

网络智能试题库系统建设研究^{*}

杨永斌

(重庆工商大学计算机科学与信息工程学院 重庆 400067)

摘要 计算机辅助教学已成为现代高等教育的一个重要组成部分。作为其主要研究方向之一,网络试题库技术与自动组卷系统得到迅速发展。网络试题库系统建设是实现无纸化考试、考试标准化与个性化、远程网络学习辅导与在线测试等的基础。本文对网络智能试题库系统建设中的相关问题进行了探讨,重点对组卷策略和算法进行研究。

关键词 网络,试题库,约束条件,组卷策略,组卷算法

A Study on AI Internet Test Databank System

YANG Yong-Bin

(College of Computer Science and Information Engineer, Chongqing Technology and Business University, Chongqing 400067)

Abstract Computer assistant instruction(CAI) being an important part of modern higher education, test databank and automatic test paper composition system have become one important field of people's research, and met with unprecedented development. The building of test databank system is the foundation of non-paper test, test standardization, learner individualization, on-line learning coaching and testing. The paper is a tentative probe into the relevant problems of building such a databank and in particular into the policy and algorithm of test paper composition.

Keywords Internet, Test databank, Restricting condition, Policy of test paper composition, Algorithm of test paper composition

1 引言

随着经济和社会的发展、网络技术的广泛应用及终身教育观念的普及,迫切需要有科学、方便、完善的网络试题库管理系统,作为积累题目、自动组卷、保存成卷、分析试卷的利器。在先进理论指导下用计算机管理的试题库,具有科学、高效、经济、灵活、方便等显著特点。试题库为题目的保存、分类和检索提供了便利,为编制试卷提供了高效率 and 低成本的手段,更为重要的是它具有作为评价工具的作用。教学过程中充分利用试题库这种教学资源,进行富有成效的形成性测试,可以帮助师生把注意力从得分、名次集中到要实现的教学目标上来,真正关心学生的学习困难和错误所在,关心学生素质的提高,促使教育观念从“应试教育”向“素质教育”转变。因此,进行网络试题库系统的建设和研究具有十分重要的意义。

2 试题库质量标准

2.1 高质量试题库管理系统的特征

1) 试题库管理系统成功地运用了现代教育理论、教育方法和新的教育心理学理论,使教学双方符合教学规律;

2) 试题库管理系统应尽可能地模拟教学专家的思维过程,根据学生的水平和学习情况,自动生成客观、科学的试卷;

3) 系统的设计与实现能充分发挥多媒体计算机的功能特点,并利用现代软件工程思想,使系统体现开放性和适应性,便于修改、扩充和操作使用。

2.2 试卷质量标准

单套高质量试卷标准: 1) 各种题型中知识点分布符合制

卷目标即考试要求; 2) 试题、试卷难度符合制卷要求; 3) 无内容重复和相近的试题。

成批组卷质量标准: 1) 各套试卷符合单套试卷的标准; 2) 任何两套试卷之间的试题重复率不超过指定限度; 3) 同一试题在所有试卷中重复出现不超过指定次数。

2.3 试题库质量标准

为保障由试题库生成的试卷质量, 试题库必须符合以下标准:

1) 试题库不仅要有足够的总题量, 还必须保证各知识点有足够的题量; 2) 同一题型中无完全表述相同的试题, 对表述不同而内容相同或相近的试题必须有标记; 3) 不同题型间内容相同或相近的试题必须有标记; 4) 在同一批组卷中试题的重复使用次数要有记录。

3 试题库结构设计

在试题库系统的建设中, 试题库的结构设计是至关重要的, 良好的试题库结构往往会收到事半功倍的效果。比如要建立一个通用试题库管理系统, 可以建立以下试题库数据模型:

1) 题目内容表(试题编号、题目内容、答案、备注);

2) 题目属性表(试题编号、题型、难度系数、认知层次、知识点、效度、答题时间、题分、备注);

3) 某门课的控制指标表(指标号、指标类型、指标取值), 如一记录值为(01, 题型, 选择题);

4) 题目其他情况表(试题编号、出题人、出题时间、有效期);

^{*}重庆市教委科学技术研究项目资助(项目编号: KJ060719)。杨永斌 副教授, 主要研究方向: 计算机算法、计算机网络应用、数据库技术、网络教育。

5) 题目反馈情况表(试题编号、答对率、良好率)。

另外,还可建立一些诸如知识点设置情况表、大纲设置情况表等辅助支持数据库,以便组卷算法的实现。在实际的试题库系统建设中,如何设计合理的试题库结构,根据系统的实际需求及组卷策略、算法等因素决定。

4 常用组卷策略与算法

4.1 常用组卷策略

一般而言,一门课程的考试通常分为阶段性考试和期末考试。所以,为了充分满足出卷人的各种需求,常有以下组卷策略:

1) 人工出卷

人工出卷是出题人员自己在试题库中逐个挑选试题形成试卷的方法,适用于对试题库内容比较熟悉的教员、管理人员等。出题者可在试题库中按任意组合条件查询,在查询结果中,手工挑出认为合适的题目,挑上哪一个,系统就将此题目的控制标志记录在一临时表中,这张临时表中包含了最新所组试卷的总信息。采用这种方法组卷,试卷的有效度较高,但工作效率较低,一般适合覆盖面小、参加考试人员数量少等特殊情况,如单元测试等。

2) 全自动出卷

全自动出卷是出题人员仅提供一个简洁的命题计划,试题库管理系统就会按照系统预设的算法模型在试题库中自动挑选合适的题目。

3) 半自动出卷

半自动出卷是介于手工出卷和全自动出卷之间的一种组卷方案,命题计划的内容要比全自动出卷详细得多。

4.2 常用组卷算法

1) 随机算法:即随机从试题库中抽取试题,其优点是算法结构简单、运行速度快;缺点是由于约束条件少,组卷中试题重复率、相似率高,试题的总体难度难以把握,尤其是试题库中题量较少时。

2) 回溯算法:将随机选题产生的每一状态类型记录下来,当搜索失败时,释放上次记录的状态类型,然后再根据一定的规律变换出一种新的状态类型进行试探,通过不停地回溯试探直到试卷生成完毕。该算法理论上可以遍历每一种可能的状态组合,但实际上当试卷总题量较大时,状态类型的组合数便成为天文数字,尤其是当试卷生成到后来,符合约束条件的可选试题量的急剧减少,使得生成的成功率大大降低。所以,此方法只能适用于状态类型和试卷总题量不多的试题库系统。而且该算法对内存占用量大,程序结构复杂,一般不适合在线组卷。

3) 随机启发式搜索算法:基于随机定位,折半搜索,其中心思想是:产生有效随机数 K ,根据题目选题状态判断是否已经选中。若选中,选下一题;否则,产生随机动态偏差数 S $[0,1]$,若 $S < 0.5$,向前搜索,否则向后搜索。确定搜索区域后进行折半搜索,产生随机数 K 再进行判断,重复上述过程直到选中为止。

5 遗传算法在组卷中的应用

5.1 组卷的数学模型

自动组卷问题实质上是一个约束的组合优化问题,即从一定题量的数据库中搜索满足组卷目标要求的一个试题组合,从而把组卷问题转化为一个多重约束目标问题,该问题将

被定义为一组约束条件与一个目标函数的组合。

1) 约束条件

为了考核学生对知识的掌握情况,无论如何设计试题库结构,其中每道试题最重要的属性有:①题分;②难度系数;③知识点,试题内容所属的知识层次范围;④认知分类,反映该题对教学内容的要求层次;⑤题型;⑥估计答题时间;⑦区分度;⑧试题编号等。组卷中决定一份试卷,就是决定一个 $n \times r$ 的目标矩阵。其中, n 是试卷所含的题目数, r 表示满足 r 重约束条件,对应 r 个重要属性,具体约束由用户自行设定。每道试题的 r 个属性决定矩阵一行元素的值,即

$$T = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1r} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2r} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nr} \end{bmatrix}$$

目标矩阵的主要约束条件有(这里假设 7 个主要约束条件):

①总分条件 $\sum_{i=1}^n a_{i1} = \text{total_mark}$,一般设定为 100

②难度系数条件 $\sum_{i=1}^n a_{i1} a_{i2} / \text{total_mark} = \text{全卷难度}$

③知识点条件 $\sum_{i=1}^n b_{ij} a_{i1} = Z_j$ (第 j 知识点的题分),其中

$$b_{ij} = \begin{cases} 1, & a_{i3} = j \\ 0, & a_{i3} < j \end{cases} \quad j \text{ 为知识点}$$

④教学要求条件 $\sum_{i=1}^n b_{2i} a_{i1} = P_j$ (第 j 认知分类的题分),其中

$$b_{2i} = \begin{cases} 1, & a_{i4} = j \\ 0, & a_{i4} < j \end{cases} \quad j \text{ 为教学要求号}$$

⑤题型要求条件 $\sum_{i=1}^n b_{2i} a_{i1} = Q_j$ (第 j 题型的题分),其中

$$b_{2i} = \begin{cases} 1, & a_{i4} = j \\ 0, & a_{i4} < j \end{cases} \quad j \text{ 为教学要求号}$$

⑥考试时间条件 $T = \sum_{i=1}^n a_{i6}$

⑦区分度条件 $D = \frac{1}{\text{总分}} \sum_{i=1}^n a_{i1} a_{i7}$

2) 目标函数

矩阵 T 的列元素的分布分别满足用户指定的试卷的总体要求,也就是所有试题的分数总和、各种难度的试题所占的分数比例、各篇章所占的分数比例、反映不同教学层次要求的试题所占的分数比例、试卷中各种题型所占的分数比例、所有试题的估时总和、总卷区分度等约束都应等于用户指定的要求,或者误差最小。在实际应用中,设定 f 来综合反映这些约束与用户要求的误差。由于它们的重要程度不同,因此整卷的误差 f 就是 r 个约束的加权和。同时,为了不至于各个误差相互抵消,这 r 个约束与用户要求的误差都取绝对值。可以用下式表示:

$$f = \sum_{i=1}^r f_i w_i$$

其中 f_i 表示第 i 个约束与用户要求的误差的绝对值, w_i 是组卷人员对第 i 个约束条件所赋予的权重。由于在实际组卷过程中,有些约束条件比较重要(如分数,知识点等),为了满足它们的需要,可以适当牺牲其他约束;而另外一些约束,可适当放松取次优解。因此,根据这些约束条件不同性质和抽卷者的不同目标,对 w_i 可自行赋值,加大权重。组卷就是从题库中抽取试题,使得整卷的误差 f 最小。因此,组卷问题

可以看作是在约束条件下的最小化问题。

5.2 组卷问题的遗传算法设计

1) 传统的遗传算法

① 染色体编码方法(基于二进制编码)

设题库中总共有 m 道题,用 S_1, S_2, \dots, S_m 表示这 m 道题,那么组卷就是从 S_1, S_2, \dots, S_m 中选出 n 道题,使得整卷目标函数 f 取最小值。可用一个 m 位的二进制串来表示问题解。形式如下: $F_1 F_2 \dots F_m$ 。如果 F_i 为 1,则表示该题被选中;若为 0,则表示该题未被选中。

② 初始化群体

通过随机的方法生成初始化的串群体。在串群体中,串长度都是相同的,群体大小根据需要,按经验或试验给出,分布相对均匀的二进制串能使算法更加有效。在实际组卷中,将群体规模设为 n ,每个个体都是通过 $\text{random}(m)$ 这个随机函数在 1 至 m 中随机选择 n 道题。

③ 计算当前群体中的每个个体的适应度

首先通过解码,得到个体的组卷参数矩阵 T 。具体过程如下:根据串 $F_1 F_2 \dots F_m$ 的值,可知试卷中包含的试题的题号,然后把把这些试题的属性写到组卷参数矩阵 T 中(由此可以看到:组卷参数矩阵 T 的值是由题库中的试题属性组成的,这就保证了按照矩阵 T 来选题时不会出现找不到题的情况)。然后调用评估函数得到每个个体的适应度。在实际应用中,可以通过计算个体的整卷指标 f 来反映个体的适应度,整卷指标 f 越小,个体的适应度越高。

④ 选择

根据适应度对群体中的个体排序,并对每个个体分配一个选择概率,所有概率的和为 1。适应度高的串(整卷指标 f 相对小的个体),就有较大的选择概率生成下一代。

⑤ 杂交

两个串按照一定的概率,对串随机地选择一个交换点进行杂交,即两个串按照一定的概率(交换概率设为 P_c)从某一位开始逐位互换。先在群体中随机地选择两个串,然后对每对串随机选择一个交换点。如果交换概率 P_c 满足 $\text{random}(1) < P_c < 1 - P_m$ ($\text{random}(1)$ 表示 0 至 1 之间的一个随机数, P_m 为变异概率),则这两个串交换,否则两个串保持原状。 P_m 一般由经验或试验决定,通常为 0.4 至 0.6,一般取 0.5。

⑥ 变异

变异是二进制串的某一位按照一定的概率(突变概率 P_m)发生反转,即 1 变为 0,0 变为 1。这里 P_m 较小,可小于 0.01。但在实践中发现,在有些遗传算子中, P_m 为 0.1 时更好。

⑦ 修正非法解。每个问题的解(个体)进行遗传操作以后,很可能变为非法,比如串 $F_1 F_2 \dots F_m$ 经过交叉或变异后,串中 1 的个数很可能大于或小于 n ,这时这个串对应的解就是非法的,必须对解进行修正。具体做法就是进行相应的运算,使得串中的 1 的个数为 n 。

⑧ 判断终止条件是否满足

$t := t + 1$, t 为进化的代数,并按照③的方法计算当前群体的每个个体的适应度。如果 t 等于指定的进化代数,或者某个个体的适应值达到要求就退出,否则转④继续执行。

2) 改进的遗传算法(基于矩阵编码)

基于传统的二进制编码的遗传算法可以从很差的个体逐渐搜索到较好的个体,对领域知识要求也很少,可以说是一种通用的有效的解决方法。但它的求解质量和效率都不理想,

计算时间也很长。通过对算法做进一步的分析,发现遗传算子是直接作用在串 $F_1 F_2 \dots F_m$ 上,而这个串本身并没有提供多少与实际问题有关的信息。也就是说,传统的遗传算法是一个共性的解决方案,而在实际应用中没有个性的方案是没有吸引力的。因此要得到理想的结果,必须改变解的编码,使它更好地反映实际问题,以便进行遗传操作时,可以利用更多的信息。另外,组卷工作是为了组成一份误差在可接受范围内的试卷,并非要求此卷的整体指标一定是全局最优,因此可以考虑选取较好的初始解,以加快搜索进程。基于以上考虑,可以采用矩阵编码,即直接在组卷约束矩阵 T 上进行遗传操作,而初始解则是采用一定的限制方法得到。采用矩阵编码的遗传算法的步骤与传统的遗传算法一样,不同的是编码和遗传操作。假设使用上述组卷约束矩阵 T 进行编码,则其某个候选解可能是如下所示:

知识点	题型	分数	难度	区分度	认知分类	估时(分钟)
02010000	填空题	2	容易	差	识记	3
01020000	选择题	2	一般	较差	理解	3
.....						
09010000	应用题	10	难	良	综合	15

在实际组卷中,对试题库的试题分布是有一定的要求的,基本上所有要考核的知识点、题型都有相关的试题,而题型跟分数一般也是紧密联系的。当计算组卷约束矩阵 T 时,试题库中都有与知识点、题型和分数属性相同的试题,误差全部出现在难度、区分度、认知分类、估时等属性上,这样遗传操作所要做的工作主要是匹配这些属性。因此,遗传算子的设计如下。

① 杂交算子的定义:令 $\text{oper}_k = \text{random}(k)$, $\text{random}(k)$ 指从 0 到 $k-1$ 的自然数中随机取一个(k 为遗传操需要匹配的的属性个数)。如果交换概率 P_c 满足 $\text{random}(1) < P_c < 1 - P_m$,就交换两个父本的第 $(3 + \text{oper})$ 列以后的列,否则直接复制两个父本。通过互换可以达到全局搜索的目的。

② 变异算子定义为:在一定的变异概率 P_m 下,从候选解(组卷约束矩阵 T)中随机选一行,根据这一行第二列的值得到题型属性,然后从题库中选取另外一道同题型的试题的属性取代候选解的这一行。通过变异算子可以达到局部搜索的目的。

3) 算法描述

不同的编码方案、选择策略和遗传算子相结合,构成了不同的遗传方法,采用的重叠种群的遗传算法如下:

```

{
  随机初始化种群 P(0), t=0;
  计算 P(0) 中个体的值;
  While(不满足终止条件) do
  {
    根据个体的适应值及选择策略,计算种群内个体的选择概率 Pi;
    根据 Pi 从 P(t) 中选择 N1 (<N) 个个体进行遗传操作,用产生的 N1 个后代替换掉 P(t) 中 N1 个较差的个体,生成新种群 P(t+1);
    计算 P(t+1) 中个体的适应值, t=t+1;
  }
}

```

选择何种编码方案一般会对算法的性能、效率产生很大的影响。根据实际问题设计编码方法,提高效率是遗传算法的一个研究热点。

结束语 在网络智能试题库系统建设中,自动组卷是系统的关键。自动组卷研究中遇到的主要难题是如何保证生成的试卷能最大程度满足用户的不同需要,并具有随机性、科学

(下转封三)

第 11 期

普适计算中间件技术的研究与进展	(1)
分布式安全存储技术	(6)
企业服务总线研究综述	(13)
分布式动态更新支持系统:研究综述	(19)
有界 Petri 网的最小化简化	(26)
Mesh 网络耐故障虫孔路由	(29)
一种超节点 P2P 网络中基于语义的协同缓存管理机制	(32)
AIMD 组播拥塞控制在卫星 IP 网络中的性能分析及改进	(37)
基于有环因子图的密度进化理论研究	(41)
一种主动网络管理系统结构模型的研究	(44)
IPv4/IPv6 网间的 SIP 通信的解决方案	(47)
网络组合攻击建模方法研究进展	(50)
一种有效的匿名分析算法	(56)
差分功耗分析单片机 DES 加密实现的旁路攻击	(58)
基于动作检测的交互模式研究与应用	(62)
基于模糊粒子群优化的计算网格工作调度算法	(64)
基于强化学习的多机群网格资源调度模型	(67)
P2P 资源的拍卖激励机制	(71)
异构分布式环境中 Alhusaini 资源映射算法的改进	(74)
基于资源约束的流媒体服务 QoS 策略	(79)
一种改进的自适应多媒体任务调度算法	(82)
基于 SVG 的实时 Web GIS 技术研究	(86)
基于二级向量描述的搜索引擎个性化服务模型	(89)
Web 服务器系统中基于反馈控制的比例延迟保证	(93)
基于 Web 服务的电子政务数据交换中心的设计和实现	(98)
社区服务管理系统中服务网点配置算法的研究	(102)
面向公民的电子政务服务框架(CCeGSF)	(106)
基于 J2ME 和混沌加密的移动商务安全方案	(110)
同构化信息温度与热点发现应用初探	(113)
基于时态逻辑的工作流分析	(118)
支持共享的 Ftp 客户端的设计	(124)
基于 VFW 的远程专家会诊系统中的实时音视频捕获	(127)
形式概念分析与集值信息系统	(129)
DCSP 和 DCOP 求解研究进展	(132)
基于模糊逻辑的智能 agent 情感建模	(137)
“软件人”感知系统的协同分类模型研究	(140)
一种基于时间-效用的 Agent 社会承诺机制	(144)
量子克隆遗传算法	(147)
一种改进的抑制早熟收敛的模糊遗传算法	(150)
基于正交多主体遗传算法的业务规则引擎推理方法	(154)
基于神经网络的 Agent 增强学习模型	(156)
一类时滞神经网络的全局指数稳定性	(159)

进化算法及其在入侵检测中的应用	(162)
粒子群算法种群结构与种群多样性的关系研究	(164)
一种基于多区域不变矩的步态识别方法	(167)
数据流的预测与分类研究	(170)
一种基于粗糙 k 均值的双聚类算法	(174)
一种基于熵的聚类算法	(178)
支持向量逐步回归机及其改进算法研究	(180)
GB18030 汉字的分形相关性研究	(183)
汉字字形形式化描述方法研究	(185)
句法标注的一般模型与参数分析	(189)
Google 地图算法研究及实现	(193)
一种基于 Contourlet 的图像低码率 SPIHT 算法	(196)
文本图像的双边总变分超分辨率恢复方法	(201)
基于 ARNOLDI 过程的参数估计	(205)
双域单向水平倾角最小化围绕凸壳新算法	(208)
四元数与彩色图像边缘检测	(212)
QR 码图像的矫正与定位方法研究	(217)
一种基于密钥回收的量子加密算法	(220)
一种保持码流结构的 JPEG2000 加密算法	(222)
数字指纹编码与嵌入假设	(224)
一种基于共轭混沌映射的数字水印算法研究	(229)
算法程序变换研究与进展	(232)
C4ISR 体系结构全信息分析模型	(239)
CMMI V1.2 版本的新特征	(241)
功能点分析方法研究	(245)
一个面向服务的应用案例研究	(252)
一种适合分布式虚拟环境的 XML 路由技术	(255)
多核处理器降低功耗技术综述	(259)
一种基于本体的软件自适应机制	(264)
无线移动环境中基于移动 Agent 的实时构件组装机制研究	(270)
结合 AOP 与反射机制动态改变软件的行为	(274)
混合投影时序逻辑与混合系统的形式化验证	(279)
软件测试中的窗口测试	(283)
硬盘读写通道的 FIR 噪声预测算法与仿真模型	(286)
面向评估的高产出率计算指标量化分析	(290)
负载感知的存储子系统调优研究	(293)
乱序执行机器上的 load 指令调度	(298)
C/S 遗留系统到 SOA 系统移植框架研究	(301)
基于 RTAI 的实时 Linux 系统构筑及其嵌入式程序移植	(304)
支持普适计算的多嵌入式系统	(308)

第 12 期(略)

(上接第 298 页)

性、合理性,而在网络交互式环境下,用户对组卷速度的要求较高。因此,高效、科学、强壮的智能组卷算法是目前题库建设领域的研究热点之一。在网络题库系统中,通过自动组卷,将人工智能技术与教育专家的组卷知识和经验结合起来,由计算机来完成试卷内容的设计,生成达到专家级水平的试卷。一般而言,在一个题库中,叙述题、分析题、设计题、应用题等的题量较少,而判断题、选择题、填空题等的题量较多,在具体组卷时,可以针对不同的题型可采用不同的参数匹配算法,并对试题参数进行动态跟踪修改,使组成的试卷能更好地满足用户的需求。

参考文献

- 1 李敏强,寇纪淞,等.遗传算法的基本理论与应用[M].北京:科学出版社,2003
- 2 王小平,等.遗传算法—理论、应用与软件实现[M].西安:西安交通大学出版社,2002
- 3 余胜泉,何克抗.网络题库系统的设计与实现[J].中国远程教育,2000(9):53~57
- 4 于洋,查建中,等.基于学习的遗传算法及其在布局中的应用[J].计算机学报,2001,24(12):1242~1249
- 5 高玮.改进的快速遗传算法及其性能研究[J].系统工程与电子技术,2003,25(11):1428~1430