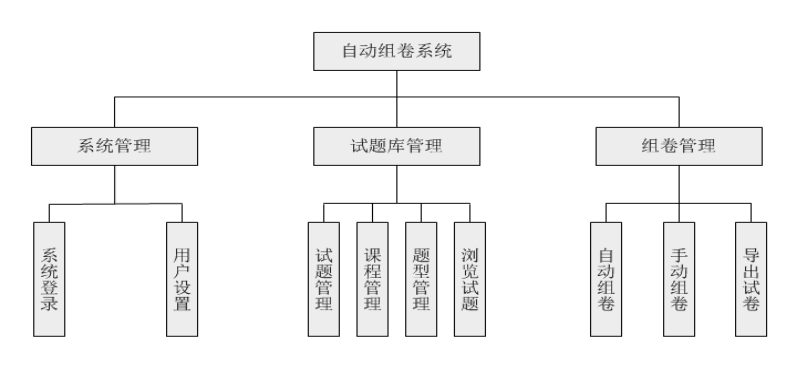
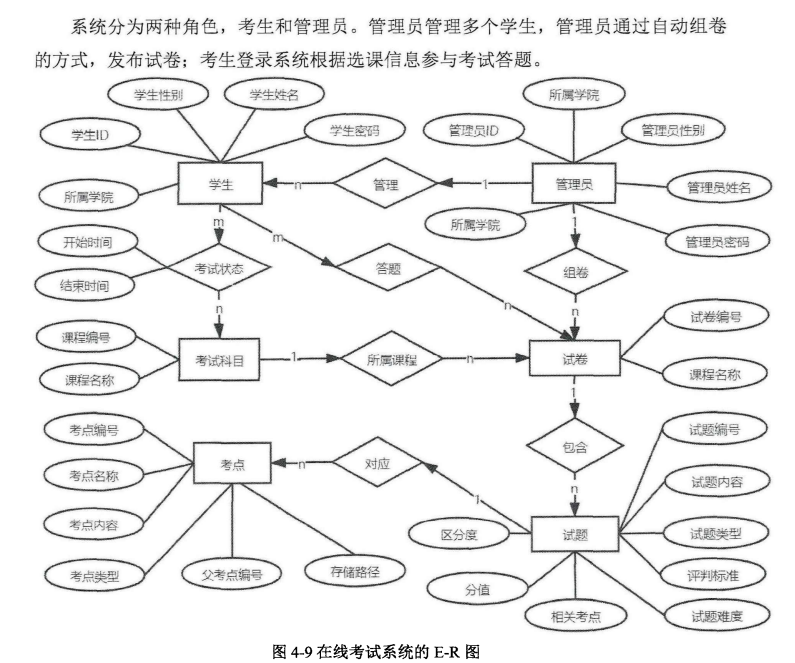
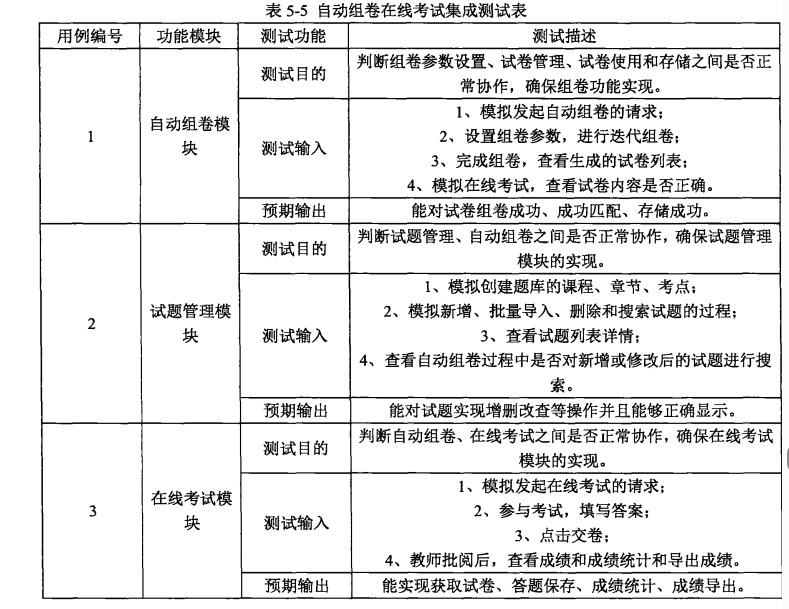
# 自动组卷







自动组卷系统需要解决的是一个满足多目标约束条件的复杂问题。其需要满

足的约束条件大体上可以有：

1) 试卷的质量约束条件，主要有试卷的难度、覆盖度、相似度、区分度等。

2) 试卷的格式布局约束，主要有试卷的分值、题型、总体题数以及各个题

型中的题数。

3) 其他约束条件，如某一章节最高分值约束，同一知识点考核次数限制等。

试卷中的约束条件相互之间并非毫无关系的，而是在试卷中相互影响、相互

联系的。

试卷质量：

是由 试题科学性 、 知 识 点 的 随机性和合理 性 、 考点 难度 的覆盖率和 自 然性决定 。 不 同 老师 出 的 试卷 ， 可能带有主观性 ， 同 时 侧 重 点 的 不 同 与 老师观念 的 不 同 ， 也会导致试卷 的 主观性太强 ， 难 以 反 映考生 的 真实 水平 ， 降低 了 考试 的科学性。

在 一 般 的 组卷流程 中 ， 考卷涵盖 的 题型 、 各类题型 的 数 目 均 未 形成偏差 ， 各个章节 、 题 型 的 难度系 数这 ２ 个制 约 要求也许存在偏差 。 偏差 的 高低通 过客户 的预期值与 各类制 约 条件 的 关键性来确 定。

参考方向：

1. 提高试卷公平性、科学性，降低学生背题可能性；降低作弊可能性；试题筛选成功度；
2. 卷在难度、章节知识点数和独一性等指标上能够达到要求。
3. 试卷的试题内容上降低出题老师的主观因素影响。
4. 避免所出试题重复出现
5. 由于传统的遗传算法存在进化后期效率变 低、进化方向可能出现逆向、容易陷入局部最优 解等问题
6. 在线考试系 统 的 核 心是 自 动组卷算法 ， 但传统 的 自 动组卷算法存在 效率低 、 随机 性高 、 试卷难 易 不 易 控制 等 问 题
7. 在线考试系 统 的 核 心 是 自 动 组卷算法 ， 使用 算法 的 好坏决 定 了 生成试卷 的 质 量 ， 也 决 定 了 是否 能真正反映学生 的 学 习 水平 。
8. 目 标试卷试题重 复 、 知识涵盖不全和 试卷不符等 问 题
9. 为 了 克服 Ｐｃ 和过大造成个体之 间 的 适应度波动较大 ， 或者 是 Ｐｃ 和 偏 小致使个体 的 平均适应度相对稳定 ， 也许形成 “ 早熟 ” 的情况 ， 以 及保护优质 考卷 的 个体 ， 部分文献指 出 鉴于 交叉几率 Ｐｃ 和变异 几率的 适应性对策
10. 因 为种群 之间 存在差别 ， 客户 的 需 求也 同 样存在差别 ， 在遗传 算法 内 无法界定种 群演化至第几代便可 以 获得最佳的 考卷
11. 遗传算法 中 的 适应度 函 数是 决 定 组卷性 能好坏 的 关键 ， 组卷过程 中 的 适应度 函 数 值越大越可 以 指 导 问 题搜索 的 进化方 向 。 适应 函 数有两个 基本要求 ： 变量输入的 非负 性 以 及 目 标 函 数对试题搜索起正反馈 。
12. 常 用 的 选择算子包括轮盘赌选择和 锦标赛选择 等方 式 ， 其 中 轮盘赌选 择会保 留 部分劣 质 个体 以 保证群体 的 多 样性和避免 陷 入局 部最优 ， 但也会 陷 入无法 收 敛的 问 题 。 锦标赛选择则 只 保 留 最优个体 ， 无法保证群体 的 多 样性和 全 局 最优解 。
13. 在遗传算法每一 轮 的 迭代 以 后 ， 一 定概率 下 会将种群 中 最优个体淘汰 。
14. （ １ ） 可 以 对系 统 的 题库管理进行完善 ， 不 断对题库进行更新 ， 增加新 的题型 ， 这 样有利 于更加 全面的进行考查 ， 让考生能够学 习 到更 多 的 知 识 点 。 （ ２ ） 随着广大研 究人员 对于组卷算法 的研 究 ， 在线考试系 统 的 组卷策略与 组卷速 度还可 以 进 一 步 的 优化 ， 以 满足更 多 用 户 的 组卷 需 求 。 此外 ， 我们可 以 优化在 线考试 系 统 的 简 答题评 阅 方式 ， 通过结 合 自 然语言 处理相 关 的 技术 ， 实现主观题 的 自 动评 阅 功 能 ， 达到 完全 自 动化 阅 卷 的 效果 。 （ ３ ） 可 以 引 入人工智 能 以 及机器学 习 的 相 关算法完善在线考试系 统 的 其他功 能 。 例 如 自 动 归 纳 易 错 点 、 错因 分析 以 及根据错题本统计考点 的 掌握情况 。
15. 传统的自动组卷算法存 在效率低、随机性高、试卷难易不易控制等问题
16. 以生物进化为理论基础的遗传算法，具有很好的全局搜索能力，能够快速的 从全体解中搜索出符合全部约束条件的最优解。在搜索的过程中可以对多个个体 同时进行比较，存有潜在并行性。利用此特点能够进行相应的分布式计算，大大 加快了求解速度。但是遗传算法也具有很明显的不足之处，当问题计算量大时， 如果单纯的只使用遗传算法，则问题在局部搜索能力上效果较差，耗时较多以及 基于遗传算法的自动组卷系统优化设计及应用 20 后期搜索的效率低。在实际应用中表现为，算法的早熟收敛，以及问题的搜索容 易收敛到局部最优解。这也是遗传算法应用于本课题重点需要改进的地方。

## 自动组卷概念：

### 总述：

组卷 从时 间 的 维度可 以 划 为三个不 同 阶段 ［ ７ ］ 。 其 中 ， 第 一 阶 段是依据老 师 的 日 常教学 以 及考试大纲 ， 按 照 经验手动 出 卷 ， 虽然耗 时长但试题质 量高 ； 第 二阶段是将试题存入数据 库 中 ， 通过手动勾 选试题进行组卷 ； 第 三阶段是智 自 动 组卷 ， 出 卷系 统会按 照 老 师设置 的 试卷属 性参数 自 动 搜索 题库 ， 并 根据 组卷算法进行 自 适应 的 调 整 ， 最终完成组卷 。

自 动 组卷系 统 的 目 标是寻找到 多 约 束 条件 的 多 目 标和 全局 最优 的 解决方案 ， 根据输入参数 的 不 同 ， 会生成在一 定 范 围 内 随机 的 解 。 最终组卷 的 效率与 质 量的 核 心在 于找到 全局 寻优和 收敛速度 的 统 一 。

自动组卷指根据限定的试题数目、题型、分数、知识点 范围、考点分布等因素，自适应地搜索题库，根据每个限定 因素的特征参数的权重匹配出最优试卷，主要用于解决多重 约束的组合优化问题 [2-3]。

### 参数：

题目知识结构、数据延展进行全方位跟踪管理。

题目知识点、难度、教学要求、分数四项指标；

可以对题库和试卷中的题型进行自定义；

考试性质(考试名称)，试题类型，课程章节，试题数，试题难度等参数。

自动组卷模块根据这些输入的参数由系统自动搜索试题库中与特征参数相匹配的试题，从而抽取最优的试题组合。

### 应用：

有一些软件和网站提供这种功能：考试云和组卷网。

随机组卷像考试云，任意抽题规则的随机策略，能智能根据试题的难度系数、试题类型、试题知识分类进行随机组卷抽题。

具有针对性的随机抽题算法是在线考试系统在人工智能 (AI)领域在不断完善发展一项重要任务。今后将会有越来越多令人想不到的智能算法问世。

## 发展历程：

自动组卷系统的主要工作是设计并实现一套能够组织并建立试题库，并对试题库进行管理的系统。然后根据出题要求，利用试题库中的试题，进行自动组卷，生成符合要求的试卷

一 些传统 的 组卷算法 （ 例 如 随机 抽 取算法 、 粒子群算法 、 遗传算法 、 回 溯法等 ） 都是最早起源于美 国 ， 因 为互联 网 的 发展传入 国 内 ， 而得到 了 广泛应用 ［ ８ ］ 。

赵永平等针对传统 随机抽 取 算法进行 了 改进 ， 扩大知 识 点 覆盖范 围 ［ １ １ ］ 。 马 跃亮等对传统粒子群算法进行 了 改进 ， 避免 了 组卷收敛速度变慢 ， 容 易 陷 入局 部最优等 问 题［ １ ２］ 。 董敏等通过 自 适应遗传算法 的 设计 ， 利用 分段编码 改进 了 传统编码方式 ， 降低 了 算法执行过程 中 的 内 存消 耗 ， 提 高 了 速度［ １ ３ ］ 。

Ｇｏｏ ｇｌ ｅ 于 ２ ０ １ ２ 年 ５ 月 正式提 出 了 知 识 图 谱的概念 ， 并且将知 识 图 谱定义为一 种表示客观世界 的 实体 、 概念及其关系 的 语义 网 络［ １ ４ ］ ， 很好的缓解 了 大规模数据组织 处理 的 难题 。

## 技术优劣：

### 优点

提高公平性：自动组卷可以根据给定的约束条件搜索最优解，减少人为因素造成的失误。

例如，在线考试系统可以实现海量试题批量导入，无限层级试题分类，随机组卷模式，以及可以在手机、微信及钉钉上进行考试等2。

智能组卷系统的试卷组织模式非常灵活，有手动选题组卷、随机抽题组卷、固定抽题组卷和综合组卷，以及可以批量导入试卷的多种试题导入方式1。

智能组卷系统将对考试数据进行多方位分析，准确统计参考率、通过率、最高分、最低分、平均分、排名等数据，并生成多维考试报告1。

总之，自动组卷技术具有许多优势，包括提高组卷质量、支持高并发能力、灵活的试卷组织模式、简化考试流程和自动评分等。这些优势使得自动组卷技术成为一种非常有用且实用的技术。

### 缺点

组卷成功率低：对于整个组卷过程来说，组卷成功率低，即使组卷成功，花费时间也令人难以忍受3。

程序结构简单：该方法结构简单，对于单道题的抽取运行速度较快3。

重复率：同类试题不同阶段出可能存在一定重复试题的出现。

### 并不能完全取代教师出题。

教师在出题时不仅需要考虑试题的难度和知识点分布，还需要根据学生的实际情况和教学目标来设计试题。

自动组卷技术虽然可以根据给定的约束条件搜索最优解，但它无法完全理解教学目标和学生的实际情况。

因此，教师在使用自动组卷技术时仍然需要对试卷进行审核和修改，以确保试卷符合教学目标和学生的实际情况。

### 无纸化考试系统：

组卷到成绩分析系统自动完成

————————————————————

五大块：题库管理模块—自动组卷模块—客户端模块—自动阅卷模块—成绩评估与试卷分析模块

题库管理子系统:

试题库指由许多适用于不同目的、技能和任务需要且具有必要参数的大量优质题目组成的集合。题库管理模块能通过界面录入包含图形、公式和文字的各类型各课程试题，生成不同科目的试题库,可以方便地完成对试题的保存、扩充、修改和查询，从而有效的管理试题。

自动组卷子系统:

根据考试要求由组卷系统自动从试题库选出考试题目并形成符合各个老师所要求的试卷称为自动组卷。

成绩评估和试卷分析子系统:

此模块能够方便、快速、准确地对每个考生的成绩及排名进行查询并对考试结果进行统计和评价，包括成绩的分段统计、成绩分布曲线的绘制、平均分数、最低分、最高分、及格率的统计等并生成规定格式的考试分析文档。

关联关键词：智能组卷，随机组卷算法，回溯组卷算法，遗传组卷算法 无纸化考试 试题库

无纸化考试形式：

1.单机模式，即每个机子装一套系统及考题，考完收集成绩。

2.C/S 模式，即在服务器上装题库，在每台机子上安装客户端程序，登录，抽题，考试，传回服务器成绩

3.B/S模式，即整个考试系统全装在服务器上，考试端只需打开浏览器(Browse)即可，输入服务器 URL 即可调出页面登录、抽题、考试、评分、返回成绩等。

主要的自动组卷算法有随机抽取算法、误差补偿算法、 回溯试探法和遗传算法等 [1]

自动组卷算法的好坏直接影响到试卷的质量，如何从试题库中选出试题组成符合用户要求的试卷，并使组卷具有较高的效率和成功率是当前研究的重点。

现存的组卷算法主要可分为 3 类:

1. 随机选取法:

根据状态空间的控制指标，由计算机依据随机函数算法随机的抽取一道试题放入试题库，此过程不断重复，直到组卷完毕,或已无法从题库中抽取满足控制指标的试题为止。该方法结构简单，对于单道题的抽取运行速度较快，但是对于整个组卷过程来说花费时间过长。

随机抽 取算法优点 由 于 结构 简单 、 抽取题 目 快 、 易 于 实现而被广泛应用 ， 但也 存在 目 标试卷 随机性高 、精度低 、 重复率高 、 不可回 溯等 问 题 ， 因 此无法保证 目 标试卷 的质 量

猜想：依靠题库分类的程度深。

1. 回溯试探法：

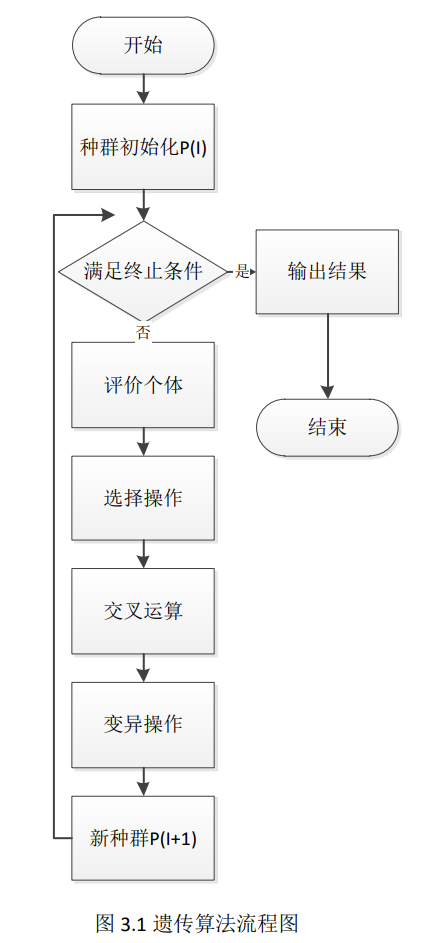
回溯试探法是一种选优搜索法，也称为试探法。它按照选优条件向前搜索，以达到目标。当探索到某一步时，如果发现原先选择并不优或达不到目标，就退回到上一步，重新选择。这种走不通就退回再走的技术被称为回溯法，而满足回溯条件的某个状态的点称为“回溯点”1。这种算法对内存的占用量很大，程序结构相对比较复杂，而且选取试题缺乏随机性，组卷时间长，后两点是用户无法接受的，因此它不是一种很好的用来自动组卷的算法

这种有条件的深度优先算法对于状态类型和出题量都较少的题库系统而言，组卷成功率较好。但是对于程序结构相对比较复杂，该算法表现就比较差了2。

总之，回溯试探法的优势在于它能够通过剪枝函数避免无效搜索，提高搜索效率，并且能够在搜索过程中动态调整搜索策略，以达到最优解。它是一种灵活、高效、智能的搜索方法。在广度或深度遍历试题时 ， 如果组卷过程 中 误差超 出 范 围 、 陷 入死锁 ， 就会返 回 上一 次 的 状态进行 回 溯 ， 直到 组卷成功 或 回 到初始状态 。 因 为采用 记忆化 的方式进行存储 状态 ， 相 当 于空 间 换时 间 的 思想 ， 实 际组卷 中 也发现 内 存消耗大 、 缺 乏 随机性 的 问 题。

猜想：轻量级的便于处理一些。会按照之前的状态反思现在的状态。

1. 遗传算法:



遗传算法从提出到发展至今已经非常成熟，关于该算法的应用领域非常多。 例如可作为机器学习算法；也可与神经网络、模糊理论等智能计算方法相结合； 除此之外经常与进化计算理论相结合。虽然算法本身已经成熟，但在遗传算法的 应用以及延伸研究等方向上依旧非常热门和活跃[21]。

此算法的实质就是一种把自然界有机体的优胜劣汰的自然选择、适者生存的进化机制与同一群体中个体与个体间的随机信息交换机制相结合的搜索算法。

首先将所要求解的问题表示成二进制编码，然后根据环境进行基本的操作:selection，crossover，mutation······这样进行不断的所谓“生存选择”，最后收敛到一个最适应环境条件的个体上，得到问题的最优解。

在遗传算法中，适应度函数主要对个体的优劣进行评估，对染色 体上优劣基因的筛选起重要作用。适应度函数的信息可以指导问题的搜索方向， 其直接影响算法的运行效率和结果的准确性。所以适应度函数对该算法来说至关 重要。

选择算子的常用方法有“轮盘赌 策略”( Roulette Wheel Selection )，它是遗传算法一直以来最为常用的选择算法。 其通常表现为一种“比例选择”，具体操作中为个体被选中的概率和它的适应度 值成正比。轮盘法的优点是简单易懂，计算量小，适用于大规模的优化问题。在遗传算法中，轮盘赌选择法可以使适应度值大的个体有更大概率被选择作为父代个体来生成新个体，从而提高算法的效率。

轮盘法的缺点是当适应度值相差不大时，个体的选择往往不是依据个体的选择概率，而是根据“累积概率”来进行选择，这可能会导致某些个体明明选中概率很小但却因为它位置靠后而导致其累积概率很大而被选中的情况发生。此外，该方法是属于有放回的选择，不适合不放回的情况；

（ ２ ） 误差补偿算法 误差补偿算法［２ １ ］ 也是组卷系 统常用 的 一 种方法 ， 主要是针对约束失效情况 的优化 。 组卷过程 中 的 约 束条件 的 理想化 ， 会导致约束失 效 的 情况 ［ ２２ ］ ， 通过对部分次要约 束条 件进行放缩 ， 可 以 避免 了 组卷过程 中 因 为约 束条件与题库容量造成 的 组卷失 效 ， 避免 系 统进行死循环 的 状态 ， 进而提高组卷 的 成功 率 ， 但也 同 时存在组成试卷精度 误差 的 问 题 。

为了解决轮盘法的缺点，可以采用其他的遗传算法选择方法，例如锦标赛选择法、随机选择法等。这些方法可以在适应度值相差不大时，更好地保证个体的选择概率与其适应度值成比例。

锦标赛选择法和轮盘法的区别在于，锦标赛选择法是通过随机选择一定数量的个体，然后从中选择适应度值最高的个体作为父代个体来生成新个体。而轮盘赌选择法是通过计算每个个体的选择概率，然后根据这些概率来进行选择。锦标赛选择法的优点是可以在适应度值相差不大时，更好地保证个体的选择概率与其适应度值成比例。此外，该方法不需要计算每个个体的选择概率，计算量较小。

常 用 的 选择算子包括轮盘赌选择和 锦标赛选择 等方 式 ， 其 中 轮盘赌选 择会保 留 部分劣 质 个体 以 保证群体 的 多 样性和避免 陷 入局 部最优 ， 但也会 陷 入无法 收 敛的 问 题 。 锦标赛选择则 只 保 留 最优个体 ， 无法保证群体 的 多 样性和 全 局 最优解 。

此外，为降低考生作弊的可能性，也给考生最大的公平感，相邻两考生的试卷往往不同。这也是无纸化考试系统优于传统考试的特点之一。种动态出卷方法是当学生登录考试系统，考试程序根据场次和机器的 IP地址分出A、B 卷，只要相邻的机器 IP 地址连续，相邻的试卷便不相同另一种动态出卷方法是当学生登录考试系统，考试程序为每个考生生成一份试卷，这种方法使得每个考生的试卷都不同。

（对试题难度有更多要求）

如何让试卷不同并能同时兼顾到考试的公平公正性?常用的方法有以下三种:

1.系统使用组卷算法生成一份试卷后，每位考生接收到的试卷的题目完全一样，选项顺序由系统随机分配，使得相邻考生的答案不一样，以此防止考生作弊的可能性。但此方法只适用于单选题和多选题。

2.系统使用组卷算法生成一份试卷后，每位考生接收到的试卷的题目完全一样，各个题型的排列顺序由系统随机分配，使得相邻考生的答案不一样，以此防止考生作弊的可能性。此方法克服了第一种方法的缺点但是题目一致使得考生仍存在偷看作弊的可能性。

3.系统使用组卷算法为每位考生生成一份试卷，但是每份试卷的试题的章节和难度一致，此方法克服了前两种方法存在的问题，最好地防止了考生作弊的可能性。是目前较常用的一种动态出卷的方法

本文介绍的自动组卷系统采用 A，B 卷形式使得相邻两考生的试卷试题完全不同的动态出卷方法。

试题库具有 3个特征:

一是组成试题库的基本单位是试题。使用试题库时，可以根据需要从试题库中挑选所需试题组成试卷。

二是试题库具有相应的各种参数指标。题库不是试题的简单集合，而是大量的、有必要参数的有机组合。题库中的每一道试题都有各种参数指标来标明它的性质和特征。试题参数的确定依据题库的性质、内容以及题库理论和建库模型的不同而不同。

三是试题库是一个系统。题库既是一个相对稳定的系统又是一个动态的管理系统。一个题库一旦建成，其内容结构和方式就构成了一个相对稳定的系统。但随着科学技术的发展，社会的进步，职业的要求将不断发生变化，因而相对应的职业技能鉴定试题库的内容与形式以及参数值都需要不断地补充、修改和完善。

综上，将自动组卷系统分为两个部分:

1. 自动组卷:即根据输入的参数从题库中利用随机函数选出符合参数条件的试题组成一份或两份试卷(分 A、B 卷)。此外，考虑到可能有不止一场的考试，例如，由于机房的限制，一门课程可能分为两场进行考试，上午一个班级考试，下午另外一个班级考试，由于是同一门课程的考试,所以两场考试的难度，章节等各参数必需一致，但试题又不能一样，基于该情况组卷系统可以同时生成 1场到3 场考试
2. 查阅修正试卷:该部分主要完成三个功能: (1)对之前自动组织的试卷试题进行查看与修改，可多条同时查看也可逐条查看。 (2)对不满意的试题可进行修改，只要选中某个章节，系统便会自动从该章节中选择符合要求的试题替换原来的试题。 (3)导出数据，可将所有试题以 Excel 表格形式导出。最后导出考试系统，把试题，文档，样例保存到相应位置，等待考试开始分发试卷。

至此，整个自动组卷的基本结构已大致形成，各部分的具体实现方法将在第五章中进行设计与分析。

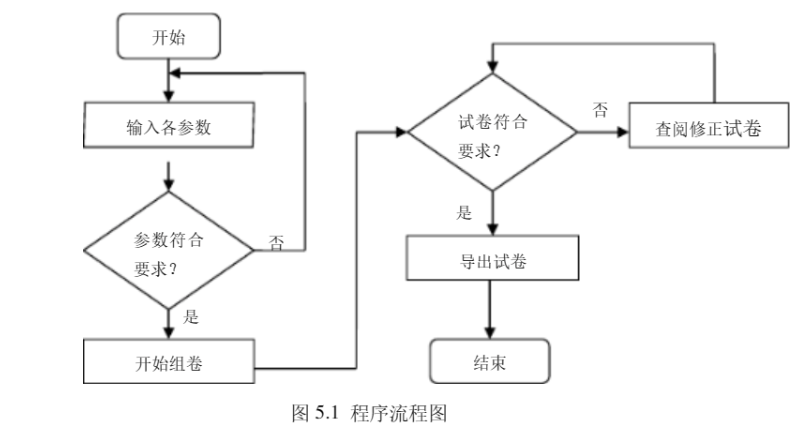
针对系统的需求,本系统中需要建立三个数据库:

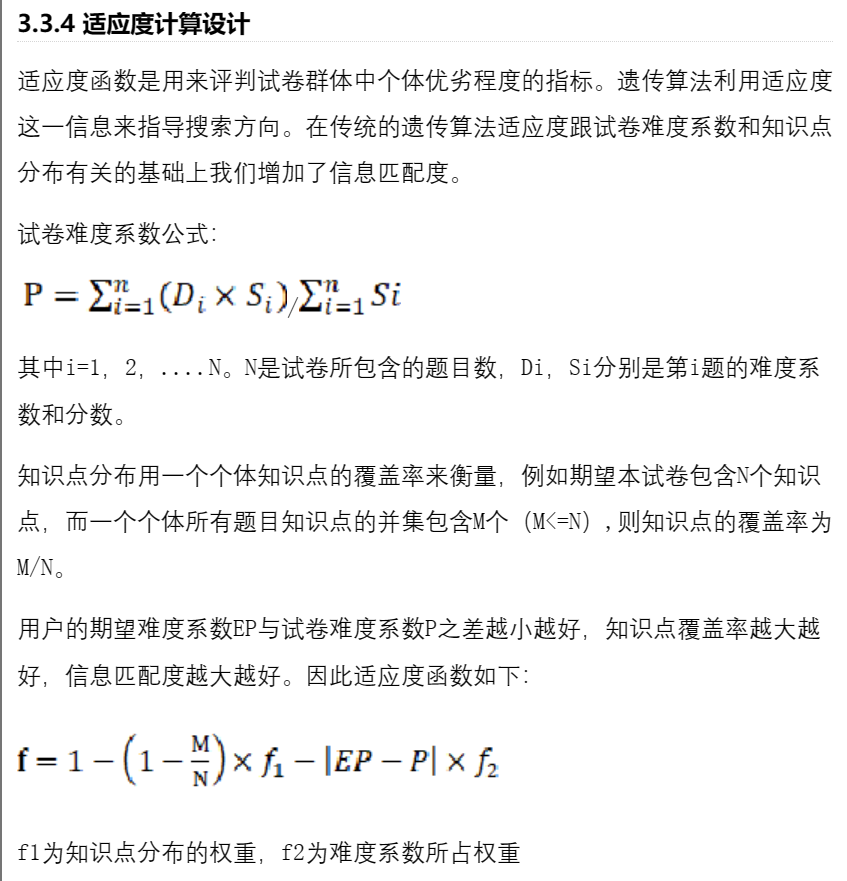
1.试题数据库:用来存放所有试题的相关信息，包括试题内容、试题答案及每个试题所对应的所有参数。此数据库是试卷数据库与考试数据库的基础。

2.试卷数据库:用来存放由组卷算法挑选出来的所有试题内容、试题答案及各个试题所对应的参数。此数据库用来为自动组卷系统的导出考试系统模块提供试题、为无纸化考试系统的自动阅卷模块提供标准答案

3.考试数据库:用来存放每个考生考试时用到的所有试题，包括试题内容，分值等。此数据用来为无纸化考试系统的客户端模块提供试题。

上述三个数据库的结构基本相同，但由于各数据库使用阶段和目的的不同使得各数据库的结构稍有差别。





组卷系 统终止条件主要考 虑 以 下三个方面 ：

（ １ ） 群体适应度总 和 达到预设 阈 值 的 某个 区 间 范 围 内 时 ，

（ ２ ） 循环次 数达到最大迭代次数时 ，

（ ３ ） 连续几代种群平均适应度达到预设 阈值 的某个 区 间 范 围 ， 算法终止

国外还有众多成熟的等级考试系统诸如美国TOEFL、 GRE等考试系统[3]。

国内在组卷系统和在线考试方面相关技术已经日趋成熟，已经从最开始的试题库简单， 系统功能单一，发展到试题库丰富全面，功能日趋完善[4]。并且产生了大量成功 的应用案例，如驾驶证考试系统，科目一和科目四的考试就是一个成熟的在线考 试系统，其通过一定的组卷策略在题库中快速抽取试题，生成试卷以及答案，在 考生答案确定后就对考生答案进行判断并显示出来，考试结束时考生的考试结果 也显示出来。除此之外计算机等级考试、普通话考试等也是一个成熟的组卷考试 系统。

组卷要求

此系统主要是用来协助教师来完成大学生的期末考试试卷的编写。需要充分

考虑学校的考试要求以及本课程的考试大纲要求。

所以此系统必须要满足以下要求：

1) 试题库内试题覆盖要足够丰富全面，高质量的试题库是组卷系统的重要

基础，是完成高质量试卷的前提。其要求试题的难易度，覆盖度，以及题型的种

类等都能满足课程考试大纲要求，这样的系统根据不同的要求时完成的试卷才更

具有代表性、科学性。

2) 能够按照用户要求的各项指标完成试题抽取并组成试卷。教师按照教学

大纲要求来指定考试范围内容、难度等信息，系统能够按照这些要求来完成相应

的组卷。

3) 能够让教师手动组卷，即教师按照自己对教学大纲以及考试大纲的理解

结合教学情况和学生整体学习情况来命题组卷。通过手动选择相应的题型和试题，

同时系统也可以生成对此套试卷的难易度，覆盖度，重复度等相关的评价公交是

作为参考。此功能可以保证在一些要求特殊情况下的组卷，使系统在实际应用中

适应性更强。

4) 系统生成的试卷具有多样性。用户在分析好教学大纲以及考试大纲要求

后在组卷系统中选择好试卷难易度、覆盖度、相似度等属性以后能够快速生成满

足要求的试卷，并且每次运行过后所产生试卷的试题具有差异性。此组卷要求可

以用完成多套备选试卷供考试使用，有利于防止作弊现象出现并有效保证考试的

公平。

5) 能够调节试卷难度。试卷的难易度是评判试卷的一个重要指标，本系统

能够将试题的难易度按照一定的方法策略换算成各大题型的难易度，或者直接将

各试题的难易度转换成试卷的难易度。

6) 能够调节试卷的覆盖度。试题库的所包含的试题覆盖范围要全面，各属

性的设置合理科学。即试题具有科学性，全面性，代表性。此组卷系统能够按照

考试大纲的要求科学调整各章节试题的比重，完成符合一定要求的考试。

基于遗传算法的自动组卷系统优化设计及应用

22

7) 试卷的相似度要尽量的低。为了提高试卷的保密性，一方面要求试卷中

与过去试卷中重复试题的总分值尽可能低；另一方面要求重复的试题距离本试卷

的使用时间尽可能的远。

除此之外，对于系统的操作管理方面，有以下要求：

1) 试题的管理

试题的管理是此系统的基础和前提，所以必须要能够安全、高效的对试题进

行正常的增删改查功能，同时还要能够让试题在系统内部实现共享。要方便管理

者的正常使用和维护。

2) 组卷管理

常见的组卷由随机选题等方法来完成，在约束条件较多的情况下组出的试卷

存在较多不足之处。本系统为基于改进型遗传算法的组卷系统，可以满足试卷区

分度条件、难易度条件、覆盖度条件、相似度条件等多个约束条件来完成高质量

的试卷。过程中主要考虑如何数学建模将试题和试卷结合起来，试卷的各评价指

标量化，以及遗传算法的具体实现过程，以便整个组卷过程的顺利的执行。

3) 用户管理

本系统将主要提供教师使用，同时有系统管理员管理功能，此功能操作简单、

安全便捷，用户可以更改自己的密码等基本信息，防止账户被非法使用。管理员

可以对用户进行管理，对用户进行增删等功能。

组卷只有在满足一定的原则的前提下，抽取满足要求的试题，才能够完成满

足质量要求的试卷。这样的试卷在实际考试当中才能更加有效的反应学生对于课

程知识的掌握情况。这样的考试结果更具科学性，更有利于教师了解自身教学过

程中的不足之处。教师和学生都能够通过这份试卷的结果检测自己，对于后续的

教学和学习具有重要积极的指导意义。组卷的基本准则主要有一下几方面：

1) 科学性原则：试卷要能够科学的反应考试大纲要求，题型选择以及试题

难度、数量等都要科学的选择排版。能够准确的反应实际情况。

2) 目的性原则：好的试卷要在适合它的条件下才能发挥其作用。试卷是用

来作为课程章节检测还是期中考试检测，还是学业水平检测亦或是期末考试检测，

首先要明确。考试的目的不同，那么试卷必然要做出相应的变化调整。

3) 层次性原则：一份试卷不可能全部都是简单或者困难的题目。试卷不仅

要让学生检测出自己学习的不足之处，也要让教师从中得到教学的不足支持。在

试卷编制过程中要充分考虑学生对知识认识掌握的差异性，以及课程中知识点的

第四章系统实现

23

难易度，让试题在试卷中呈现出一定的梯度。一方面是试题本身要有梯度，特别

是一些主观题小问题的难度要递进排布；另一方面，整张试卷在试题排版的时候

要遵循由浅入深，先易后难的排布。

4) 试卷所考核的知识要能够对学生的后期学习方法有一定的指导意义。考

试对学生来说作用非常明显，通过成绩的形式让学生更加深刻的认识到自己与别

人的差距，了解自己对知识的学习掌握程度，认识到不足之处促进其做出改变。

5) 试卷中语言要简洁准确，试卷中不要出现与考试无关的话语。这些有助

于提高试卷科学性、专业性、权威性。