育可分为2个方面：

（l）计算机辅助教学(CAI)：通过电子计算机为运算核心和辅助媒介，教师将一些学习信息交由电脑传递给学生，在此同时教授学生相关的专业知识技能。

（2）计算机管理教学(CMI)：在教学的监管和督导中引入电子计算机技术，发挥其技术优势对与教学相关的计划、安排和调度等诸多资源进行细致的管理。

学生在学习过程中发展是不均衡的，德育体美劳等各项指标往往需要通过考试来进行评测，通过考试得到的具体信息总结出相关的教学讯息，同时得出教学问题，进一步改进教学方式，提高教学质量。测试试题的制作水平很大程度上决定了考试活动是否能达到预期的目标。

对试题有如下诸方面的要求：

其一，可以如实测评出学生通过相关的教学对知识的掌握程度，能够全面的了解学生的具体的接受知识的能力。

其二，在测定出考生对相关知识的实际收获的程度同时，又要反映出考生的应用能力。

其三，可信度高、结果稳定、一致性高的试题。其四，考试题目能够尽量覆盖教学要求的范围。

其五，内容简单扼要、通俗易懂，语句通畅，叙述清晰明了，涵括的知识点全。

其六，测试结束后后续工作轻易，取分、排名等后续处理方便。

从中我们可以看出教学考试中对出题工作的要求是非常高的。反映出有关考生群体对知识的掌握程度的重要因素之一，就是要编制一套非常好的试题，这样用来分析考生的学习状况就变得可用，如成绩分布、最低成绩、最高成绩、平均成绩、试题检验后的难易程度、区分度等。而且也可以在一定程度上的反映教育工作者的教育情况。但是，从另一方面讲，如果不能为考生以及教育工作者的教

与学的成果作出正确的衡量的一份试卷，是没有达到最终要求的，这样就可能影响学生和从事教育方面人员的真实情况，甚至得到完全不准确的结果。现阶段，考察学习人员对知识的接受程度的手段比较匮乏，就目前来讲，依然是考试，但是制定一套试卷对于教师而言不是一件简单的事情。而且传统标准化试卷的命题，通常有命题的教学专家组成命题团队，以教学大纲中所提出的具体标准为准线，结合命题团队的丰富经验，确定试题所要考察的知识范围和目的，为了使得到的试卷符合要求甚至足够标准，还需要经过反复的修改和推敲完善。但是这种制定试卷的方法的工作量大，试题的难易很难掌握、题量大小、标准化程度等，每次考试出卷需重复劳动，且受出卷老师主观因素影响，对管理试题和试卷不够科学性、系统性。目前许多考试的执行过程存在着不科学，考试手段不先进，对评分标准、命题内容等的人为因素的较多，这些缺点在于：对于相关的试题无法做出正确的评价和更新；选题的范围比较狭窄；试题产生的好坏很大程度上取决于个人素质。

由于考试内容改革的要求，再加上科学技术，特别互联网技术的快速发展，考试模式等进行改革。计算机辅助教学(CAD)工程的重要组成部分是通过使用电脑来建立考试试题库，实现计算机自动选题组卷，这样，考试规范化、科学化、教考分离等就容易实现了。

所以，由计算机对试卷进行自动的生成，并且形成考试题库。试卷的管理要高效而快捷、自动化的前提就是要实现试卷的自动生成，具有非常大的意义。这样使得试卷的制作工作十分轻松愉快，在一定的程度上，也减轻了教师的劳动力和宝贵时间，便于教师有更多的精力和时间在教学和学生的管理上。教师使用计算机，通过自动生成程序制作试题，同时可以对相应的试卷进行评估，在减少教师劳动量的同时，也进一步增加整个过程的正确率。

## **1.2** 背景与现状

在1922-1928年间，普莱西就将教学与程序相结合，并对其进行了进一步的探索。这同时也导致了程序教学与教学机器的开始。1958年，世界上第一个计算机教学系统在美国IBM公司沃斯顿研究中心设计出，在平常的教育和生活领域，电脑逐步的被大范围的采用，伴随电脑的发展，CAI也得到快速发展。

在CAI的基本内容中，试卷的整理、分析等管理工作是其关键组成部分。试题题库的构建是一个繁琐复杂的过程，是在明确了题目的基本属性和组成结构后，以相应的数学模型作为核心构建起来的，其中试题由优秀教师编写。在国家相关项目的大力支持和帮助下，越来越多的试题提举相继建成，例如高等教育基础科学系列题库，国家医学水平考试题库等。因为试题库涉及的各个方面都很精

密复杂，各科教师亲自负责与试题库相关的管理、维修，保护、更新、数据统计等工作，因此，基本上都是在比较封闭的状态下运行的，对于较小的科研单位，也许就没有使用的机会，因此，提供得不到应有的使用，也发挥不了较好的作用。

普及和迅猛发展的计算机网络，友好界面、网络共享，为了能够有效的解决系统在运行的过程中出现的错误。现阶段，爱单击题库系统开发的基础上，开始了以遗传算法的试卷生成为基础的新系统。这种题库系统的建立在是Web站点上运行，用户通过浏览器进行访问。并且具有一定能够的新功能：

#### （1）集中管理、共享使用

试题在经过专人确认以后，在第试题进行删减、添加、修改等。对于比较分散的题库，主要是通过设立一个Web站点，以此对整个数据库中的试题进行操作。

#### （2) 集中存储测试数据

网络的优越性，使及时存储数据变得十分容易，利用统计与分析、数据挖掘，以完成对试题参数的错误值进行校正，从而保证试题的高质量。

#### （3) 开放使用、开放建设

网络的优越性，决定了系统的开放性，无论是搜索、抽样等操作，经过简单使用浏览器登陆网络，就可以实现，用户制定的试题，经过提交，试题经过权威专家的检测，符合条件，方可纳入试题库中。通过网络形式，制定的试题，能够有效的节约成本，缩短了时间。

#### （4）方便实用易懂

当客户独自利用网络，由系统的界面引导就能完成所有的操作，不用安装、配置和管理，进而为使用题库的用户提供的方便。

目前，应用于中小学校的大部分试题库与试卷系统，多数是利用的数据库是

Paradox、Access及开发工具是VF、VB、C++等程序语言开发的适合单机的系统。如今电子计算机与网络处在高速发展的时期，单一的试题题库已经不能满足人们的需要，对于题库的管理也不断的加大，同时试题库的安全问题也越来越受到重视，而且试题确定后交由审阅机关审查的过程还是完全依靠人工，还未实现网上命题和自动提交的OA模式。

用于中小学，这些系统可能是合适的，但在中职学校这些系统也往往不能满足了，因为中职学校里面专业和科目都比较多且常变化的特点。

## **1.3** 论文进行的工作和所要达到的目标

由论文之前的发展现状和课题背景可以总结得出，本论文的目的是：利用互联网和计算机，对试题库和试卷系统形成自动生成模式，以方便，出卷人，同时

在抽取考试题目的时候要依据着大纲的具体安排有序合理的展开，保证产生的试卷能与考生的学习现状相符，能够充分的反应学生的真实学习情况和教师的教学能力，在通过网络环境进行提交，试卷的审核人进行在线审查，审查通过后形成正式的试卷，归纳到试题库中。

本课题有以下几个研究内容：

（1）试卷生成系统的数学建模1．题型难度系数数学模型

一份科学准确的试卷应该是：拥有难以适中的题型，实现对学生考试的目的，例如：对考试优秀率的语气，反应出一个班级学生的分布情况。一般的题库中，关于试题的难易程度，都有明确的分类。传统的命题模式，要把握题型的难度程度，是很难控制的，因此，在不同的时间对于同一个类型的试题的难易程度的判断也存在差异，所以这在一定的程度上也对试卷的公平性和客观性有一定的影响。因此可以对试题的难易程度建立具体的数据模型，为了保证试题的准确性，其模型为：

设Q参加考试的全体学生，L为针对某一道题，回答正确的学生数量，N为该题的难度系数，令N= (Q-L) /Q，其中，当L=Q时，取N=1 （1.1）

从模型可以看出，L增大，难度系数就变小，L变小，难度系数就增大。按通熟的说法就是，起容易被人做对的题目，就说明这个题目越容易，反之，就是相反。

例如，给定题库的难度指标是5级，分别为5，4，3，2，1，则把模型作取整处理，即是：N= INT((Q-L) /Q)×5)，其中，当L=Q时，取N=1。这里的INT是表示取整函数，详细情况见表1。

表1

| 考生号 | 试题 | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | | 2 | 3 | 4 | 5 |
| 1 | √ | √ | √ | √ | × | × |
| 2 | √ | √ | √ | × | × | × |
| 3 | √ | × | √ | × | √ | × |
| 4 | √ | √ | × | √ | × | × |
| 5 | √ | √ | × | × | × | × |
| 难度等级 | 1 | | 2 | 3 | 4 | 5 |

从表1可知，简单地去理解难易系数就是，以5个学生为例子，难度指标为 5

级，那么，某道题的难度等级就是这5个学生中做错的人数。这样子，在后面的各自模型中，理解起来就容易多了。该模型的优点就是，反映客观事实的同时，

算法本身也很简单明了。

2．试题在章节中的分布模型

确定各章节在试卷中的分布比例的方法是这样的，依据课程内容和考试大纲，在试卷的生成过程中，采用人工输入的形式，把各科中各章节的所占比例录入题库.说明如下，第j部分章节在试卷中所占的百分比用来Mj表示，另外，用

F、b、Nj分别表示试卷的总分、分值、试题数，那么，Nj= INT((MjF) /b) （1.2）

举例，第5章在试题中占20%，总分为120，b= 3，就是说每题2分，那么

N5= 8，也就是说第5章要出8题，该数学模型转换成算法也不难。

3．各级难度题型分布模型

被测试对象水平的反映其中之一就是试题的难度。期望值直接影响了难度，当期望值增高，就说明了试卷中的题平均难度变低，这样子，用用户指定的期望值求得各级难度的试题所占的比例。

被测试对象达到一个平均水平，学生所做的每一道题的结果相对于其他的题目而言，都是相互独立的，每道题的只有正确和错误两种结果，结果具有较强的随意性。所以，随机答题时间和离散型随机的变量的二次函数B(m, r)一般会认为是相符合的。

Pm(j) =Cjmrjqm-j，其中 j= 0,1,2,…，m，m∈N+；r> 0，q> 0, r+q= 1 (1.3)

T=mr是二项分布均值，在以上模型中，其中一个相对普通的被测试的对象与之相对应的相同难度的概率，用r表示。难度层次由j表示。而Pm（j）表示的是在被测试的对象中，对于j级别的试题做正确的概率。整篇试卷的平均难度系数值用T表示。在固定了m, r的二项分布中，P {x=j}会随着j值的增加，而递增，其递增的规律是，先递增至最大值，然后在进行单调递减。

4．试卷的生成

设定计算机考试，考试的范围囊括了五个章节，在试题库中，只有选择题和填空题二种类型，难度的系数从1到5分为了不同的级别，在试题库中的结构为：

题号 内容 答案 难度等级 章节号 正确与否第2章试卷的输入，按照如下的条件进行：

①填空题40%，选择题60%，分值均为1分，满分100分。

②比例分析：第1章30%，第2章20%，第3章15%，第4章20%，第 5

章15%，如表2所示。

表2

| 题型\章节 | 1 | 2 | 3 | 4 | 5 |
| --- | --- | --- | --- | --- | --- |
| 填空 | 12 | 8 | 6 | 8 | 6 |
| 选择 | 18 | 12 | 9 | 12 | 9 |

③对于试卷的期望值是2，试卷的难度系数也是2。根据以上的数学模型

（1.2），可以得到上述表2，然后在通过数学模式（1.3），这里，T=2.2，m=6，计算得出r= 0.56. 代入(3)式数学模型，得出P(0) = 0.007; P(1) = 0.062; P(2) = 0.179; P(3) = 0.310; P(4) = 0.287; P(5) = 0.129; P(6) = 0.030.

在实际的操作过程中，可以把比率相对较小的两头忽略掉或者加到P（1）和

P（5）中，这样可以得到表3。

表 3 不同章节、不同难易程度和题型

| 难度等级 | | 1 | 2 | 3 | 4 | 5 |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 填空 | 0.6 | 0.8 | 0.56 | 0.4 | 0.4 |
| 选择 | 0.8 | 2.1 | 0.8 | 0.6 | 0.6 |
| 2 | 填空 | 1.4 | 2.1 | 1.4 | 1 | 1 |
| 选择 | 2 | 3.2 | 2 | 1.6 | 1.6 |
| 3 | 填空 | 2.4 | 3.6 | 2.4 | 1.8 | 1.8 |
| 选择 | 3.6 | 5.4 | 3.6 | 2.7 | 2.7 |
| 4 | 填空 | 2.2 | 3.4 | 2.2 | 1.7 | 1.7 |
| 选择 | 3.3 | 5 | 3.3 | 2.5 | 2.5 |
| 5 | 填空 | 1.3 | 2 | 1.3 | 1 | 1 |
| 选择 | 2 | 3 | 2 | 1.5 | -1.5 |

对表3加以处理，就能据其生成一份满足整体要求的试卷。

(2)分析与设计试卷自动生成系统的算法：其中包括了：试卷的特征，响应时间与效率，算法的设计，试卷的生成算法。智能技术能在辅助试卷生成系统中的运用中体现。

(3)总体设计系统是在需求分析基础上进行的：系统总体设计、系统的模块划分、系统的软硬件体系结构设计都是依据系统的需求分析进行的。

（4）试卷自动生成系统的标准：合理分布题型难易，预期的考试目标实现了，这些都是一份科学的试卷应该达到的。

## **1.4** 论文整体结构和撰写大纲

第一章在篇首的第一部分—绪论中。主要对所探究课题的相关现状和背景进行了全面的介绍，本课题主要的研究目的和研究内容也在这章进行了分析。

第二章研究当前自动生成的试卷算法。目前常用的算法，对其存在的算法的优缺点进行分析，最后运用遗传算法对试卷进行生成。在此过程还全面细致的介绍了算法产生的所有环节，最后由具体的实验进行验证，明确了所采用的算法

的实用性和准确性。

第三章从整体的角度对该系统设计和具体分析进行了介绍。依照不同学校每个办公室和人员对该系统的不同需求，对于试卷生成系统的相应指标进行调整，确定其应用范围和需求范围；根据系统的需求，对系统软硬件进行相应的安排和设计，工作流程图和系统功能模块图也在这章里给出，并对系统中所使用的几个重要的数据库表结构进行了说明。

第四章系统实现及难点解决方案。对该系统的研究和构建过程的背景进行了描述，还简介了支撑系统实现的关键技术，在此过程中，还将系统研发中遇到的问题和关键点进行了着重的分析和描述。

第五章在此章中对以前的工作进行了总结并对未来工作进行了展望。第一总结了自己读研究生以来所做的工作，第二下一步工作给出了方向和指引是根据系统的后继需要提出来的，同时对未来工作进行了相关展望。

# 第二章 试卷生成算法研究

## **2.1** 组卷问题分析

从本质方面来说，为了既能满足教学要求又能满足用户要求，利用电脑按照确定的系统算法从已有的试题题库中选取最优试题自动生成所需的考试试卷。

在保证试题库试题质量不变的前提下，生成算法的设计决定了试卷生成系统的试卷生成工作的效率与质量。不同条件下的许多系统随着国内外的许多学校机构、科研单位对试卷生成系统的深入研究应运而生。但这些系统都存在使用繁琐，实用性不高等的问题和不足，例如，让用户详细地提出每道题的题型、难度、章节分布等要求，然后在试题库中利用随机函数进行选题。对现有题库资源进行有效管理、研究更好的试卷生成方法在当今现有无统一标准题库管理系统下，现阶段还没有办法实现优秀试题数据的完全共享和试卷的自动组合与优化，在相应的教育培养过程中充分利用其有限条件使其更加实用，这就显得颇为重要了。试卷生成的过程包括：对考试的了解，根据所学的知识，对试卷进行研究和整合，对试卷的相关要求进行确定，进而对考试的模式做好设计，最后在按照设计好的模式，更好的处理试卷中试题的组织安排任务。

### **2.1.1** 测量理论

一般来说，我们将测量理论分为三个主要部分，分别是概括化理论、经典测量理论和项目反应理论三种理论。以下分别对三种理论一一分析介绍。

#### （1）概括化理论（Generalizablity Theory，简称GT)

作为现代测量理论之一的概括化理论，是理论中最重要的一部分内容。通常情况下，我们又将其叫做概括力理论。在Stuffebeam和Cronbach. L，D. L.为弥补传统的真分数理论及信度理论之不足的情况下，概括化理论才应运而生的。经典信度理论的缺点是：不能处理影响信度各方面中的某一方或同时处理所有各方面。因为对于误差来源的成分，它不能加区分，只能分析两个测验形式之间得分的一致性程度。概括化理论的产生为行为测量提供了可靠性的证据。

为常模参照决策以及准则参照决策解决了，在测量多个误差来源的维度方面，经典测量理论是概括化理论延伸，提出了最大化信度（概括化）系数和最小化误差的测量最优化方案。因此它为检查行为测量提供了灵活的、可行的理论框架。它突破了作为单一含混的简单统计量的束缚，采用方差分析方法，明确分解测量误差。

#### （2）经典测量理论(Classical Test Theory，简称CTT)

自1920年开始，经典测量理论开始产生，但直到1960年该理论才得到基本完善。真分数理论是传统的经典测量理论别名，在测量的过程中，没有出现任何误差的情况下所得出的数值就是真值。

只有满足以下的三个条件才能实现经典测量的理论：

l）某一个学生的真实知识水平为X，和存在的误差分数x，这两者之间是线性的关系。

2）这名学生，在经过大量的测验进行误差的统计。

3）在测量的数据中，任何两个误差都是彼此独立的。

依此上假设，得出结论：误差分数和真分数的总和就是观察分数。也就是多次测量以后的平均值，是最接近真实分数的。经典测理论的应用主要表现在，测试通过率的难度，区分度是用总分和题目分数来代表的。

长期以来，经典式理论都作为建立题库的基本理论。对于相关题目公式的分析，都是根据该理论进行建立的。例如：估计分数真值的标准误差，表示区分度的题目，题目数与信度关系式，难度的p值，以及测试相关系数等等

#### （3）项目反应理论(Item Response Theory，简称IRT)

在二十世纪初的现代测量技术的基础上发展起来项目反应理论的，我们又将其叫做潜在特质理论，是在二十世纪五十年代真正发展，在二十世纪七十年代得到了突破性的发展。潜在特质是在观察分析测验反应基础上提出的统计构想，这种能力一般是潜在的，并且对这种潜力的计算是采用测验总分作为标准的。项目反应理论中，和潜在特质有关系的是被测验的项目的成绩。项目反应理论的建立是和测量量表的分数是否统一的标志，该项目参数的主要特点在于恒久性。我们可以通过一些项目分析的具体内容特点，包括鉴别度和难度系数，以便为项目的编制等操作起到重要的指导作用。

双参数模式、三参数模式和单参数模式三个相对复杂的模式组成了项目反应理论。在考试的过程中，学生可以根据这些参与和考试回答问题的实际情况进行估算。局部独立性、完成测试时间、能力的单维性等是项目反映理论的集中假设，项目理论反应是和能力测试有关的理论。

### **2.1.2** 试题属性

试卷组成的过程是一个具有复杂性、多目标性和难度性的特点。需要对多个试卷的分析，然后组成一个试卷。对于试卷的区分度、难度等以及考试的章节等都能够在很大程度上决定试卷水平的高低。此外，我们对试题的各种变量在进行取值时要确保数据的合适和恰当，因为它对试卷的质量产生了重要的影响。所以，合理有效的理论基础，对试题的指标和评价标准的确定有决定性的作用。

此系统是在经典测量理论的基础上确立起来的，所以，我们通过该理论内容所制定的各种用于题目分析的关系式，都会对试题的属性产生影响，而具体的试题属性主要包括以下几方面：

试题的知识点：这是由试题所对应的知识点的编号决定的。试题的标号：是指在试题库中试题唯一的符号。

试题章节：这是由于教学内容的章节号码进行确定的。试题的分数(FS):试题的分值。

试题的认知层次(RZ):我们将试题和教育目标分类学相联系。一般情况下，试题主要分为不同的把试题分为认知阶段，分别是知道、了解、运用等。而这也是和大纲的层次分析相一致的。

试题的难度(ND):是有组成试卷的各个题目的难度系数所决定的。

试题的答题时间(DTSJ):在做题的过程中，所消耗的时间，单位是分钟。在试卷中，针对一个试题，如果绝大部分的学生都作对了，那么说明该题的

难度系数较大，反之，该题的难度系数较小。

#### （1）选择题的难系数

学生在做题的过程中，由于备选答案的数量脚下，在选择题中，概率表现的越明显，这样反而不能更好的反应出题目的真是难度，所以一般情况下，为了平衡机遇现象，而采用吉尔福特公式

*CP**KP*1

*K*1

(2.1)

在上述公式中，对于没有校正的机遇测试难度系数的表示为P。每一道题目中能够用来选择答案的数量用K进行表示，而CP就反应的是校正以后猜测机遇的测题难度系数。

#### （2）客观题难度系数

*ND*  *R*

*n* (2.2)

以上公式中，针对一道题，答对的人数用R进行表示；参加考试的总人数由n表示，而该题目的难度系数则由ND表示。

*ND* *NDH* *NDL*

2

*ND*