

文章编号: 1672-058X(2009)04-0356-04

计算机操作测评系统探索

徐子珊, 李盛瑜

(重庆工商大学 计算机科学与信息工程学院, 重庆 400067)

摘要: 研究并分析了对同一个文档生成系统产生的两个文档进行自动比对, 并计算相似程度的系统, 此系统是很多应用的关键; 例如在测试学员对某种文档生成系统使用、掌握水平的测评系统中, 必须对学员生成的文档与样板文档加以比对并且作出评估; 分析了用于考察评判计算机操作技能 (MS-OFFICE 软件操作) 的智能系统的构成, 核心算法和实现细节。

关键词: 计算机操作技能测评; MS-OFFICE 文档结构; 最长公共子序列; 批处理; 在线处理。

中图分类号: TP391

文献标志码: A

1 问题的提出

对同一个文档生成系统产生的两个文档进行比对是很多应用的关键。例如, 在测试学员对某种文档生成系统使用、掌握水平的测评系统中, 必须对学员生成的文档与样板文档加以比对并且作出评估。随着信息技术在社会生活中的不断发展、深化, 需要对大量的学生和从业人员进行这样的技能测评。所以, 需要对大量的受测评者的作业进行高效、准确的测试评判工作。必须指出, 由于对一种文档生成系统的操作技能与做一道简单数学计算问题不同, 它包括受训者对系统工作的理解、系统中各个操作的意义、正确的操作程序等方面。所以, 在此所谈的准确评判, 指的是对测评对象掌握计算机操作技能水平与程度的更深意义的准确, 而不是进行简单的错与对的评判。对文档产生系统具有资历的测评人来说, 对这样的测评工作是可以胜任的。但是, 当需要测评的作业是海量的时候, 测评者就要受到数量、精力甚至是体力的挑战。虽然近年来有一些测评系统面市, 然而, 由于对文档生成系统所产生文档的内部结构不甚了解, 系统作出的评判往往是简单且对测评点的涵盖面是有限的。一个能对受测评作业作出智能评判的系统, 对于这一社会生活领域而言, 需求是迫切的。

2 系统描述与构成

2.1 系统架构的考虑

一个实用的测评软件系统必须满足如下的条件:

(1) 普适性。系统能在最广泛意义下用于测评者和受测评者;

收稿日期: 2008 - 12 - 13; 修回日期: 2009 - 01 - 11。

作者简介: 徐子珊 (1954 -), 上海市人, 副教授, 从事算法设计与分析研究。

(2) 易用性。测评者和受测评者能以最通行的操作方式来使用;

(3) 可靠性。系统得到的评价信息的保存、检索和迁移的可靠程度必须不低于传统方式下对这些信息管理的水平。

鉴于这些要求,将构建一个跨平台的系统,除了受测评者完成作业因为需视文档生成系统所适应平台外(这个环节本质上并不包括本系统),受测评者作业的提交,测评方对数据库中信息的访问与管理可以在任何一种平台上进行,服务器端的程序亦可运行各种平台上。

Internet 的 Web 服务已深入人心。即使是首次接触计算机的人,最先尝试的操作一定不会缺少对 Internet 浏览器的使用体验。其次,通过 Internet,系统的使用就能最大限度地摆脱对地理位置的限制——能接入 Internet 的地点就能适用本系统。因此,测评系统应当是基于 Web 的 B/S 构架。

信息系统(考试评判系统本质上是一个信息系统)的安全性和可靠性是离不开数据库技术的。为使系统本身既能利用数据库技术获得信息的可靠性要求,又能保证系统开发的合法性(主要是知识产权的合法性),开发技术使用开源的 JAVA 开发平台和 MYSQL 数据库服务器。

2.2 系统的基本功能

作为一个作业(试卷)评判系统,应当具有如下的基本功能:

(1) 题库的建立与管理。题库是一个数据库,包括题面数据表、题目考点解析表等;

(2) 受测评者信息数据库的建立与管理。受测评者信息数据库包括受测评者基本信息表、作业(试卷)表、试卷得分点解析表等;

(3) 受测评者作业(试卷)的上传。这是系统对受测评者的服务接口;

(4) 对受测评者作业(试卷)分析、评判。这是系统的核心服务,利用合适的技术和算法实现对受测评者作业(试卷)的分析和评判;

(5) 对受测评者作业评判结果的统计、分析、存储、检索。这是主要面向测评方的服务接口,测评者可通过这个接口得到受测评者的作业评价信息。

2.3 系统的构成

根据上面的描述,用图 1 来展示系统的构成。

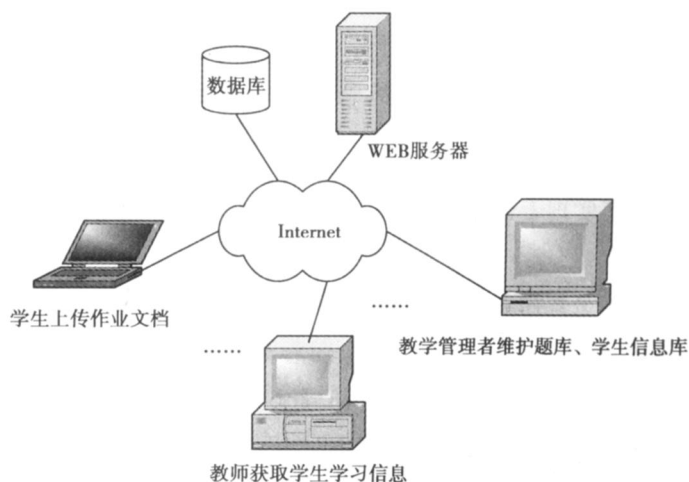


图 1 计算机操作作业(试卷)智能评判系统构成图

3 核心算法

3.1 算法的基本描述

系统的核心是对学生的作业(试卷)分析、评判的服务程序。该程序的算法可描述为:对题库中相应的文档析取文档内容(通常是文本、图形、表格等)和格式信息(通常包括文档格式、页面格式、对象(文本、图形、表格)格式等);对学生数据库中相应批次的学生作业(试卷)逐一析取文档内容和格式信息;将每个学生的作业(试卷)的内容和格式信息与题目内容和格式信息作为两个字符序列,利用生物信息学中使用的计算最长公共子序列算法,计算出这两个字符序列的最长公共子序列;用合适的记分方法计算出每个学生的作业(试卷)成绩。

其中,析取工作是算法的“物质基础”,能够正确、完整地析取所需文档信息,是求最长公共子序列进行对比计算的基础。理论上只有能够解析文档的完整格式,才可能析取出各类信息。然而,商业软件的文档格式一般不开源,所以这是系统开发的一个难点,也是重点。早期面市的一些系统往往在这一点上处理的捉襟见肘,难以推广实用。每个学生的作业成绩的计算视文档产生系统的差别、测试水平标准的差别而定。

3.2 最长公共子序列及其算法

早期的评判系统往往是将学生文档与样本文档作简单比对,或对或错。这样的系统距离实用性确实还有距离。本系统运用“最长公共子序列”的概念,计算学生文档与样本文档的最长公共子序列,用以计算两者的“相似”程度,并以此作为对学生成绩的评判。

已知序列的子序列是在已知序列中去掉零个或多个元素后形成的序列。例如, $Z = \{B, C, D, B\}$, Z 是 $X = \{A, B, C, B, D, A, B\}$, Z 的一个子序列。

给定两个序列 X 和 Y ,序列 Z 是 X 和 Y 的一个公共子序列,若 Z 同时为 X 和 Y 的子序列。例如,若 $X = \{A, B, C, B, D, A, B\}$ 且 $Y = \{B, D, C, A, B, A\}$,序列 $\{B, C, A\}$ 是 X 和 Y 的一个公共子序列。然而,它不是 X 和 Y 的一个最长公共子序列,这是因为它的长度为 3,而长度为 4 的序列 $\{B, C, B, A\}$ 也是 X 和 Y 的一个公共子序列。序列 $\{B, C, B, A\}$ 是 X 和 Y 的一个最长公共子序列, $\{B, D, A, B\}$ 也是,这是因为没有长度为 5 或更大的公共子序列了。直观地看,两个序列的最长公共子序列越长,相似程度就越高。生物信息学中常用这样的基本想法来比对两个 DNA 链,以发现两者之间存在的生物学方面的关系。利用这一方法来比对学生作业与样本之间的相似程度,得到对学生掌握的计算机操作技能的客观且人性化的评价,使其体现出一定程度的智能性。

4 细节处理技术

为了实现上述算法,成为实用的作业(试卷)评判程序,需要考虑很多技术细节。下面讨论其中最重要的 4 点。

(1) 分离作业内容与文件中的格式信息。对受测试者所做的办公文档进行评判,常常需要做两个方面的考虑:文档内容与对象(字符、段落等)的格式设置。对两个方面分别作出评判。可以通过分别析取文件中内容与格式设置而将它们分离开来,分别用最长公共子序列的方法进行评判,使得我们可以对学生的技能做更细致的分析,做出的评判更准确更具人性化。

(2) 析取文档内容与格式信息的技术。以 MS-OFFICE 文档为例。早期的方法是将文档转换成 HTML 格式文件,然后在 HTML 文件中析取信息并加以计算。这个方法受很多限制。首先,不是所有的 Office 文档的格式都有对应 HTML 标志。所以一些操作技能测试点将无法用这个方法体现出来。其次,Office 文档转换成 HTML 文件后,有很多属于 HTML 文件本身的信息,这些信息对本系统来说是冗余的。

可以考虑对 Office 文档本身的结构加以分析,从中析取足够的有用信息。这是能够做到的。Office 文档在 2007 版之前是 MS 的 OLE 2 组件文档格式,利用当前成熟的对 OLE 2 组件文档的读写操作技术就能达到设计的目标。2007 版以后,Office 文档格式为 Open XML,这是一种更利于用户使用的开放式格式文档。

(3) 胖客户与瘦客户。早期的基于 Web 的应用往往把所有的事务处理都交给服务器计算,使得客户端很轻巧,这对于事务处理(数据库存取操作)计算量较小的应用应该是不错的结构。但是,随着 Web 技术应用不断深入到社会生活的各个层面,事务处理计算越来越复杂,服务器需要响应的事务计算量越来越大。如果把所有的事务计算都交给服务器,就可能引起服务器堵塞甚至崩溃。反观客户端,如今的 PC 计算能力完全相当于当年小型机的能力。所以,在不增加网络数据吞吐量的前提下,让客户端程序完成更多的事务

计算可大大提高系统的运行效率。系统可利用 Ajax 技术,将学生作业文档内容的析取等部分计算在客户端进行,适当地给服务器“减肥”。

(4) 批处理与在线处理。如果是开发成一般的基于 Web 的系统,若租用的服务器(即使是专用服务器,若进出口带宽受限其效率也未必能达到实用要求)负担较重,在线系统不一定能及时响应受测试者的请求(对其提交的作业加以评判并给出结果),对于这样的基础设施,不妨将系统设计成批处理方式,测评方可在适宜的时候对存储在数据库中的信息成批计算,并在数据库中存储计算结果供受测试者查询。当然,高速的在线系统应当是发展方向,必要时(在同一时间段有众多考生接受测评)可以考虑用服务器群来提供更实时化的系统。

5 测评环境与题库建设

尽管测评环境(受测评者使用的文档生成系统及其运行平台)不必集成于测评系统内,但是任何种类,任何级别的测评都有关于测评内容和难度的标准大纲。作为测评系统,题库建设与评判标准都应当遵照国家(地方)颁布的相应等级的测评大纲。此外,测评时需要考虑软硬件环境的一致性,譬如,所用的文档生成软件系统版本应当是一致的。这样,才能保证存储的信息安全、一致,对学生成绩的评判保持一致。

题库建设的成功与否,是一个测评系统能否成功的关键所在,也是最重要的基础设施。而好的题库最重要的指标就是要有足够的与测评大纲相匹配的题目。这件工作离开了教师的日常维护是不行的。很多测试评判系统把题库做成封闭式的或只能由专门人员维护的半封闭式的,这样的系统其生命周期是可想而知的。仍以 MS-OFFICE 作业(考试)评判系统为例,其题库中的题目如果就是 Office 的原文档,测评方的教师只要向题库上传文档就可以参与题库建设,则题库中的题目就能保证与时俱进,延长系统的生命周期。

参考文献:

- [1] THOMAS H. 算法导论 [M]. 北京:高等教育出版社,2002
- [2] ECKEL B. JAVA 编程思考 [M]. 北京:机械工业出版社,2002
- [3] 何自聪. Ajax 开发精要 [M]. 北京:电子工业出版社,2006
- [4] 扎卡斯,曹力. JavaScript 高级程序设计 [M]. 北京:人民邮电出版社,2006

On computer operation evaluation system exploration

XU Zi-shan, LI Sheng-yu

(School of Computer Science and Information Engineering,
Chongqing Technology and Business University, Chongqing 400067, China)

Abstract: According to the demand of teaching reform, this paper studies and analyzes intelligent system components used to examine computer operation skill (MS-OFFICE software operation) and its core algorithm and realization detail and gives suggestions for examination environment and test-paper database construction.

Key words: computer operation skill evaluation; MS-OFFICE file structure; longest public sequence; batch treatment; online treatment

责任编辑:代晓红