

基于 web 的远程教育系统研究

摘 要

本文通过对基于 web 的远程教育系统的深入研究,利用面向对象的软件工程方法对系统进行需求分析,系统设计和开发。

本系统针对传统的远程教育系统模式单一,学生只能被动地接受完全相同的学习内容,不能主动地控制自己的学习过程,给出了个性化学习的概念,并且采用数据挖掘技术,得出个性化的远程教育系统的模型,实现后的个性化的远程学习是一种自主的带反馈机制的学习,它使学生能够更多更高效地获取所需要的教育信息资源。

在需求分析过程中,针对个性化的远程教育系统,重点分析了学习子系统中的自主学习和答疑两个部分,采用了面向对象的分析方法,运用 UML 标准建模语言和 Rational Rose 建模工具对远程教育系统进行了建模,给出了一个反映问题域的 OOA 模型——类图;在系统设计上,采用面向对象的设计方法,其中数据库的设计分为两个库,一个库是存放针对每个学生的详实的个人资料的学生库,另一个是存放教学资源的教学资源库,并且给出了基于 Web 的 Data Mining 技术 (Web Mining) 的个性处理模块的设计;在实现上,采用了 B/S 模式的总体结构,从逻辑上分为三层,即客户端,事务逻辑层和数据存储层。在设计模式上采用了模型视图控制器,通过 Struts 框架(一种基于 MVC 设计模式的 WEB 应用框架)来实现,在该模式设计中具体用到了三个主要的设计技术——JavaBean, Jsp 和 Servlet;在测试过程中继续运用 OO 技术,进行以对象概念为中心的软件测试。

在系统的功能实现上，基于 Web 的远程教育系统是一个规模庞大的系统，涉及范围大，且结构复杂，，本文着重阐述了学习系统中的自主学习子系统和答疑子系统。

基于 web 的远程教育系统主要以面向对象技术为系统开发的指导原则，并采用 web 技术和数据库技术，使用 Java, SQL Server 2000 等工具加以实现的。

关键词：面向对象，UML，Rational Rose，B/S 模式，数据挖掘，模型视图控制器，远程教育系统

THE DEVELOPMENT RESEARCH OF REMOTE EDUCATION SYSTEM BASED ON WEB

Abstract

This text, through the development research of remote education system based on web, uses the theory and method by software engineering to carry on the system analysis, design and develop.

Because of the simplex of current distance learning system, students can't learn according to their needs actively. The system solve this problem, we present the concept of personalized service, design the personalized distance learning model by data mining. Personalized long-range realized is a learns being that act oneself as well as coupling back study , it makes a student can gains education information resources effetely.

In the process of analysis, specifically for the personalized distance learning system, priority analyses the autonomously studying and answering questions in studying subsystem. we use the Object-Oriented method as the software development thought-way in the analytical process, make use of UML standard model language and Rational Rose model tool to carry on setting up a model to the educate system. And we product the OOA model which reflect the problem regions. On designing, we use the Object-Oriented design method, Among them, the data base have two mark, One part is a student warehouse which store every

student's individual data, another is a teaching resource warehouse. We give the personalized treatment module design based the Web Data Mining technology (Web Mining). On realizing, the text's ensemble architecture adopted B/S mode, divide into three layer, they are customer end, affairs logic layer and data are stored layer. In model designing , the text adopted model-view-controller(MVC)model, and realize through struts frame(the webs of mode use frames on the basis of MVC). There are three main implementation techniques used in MVC model of designing——bean, Jsp and Servlet. In testing process , We Continue use the OO to carry out the software testing .

This system is very big. The function elaborate emphasized the studying sub- system, answer question sub-system and examine sub-system in study management system.

The remote education system based on web wield the Object-Oriented technology to be the principle, adopt the web technology and the data base technology, , and use Java , SQL Server2000 to realize.

KEY WORDS: Object-Oriented, Unified Modeling Language (UML), Rational Rose , B/S model, Data Mining, model-view-controller(MVC), Remote Education System

独立完成与诚信声明

本人郑重声明：所提交的学位论文，是本人在指导教师的指导下，独立进行研究工作所取得的研究成果并撰写完成的。没有剽窃、抄袭等违反学术道德、学术规范的侵权行为。文中除已经标注引用的内容外，本学位论文中不包含其他人或集体已经发表或撰写过的研究成果，也不包含为获得华北水利水电学院或其它教育机构的学位或证书所使用过的材料。对本文的研究做出重要贡献的个人和集体，均已在文中作了明确的说明并表示了谢意。本人完全意识到本声明的法律后果由本人承担。

学位论文作者签名：任晓霞

签字日期：2007.6.6

保证人（导师）签名：

任晋林

签字日期：2007.6.7

学位论文版权使用授权书

本人完全了解华北水利水电学院有关保管、使用学位论文的规定。特授权华北水利水电学院可以将学位论文的全部或部分公开和编入有关数据库提供检索，并采用影印、缩印或扫描等复制手段复制、保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文原件或复印件和电子文档。（涉密的学位论文在解密后应遵守此规定）

学位论文作者签名：任晓霞

签字日期：2007.6.6

导师签名：

任晋林

签字日期：2007.6.7

前言

题目的来源

自选。

选题的目的

随着互连网应用的发展,人们越来越多地把学习过程放到网络上,而不再局限于传统的教室学习,因此基于 web 的远程学习得到了迅速推广。但是现有的远程教育系统模式单一,学生只能被动地接受完全相同的学习内容,不能主动地根据自己的需要选择,控制自己的学习过程。本课题采用基于个性化的远程学习模型,运用面向对象的理论和技术,以及基于 web 的数据库技术,来实现个性化的远程教育系统,以此来达到改善现有的远程教育系统,并推广个性化的远程学习模型能成为远程教育应用的新模式。

课题所属研究领域

《基于 web 的远程教育系统的研究》课题属于远程教育系统的研究领域。

本文的主要工作

1. 总体结构设计

本系统设计采用B/S结构的Web方案,从逻辑上分成三层,即客户端、事务逻辑层和数据存储层,即采用Browser/Web/Database (BWD)的体系结构。

2. 建立基于面向对象的系统分析模型

本系统采用面向对象的分析方法建立系统模型。经过深入分析远程教育系统,运用统一建模语言UML和Rational Rose建模工具得到OOA的基本模型——类图。

3. 设计模式的选择

采用模型视图控制器模式(MVC模式)作为本课题的设计模式。MVC模式可以将功能划分为相互关联的三个组件——模型、视图和控制器。本课题在实现过程中,是通过使用Struts框架来具体实现MVC模式。

4. 功能设计

具体包括以下主要功能:

一个远程教育系统通常应该包括以下几个主要的子系统:入学系统,学习系

统, 作业 / 考试系统, 毕业系统, 教师管理系统, 课件管理系统和教务管理系统。其中学习系统中又包含两个子系统, 学习指导系统和答疑系统。在此系统中, 重点分析和实现学习系统和作业/考试系统。

本课题的研究方法与技术路线

基于 Web 的远程教育系统是一个规模庞大的系统, 涉及范围大, 且结构复杂, 为了保证系统产品质量和开发效率, 采用面向对象的软件工程方法进行开发, 它包括面向对象的分析 (object-oriented Analysis, 简称 OOA)、面向对象的设计 (object-oriented Design, 简称 OOD)、面向对象的编程 (object-oriented Program, 简称 OOP)、面向对象的测试 (object-oriented Test, 简称 OOT) 和面向对象的软件维护 (object-oriented Software Maintain, 简称 OOSM)^[5], 每个阶段的基本任务是:

1. 面向对象的分析: 针对问题域运用 OO 方法, 建立一个反映问题域的 OOA 模型。
2. 面向对象的设计: 针对系统的一个具体的实现运用 OO 方法。其中包括两方面的工作, 一是把 OOA 模型直接搬到 OOD(不经过转换, 仅作某些必要的修改和调整), 作为 OOD 的一部分; 另外是针对具体实现中的人机界面、数据存储, 任务管理等因素补充一些与实现有关的部分。这些部分与 OOA 采用相同的表示法和模型结构。
3. 面向对象的编程: 又称面向对象的实现, 它的工作就是用同一种面向对象的编程语言把 OOD 模型中的每个成分书写出来。
4. 面向对象的测试: 对于用 OO 技术开发的软件, 在测试过程中继续运用 OO 技术, 进行以对象概念为中心的软件测试。
5. 面向对象的软件维护: 通过各种必要的维护活动使系统持久地满足用户的需要。

本课题是采用 web 技术和数据库技术, 并运用软件工程, 面向对象技术为原则, 使用 Java, SQL Server2000 的技术加以实现的。

课题的理论意义

随着知识经济时代的降临和信息社会的飞速发展, 作为工业化产物的线性教育系统正面临越来越严峻的挑战, 同时也面临跳跃式发展的重大机遇。21 世纪

伊始,以多媒体和计算机互联网为代表的当代信息技术正以惊人的速度改变着人们的生存方式和学习方式。许多有远见的战略家都认为,近5-10年内,从基础教育到高等教育,都将发生一场历史性的变革,这场变革将使教育模式迅速步入信息时代。

随着计算机技术的飞速发展,信息高速公路的建立,Internet的普及,基于计算机网络的应用模式渐渐成为各行各业开发人员的研究热点,基于Internet的远程教育也应运而生。Internet的广泛使用为学习,教学和科研提供了取自不尽,用之不竭的巨大信息资源。这为教育带来了机遇同时也带来了冲击。比如,通过Internet可以实现教师,媒体和学生的自主交流,使学习向着个性化的方向发展;而同时基于Internet的教育方式又在改变着传统的教育方式和教学过程。为了满足社会教育,在职教育和终身教育的需求,适应教育现代化的发展,基于Internet的远程教育——现代远程教育——正在悄悄却又坚实地崛起,对诸如广播电视和函授的传统的远程教育方式产生了强烈冲击,为教育注入了新鲜的血液,使现有的教育资源的作用得以更加充分的发挥,也大大提高了教育过程中的实时交互性。现在全球已经开办现代远程教育和继续教育并且具有相当规模的学校有将近200多所。国内外也有相当数量的大学已经开办了远程与继续教育,但现代远程教育的兴起不仅仅是多了一种教育方式和一种所谓的“网上大学”,它的发展可能会改变我们对教育的一些传统概念,其特点是运用计算机网络技术和多媒体数字技术,在数字化环境下进行交互式的学习,突出地表示为教学层次的多样性,教学方式的灵活性与教学管理的高效性。

现代远程教育是全球信息化进程在教育领域的必然结果,以信息高速公路为代表的信息化给教育带来的,不仅是全新的技术和媒体,而且必将全面影响到教育观念,教育体制,教育模式和教育内容^[9]。二十一世纪的教育必将围绕着基本学习能力从新设计和从新组织,使整个社会成为学习社会,是学习和掌握知识成为社会进步的主要的推动力和生活的第一需要。教育的目标会改变,不是为了一纸文凭,而是为了终身受到教育。

在家庭办公推动着传统产业向集成化,无人化方向发展的时候,越来越多的人将用一台连接于网络的电脑或交互电视在家里,或从事信息工作,或经营商务,或对自己的公司进行远距离管理。从这个意义上讲,终身教育必然要通过网络把

触角伸进家庭，到传统教室里上课已无法吸引越来越多的家庭办公族了。终身教育的概念是教育进入知识经济时代的关键，终身教育体制将在一系列变革中逐渐确立其中心地位。强化终身教育迫切需要变革现行的教育模式，迅速创建一体化的学习社会和虚拟学校，以便充分利用电脑网络所能提供的一切先进手段，为社会成员创造更多的受教育机会。

由于目前一次性的正规学校教育模式根本无法给每个人以真正意义上终身教育的可能性，社会对学习的需求呼唤着信息化教育模式尽快发育和成熟。因此，发展现代远程教育，构建终身学习体制是教育发展和社会进步的共同要求。我国现代远程教育的目标和任务是到 2010 年，基本形成多规格，多层次，多形式，多功能，具有中国特色的终身教育体制。

国内外研究动向及进展

随着巨大的社会需要和信息技术革命的深入，人类教育史上的一种革命性的教育模式——现代远程教育应运而生。现代远程教育的孕育和发展虽然还只是近一个时期的事情，但其发展势头之猛烈，应用之广泛，影响是深远，均为教育史上所罕见。

远程教育也称远程学习，过去常常叫做远距离教育，第一代远程教育是函授教育，它至今仍在成人教育中发挥着作用。随着电子信息技术的进步，传播媒体出现了较大的变化，幻灯，录音，电话，电影，电视逐步介入教学领域，印刷品加信件的函授教学方式开始向广播，电视教学方式转化，远程学习增加了声音和视频图像等媒体，远程教育也进入了它的第二代历程^[3]。习惯上把这一代远程教育称为远距离教育。当前发展的方兴未艾的远程教育则属于第三代，它建立在互联网络的基础上，从某种意义上讲，它就是“网络教育”的同义词。在我国教育界，这一代远程教育被称为“现代远程教育”。

现代远程教育是一个发展的概念，通常指的是上述远程教育形态中的第三代。换言之，现代远程教育是计算机信息技术和 Internet 在远程教育领域的新兴应用。纵观远程教育的发展史，我们可以看到，向交互性，多种媒体交流方向发展是远程教育发展的主线。

现代远程教育是基于网络的远程教育，与前两代相比，现代远程教育使用的是双向交互式通讯媒体，是建立在电脑技术，网络技术，多媒体技术，双向电子

通信技术基础上的，与教育传播理论，现代学习理论紧密结合，它以交互性，网络化，实时性，综合性和适应性为基本特征。

现在全球已经开办现代远程教育和继续教育并且具有相当规模的学校有将近 200 多所。它们大致可分为两类，一类是附属于大学的，一类是独立地只进行远程与继续教育的。第一类中有许多国际著名的大学，其中在美国就有哈佛大学，斯坦福大学，马里兰大学，加州大学的各个分校和弗吉尼亚理工学院等；第二类中有英国公开大学，香港公开大学，荷兰公开大学等等。国内也有相当数量的大学已经开办了远程与继续教育，如清华大学，上海交通大学，北京大学，北京邮电学院等。但现代远程教育的兴起不仅仅是多了一种教育方式和一种所谓的网上大学，它的发展可能会改变我们对教育的一些传统概念，它的特点是运用计算机网络技术和多媒体数字技术，在数字化环境下进行交互式的学习，突出地表现为教学层次的多样性，教学方式的灵活性与教学管理的高效性。

而现有的远程教育系统，在交互性方面做的还不是很好。另外，现有的系统虽然自身信息量巨大，但是，浏览的课件、考试的题目与整个学习的流程都是单一固定的，教育方法与模式上就显得非常单一，相比而言，接受教育的对象存在个性差异，个人的学习能力、个人兴趣与习惯、个人学习基础、努力程度，都存在巨大的差异。这造成了两者之间越来越大的矛盾。目前的远程教育忽略了学习本身是一种个性化的过程，没有做到因人而异、因材施教、兼顾到个体的差异。因此，为了从根本上解决这一矛盾，我们将把个性化的概念引入远程教育中。

引入个性化后的远程教育是一个智能化的系统，可以更多且及时地了解到学生的状况、需求、能力差异、学习进度、兴趣爱好，并动态地根据这诸多因素调整学习计划及进度，让用户得到针对其“个性”的教育。

远程教育在中国

远程教育在中国的发展经历了三个阶段：第一个阶段是函授教育，这一方式为我国培养了很多人才；第二个阶段是 70 年代兴起的广播电视教育，我国的这一远程教育方式和中央电视广播大学在世界上享有盛名；90 年代，随着信息和网络技术的发展，我国产生了以信息和网络技术为基础的现代远程教育。

在二十世纪九十年代末，我国出现了现代远程教育。在最近的这几年里我国的远程教育又经历了一个高速发展的过程，而且这一进程前进的速度还在加快，

在这一期间里，许多学科和专业都找到了和摸索出了一套切实可行的技术和手段，借助互联网来进行自己的远程教育，而且已经有许多学校把远程教育商业化了，并已经取得了不小的成绩。

在国外，远程教育发展较早，市场稳步增长。发展以互联网为主要手段的现代远程教育如今已成为一种不可阻挡的发展趋势。2000年，世界上已经有100多个国家开展了现代远程教育，全世界86%左右的大学在网上拥有自己的网站，其中的1/6推出了网络教育课程。由于网络风暴首先席卷欧美，他们的网络教育很自然走在了世界前列，已具有不小的规模。而国内的远程教育刚刚起步，前景日益广阔。以高校远程教育为主的整个远程教育市场目前发展迅速、潜力巨大。但是，由于我国的远程教育还在发展阶段，如果不算电大系统的远程教育，我国真正意义上的利用网络技术开展的远程教育在2000年才刚刚开始起步。这就需要我们努力发展本国的远程教育系统，使其尽快走在世界前列。

本文的组织结构

本文共分七章。

第一章绪论，讲述了题目来源，选题目的，课题所属研究流域，本文的主要工作，给出了研究方法和技术路线，并介绍了远程教育系统的发展。

第二章是基本方法与技术。

第三章是需求分析。

第四章是系统设计。

第五章是系统实现。

第六章为系统测试。

第七章为全文的总结和进一步研究的展望。

第一章 基本方法与技术

本课题主要涉及到的方法与工具有：OOA/OOD 方法，统一建模语言 UML 及建模工具 Rational Rose，web 原理，web 数据库访问技术，数据挖掘技术，MVC 设计模式，Servlet 技术，JSP 技术，JavaBean 技术等，本章针对这些基本原理进行详细阐述，并说明如何将这些知识结合在一起应用到实际项目中。

1.1 面向对象分析与设计方法

OOA/OOD方法是由Coad和Yourdon于1991年提出来的。与传统的分析方法相比，OOA/OOD方法具有如下优势：可以处理更有挑战性的问题域，改善了分析人员与问题领域专家的交流，通过分析，设计和编程增加内部的一致性，显式地表示类和对象间的共性^[4]。

可以建立有弹性的规范，面向对象分析，面向对象设计和面向对象编程的结果可重用，为分析，设计和编程提供一致的基本表示。

1.1.1 面向对象分析

对于一个大系统，最好先将问题域细化为几个主题，然后再开始分析过程。

OOA 的活动如下：

1) 识别出问题域中的类和对象

在这个活动中，分析人员通过对问题域深入地分析和理解，识别出组成系统核心的相关、稳定的类和对象。找出类和对象的第一步是研究问题域，可以通过审视下列选项来发现可能的类和对象：结构、其他系统、设备、被记住的事情或事件、所扮演的角色、操作的程序、地点（物理位置）、有组织的单元。找出可能的类和对象后，对这些候选的类和对象进行筛选，将筛选后的类和对象添加到 OOA 图中。

2) 确定结构

结构可以分为两种：“一般——特殊”结构和“整体——部分”结构。

在找出“一般——特殊”结构和“整体——部分”结构后，就可以识别出多重结构，多重结构是“一般——特殊”结构和“整体——部分”结构的各种组合。识别出多重结构后，将结构添加到 OOA 图中。

3) 确定主题

在这个活动中，将模型分解为更易管理和理解的主题域，从而降低所产生模

型的复杂性。

4) 定义属性

在初步识别出属性后，可以识别出对象间的实例连接，对识别出的属性和实例连接进行检查，然后规定属性，将属性和实例连接添加到OOA图中。

5) 定义服务

在初步识别出服务后，可以识别出消息连接，然后规定服务，并将服务和消息连接添加到OOA图中。

6) 准备文档

OOA部分的最后一步是整理OOA文档。主要文档包括：完整的OOA图，类和对象的规格定义。

根据这些主要活动，在分析阶段建立的模型由5层组成：

- 1) 主题层：将系统分为模块，从而降低模型的复杂性。
- 2) 类和对象层：类和对象层描述了系统中的类和对象。
- 3) 结构层：结构层描述了类之间的继承关系和“整体——部分”结构。
- 4) 属性层：属性层描述了属性以及类和对象之间的实例连接。
- 5) 服务层：服务层描述了方法以及类和对象之间的消息连接。

1.1.2 面向对象设计

OOD的活动如下：设计问题域组件、设计人机交互组件、设计任务管理组件、设计数据管理组件^[7]。这些组件的设计没有时间顺序。OOD阶段扩充了OOA阶段创建的5层，OOD阶段将OOA阶段产生的结果放入组件中，如图所示。

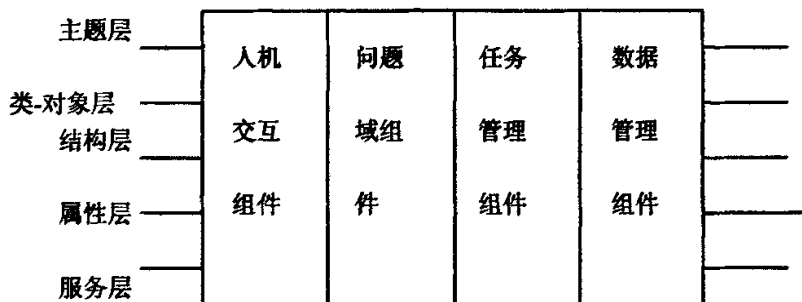


图 2-1 OOA 与 OOD 的连接
Figure 2-1 OOA and OOD connection

1.2 统一建模语言及工具

1.2.1 统一建模语言 UML

使用统一建模语言(Unified Modeling Language, 简称 UML)来建立系统分析模型。它是一种对软件密集型系统的制品进行可视化、详述、构造和文档化的图形语言。UML 给出了一种描绘系统蓝图的标准方法,其中即包括概念性的事物,如业务过程和系统功能,也包括了具体的事物,如用特定的编程语言编写的类、数据库模式和可复用的软件构件。可视化建模语言 UML 将模型中的信息用标准图形元素直观地表现出来,实现模型内部及外部的各种通信。UML 为对象的结构模型和行为模型定义语义,描述基于 UML 的元模型定义,从而为 UML 的所有元素在语义上提供了通用的定义性说明,消除了因人而异的表达方法的影响。

UML包括概念的语义,表示法和说明,提供了静态、动态、系统环境及组织结构的模型^[9]。它可被交互的可视化建模工具所支持,比如本文所用到Rational Rose工具,这些工具提供了代码生成器和报表生成器。UML标准并没有定义一种标准的开发过程,但它适用于迭代式的开发过程。它是为支持大部分现存的面向对象开发过程而设计的。

UML描述了一个系统的静态结构和动态行为。UML将系统描述为一些离散的相互作用的对象并最终为外部用户提供一定的功能的模型结构。静态结构定义了系统中的重要对象的属性和操作以及这些对象之间的相互关系。动态行为定义了对象的时间特性和对象为完成目标而相互进行通信的机制。从不同但相互联系的角度对系统建立的模型可用于不同的目的。

UML还包括可将模型分解成包的结构组件,以便于软件小组将大的系统分解成易于处理的块结构,并理解和控制各个包之间的依赖关系,在复杂的开发环境中管理模型单元。它还包括用于显示系统实现和组织运行的组件。

UML不是一门程序设计语言,但可以使用代码生成器工具将UML模型转换为多种程序设计语言代码,或使用反向生成器工具将程序源代码转换为UML,UML不是一种可用于定理证明的高度形式化的语言,这样的语言有很多种,但它们通用性较差,不易理解和使用。UML是一种通用建模语言。对于一些专门领域,例如用户图形界面(G UI)设计、超大规模集成电路(VLSI)设计、基于规则的人工智能领域,使用专门的语言和工具可能会更适合些。UML是一种离散的建模语言,不

适合对诸如工程和物理学领域中的连续系统建模。它是一个综合的通用建模语言，适合对诸如由计算机软件、固件或数字逻辑构成的离散系统建模。从企业信息系统到基于Web的分布式应用，甚至到严格的实时嵌入式系统都适合于用UML来建模。它是一种富有表达力的语言，可以描述开发所需要的各种视图，然后以此为基础装配系统。

1.2.2 建模工具Rational Rose

目前在面向对象分析、建模、设计工具市场上起着主导作用的 Rational Rose，其主要特点是：

1) 支持 UML 对象建模符号标准

Rational Rose 提供完整的 UML 支援。由系统分析、设计至程序设计，均可用 UML 来表示。UML 让系统开发的所有参与者，在此世界通用的建模语言的基础上沟通。

2) 支持构件化软件发展

对复杂系统而言，构件式的软件开发已跃升为最有效率的模式。Rational Rose 允许使用者使用构件图，清楚地表达构件及其界面之间的关联。

3) 支持多种语言的软件开发

在大型软件开发中，使用数种程序语言是很常见的。Rational Rose 企业提供了多种语言的支持，让使用者可以发展不同程序语言的构件。Rational Rose 可从模型产生 C++、Java、VB、Delphi、PowerBuilder、SmallTalk 或 Ada 的程序代码。

4) 反复式工程

Rational Rose 具有正向工程、反向工程等功能特性，可以在修改实际程序后将最新的程序代码状况，迅速地反应到设计模型中，以保持对象设计模型与程序源代码的一致性。

5) 强力支持团队开发

无论团队有 10 或 100 个成员，Rational Rose 均能有效地协调他们之间的工作。

6) 扩充性

Rational Rose 提供了比较好的扩充性。使用者可透过 Rational Rose 之

IE automation server 界面 API, 取得在此系统中系统开发者所设计的各种模型的资料, 并加以运用, 透过 Rational Rose 外挂功能 (Add-In), 达到扩充 Rational Rose 功能的目的。

Rational Rose 是系统的图形, 包括所有的 UML 框图、角色、使用案例、对象、类、组件和部署结点。它详细描述系统的内容和工作方法, 开发人员可以用模型作为所建系统的蓝图。Rational Rose 支持开发的模型有: Business Use Case 框图、Use Case 框图、Activity 框图、Sequence 框图、Collaboration 框图、Class 框图、Statechart 框图、Component 框图和 Deployment 框图。

Rational Rose 在软件开发过程的几个阶段很有用。在项目开始阶段, Rose 可以产生业务模型使用案例模型。在细化与构造阶段, Rose 可以开发活动框图, 显示事件流程。序列图和协作图显示要开发的对象, 及其相互间的交互。Rose 开发的类框图显示对象之间的相互关系。组件框图显示系统组件间的相关性, 并可以产生系统的框架代码。在构造阶段, Rose 可以将新开发代码逆向转出工程代码到模型中, 根据现有的系统产生模型 (好处: 模型发生改变时, Rose 可以修改代码, 作出相应改变; 代码发生改变时, Rose 可以自动将这个改变加进模型中), 将开发阶段出现的变化反映在模型中。构造之后进入交接阶段, Rose 主要用于在软件产品完成时更新模型。

Rational Rose 提供完整的 UML 支持, 可从模型产生 C++、JAVA、ADA 等的框架代码, 并能通过反复式工程保证模型与代码的一致, 还能支持团队开发。因此, 采用 Rational Rose 进行 UML 建模, 能够比较完整地实现 UML 建模思想, 准确地表示出图形化模型, 提高系统开发的效率。

1.3 Web 基础

什么是 Web Service? 从表面上看, Web service 就是一个应用程序, 它向外界暴露出一个能够通过 Web 进行调用的 API。这就是说, 你能够用编程的方法通过 Web 来调用这个应用程序。我们把调用这个 Web service 的应用程序叫做客户。

Web service 有更精确的解释: Web services 是建立可互操作的分布式应用程序的新平台。Web service 平台是一套标准, 它定义了应用程序如何在 Web 上实现互操作性。你可以用任何你喜欢的语言, 在任何你喜欢的平台上写 Web

service，只要我们可以通过 Web service 标准对这些服务进行查询和访问。

Web service 平台需要一套协议来实现分布式应用程序的创建。任何平台都有它的数据表示方法和类型系统。要实现互操作性，Web service 平台必须提供一套标准的类型系统，用于沟通不同平台、编程语言和组件模型中的不同类型系统。在传统的分布式系统中，基于界面(interface)的平台提供了一些方法来描述界面、方法和参数(译注:如 COM 和 COBAR 中的 IDL 语言)。同样的,Web service 平台也必须提供一种标准来描述 Web service，让客户可以得到足够的信息来调用这个 Web service。最后，我们还必须有一种方法来对这个 Web service 进行远程调用。这种方法实际是一种远程过程调用协议(RPC)。为了达到互操作性，这种 RPC 协议还必须与平台和编程语言无关。

总的来说，Web服务(Web Services)是由企业发布的完成其特别商务需求的在线应用服务，其他公司或应用软件能够通过Internet来访问并使用这项应用服务，它是基于网络的、分布式的模块化组件，执行特定的任务，遵守具体的技术规范，这些规范使得Web Service能与其他兼容的组件进行互操作。实质上讲Web服务是自包含的、模块化的应用程序，它可以在网络中被描述、发布、查找以及调用。

Web服务在Internet上不同服务器接口之间应用程序的相互调用特性，省去了开发人员为编写每个应用程序进行的重复而烦琐的工作，以便集中精力挖掘软件独特的商业价值。Web服务的关键是WEB服务体系结构，它是由平台搭建商提供基于Internet的应用解决方案，这些方案可以用标准的格式通过Internet进行调用。而那些在外部通过Internet调用这些解决方案的特定进程的企业则构成了WEB服务的客户群。

Internet向人们提供电子邮件、文件传输、远程登录、电子公告、游戏娱乐、WWW等服务功能，其中WWW信息服务称为Web服务，是目前Internet上最受欢迎、发展最快的技术。

在Web环境中，信息是以Web页的形式出现的，超文本和超媒体是Web页上信息的基本组织方式，通过超文本或超媒体的超级链接，可以实现Web页面间的跳转，从而使整个Internet上的Web环境成为一个整体。

1.4 开发模式

现在管理系统的开发模式主要有两类:客户机(Client)/服务器(Server)模式,浏览器(Browser)/服务器(Server)模式。

1.4.1 客户机(Client)/服务器(Server)模式

C/S结构也称为客户机(Client)/服务器(Server)结构模式。它根据网络中各个计算机的特点进行分工,克服了集中式网络中各个计算机分工不明确的缺点。

1.4.1.1 客户机/服务器模式概念

客户机/服务器(Client/Server)这一概念最早用于描述软件的体系结构,表示两个程序间的关系,即一个是提出请求的应用程序,另一个是服务程序,它们之间的交互不需要任何通信设施。此时,它们可能就是在同一台机器上运行的调用和被调程序,而且请求服务交互可以是不同步的。

从概念上讲,C/S(Client/Server)模式可定义为一种特殊的协作式处理模式,整个应用程序被分布在客户机和服务器上。客户机和服务器两者都参与一个应用程序的处理,其中软件成分相互协作以完成特定的应用程序功能,而硬件资源则提供软件成分相互协作的设施。

在Client/Server体系结构中,Client向Server发送请求,Server响应Client发出的请求并返回Client所需要的结果。它将I/O操作快的计算机用作服务器,其它计算机作为客户机。同时应用程序也进行分工,Server上是数据库应用SQL程序,Client上主要是I/O界面及处理、分析程序,提高了计算机的运行效率,减少了网络上的数据传输,系统可维护性及坚固性加强。客户机/服务器模式的定义强调协作处理。客户机/服务器模式的协作处理是一种特殊的分布式处理。它把一个计算机应用系统分成三个基本组成部分:

服务器:专门从事提供某项目服务功能的计算机系统。

客户机:面向最终用户,完成各自业务处理及用户界面的计算机系统。

中间件:连接服务器和客户机的部分。

在同一个网络中,客户机根据用户的业务处理要求,提供人机交互界面,提出数据服务请求,通过网络将请求传送给服务器,服务器端按请求组织数据,并通过网络把结果传送给提出请求的客户机;客户机处理或显示结果数据。也就是说,在客户机/服务器模式中,网络上的计算机系统是以协作方式来完成某项功能的。

1.4.1.2 客户机/服务器环境下应用成分的分布

Client/Server将软件分成三部分:Clients, Servers和Interfaces(接口), 能够将不同的数据源、应用及服务集成在一起。在 Client/Server体系结构下, Server控制所有的数据, 具有灵活性和适应性, 能够满足不断增长的需求, 允许任何数据源以标准接口响应Client请求。而Client应用十分灵活, 任何能够通过网络与Server通信的计算机均可运行Client应用。通常, 一个典型的应用程序可分解为四个组成部分:

- 1) 界面表示逻辑:这是与用户交互的应用代码。
- 2) 业务处理逻辑:这是使用输入数据来完成业务处理和规则的应用代码。
- 3) 数据处理逻辑:这是应用程序中负责数据逻辑如完整性判别等的代码。

4) 数据库管理:应用程序由DBMS来完成实际数据存取处理的程序。虽然DBMS不属于应用程序本身, 但它是分布处理的基本组成部分。在 C/S环境下, 通常把界面表示逻辑和业务处理逻辑驻留在客户机上, 而把数据处理逻辑和DBMS功能驻留在服务器上。

一个数据库应用程序, 应该在服务器端完成的是数据的管理、组织和处理, 即提供数据服务; 在网络上传输的应该是客户机的请求和服务器返回的结果或信息; 在客户端完成的是业务处理的表示和用户的交互的界面。

1.4.2 浏览器(Browser)/服务器(Server)模式

在企业Intranet中, 运行在客户端的应用软件将移植到服务器端。客户端将不再需要应用程序, 它们完全集中在服务器端, 这将永久地简化实际应用。这意味着用户完全可以通过浏览器来执行应用程序, 同时, 对诸如并行处理技术的支持也会接踵而至, 而它的数据库容量会变得越来越庞大。这些数据将统一集中在少数的大型服务器上, 而非大量的小型服务器上。如果将程序集中到服务器端, 则情形将大为不同, 可以使用低成本的网络, 通过浏览器传递网上众多的数据。它不再需要应用程序来完成客户端与服务器端的通信, 应用将以广域网而不是局域网为基础。而基于互联网应用系统都采用Browser/Server(浏览器/服务器模式)结构, 在客户端中除Web浏览器外, 要安装的软件便是一些插件和控件。这些插件被设计于服务器端, 并通过网络下载到客户端, 而控件则是在第一次浏览时下载并注册的。这是一种最小化的客户端安装及瘦客户机模式, 是目前客户/服务

器体系结构的发展方向。其分布结构如图2-2所示：

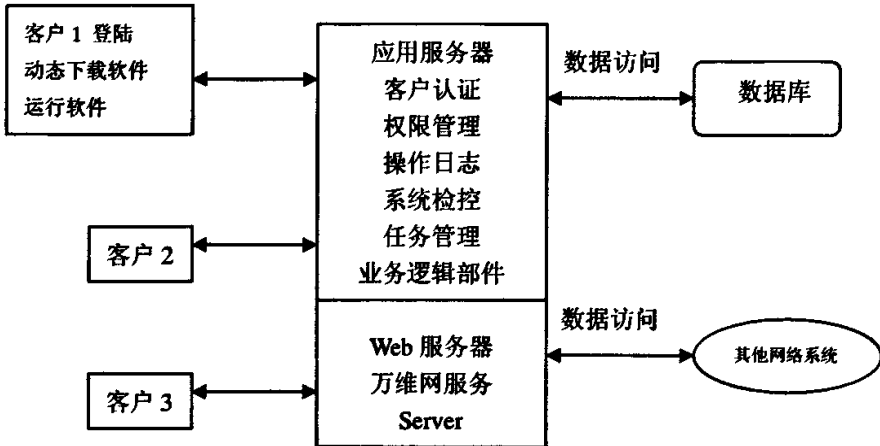


图2-2 B/S结构图

Figure 2-2 B/S structure

B/S环境下的应用系统将传统的C/S体系结构扩展到由Web浏览器、Web服务器、数据库服务器组成的三层B/S结构，B/S结构极大的简化了客户机的工作，客户机上只需安装配置少量的客户端软件，服务器将承担更多的工作，对数据库的访问和应用程序将在服务器上完成。三层体系结构被分割成三个相对独立的单元：表示层(Presentation)、功能层(Business Logic)、数据层(Data Source)^[12]。

1) 表示层：Web浏览器

在表示层中包含系统的显示逻辑，位于客户端，主要负责人机交互，包括一些与数据和应用关系的图形和界面运算。它的任务是由Web浏览器向网络上的某一Web服务器提出服务请求，Web服务器对用户身份进行验证后用http协议把所需的主页传送给客户端，客户机接受传来的主页文件，并把它显示在Web浏览上。

2) 功能层：应用服务器

具有应用程序扩展的Web服务器和应用服务器，在功能层中包含系统的事务处理逻辑。Web服务器主要负责对客户端应用程序的集中管理。Web服务的任务是接受用户的请求，首先需要执行相应的扩展应用程序与数据库进行连接，通过SQL等方式向数据库服务器提出处理申请，而后台数据库服务器将数据处理的结果提

交给Web服务器,再由Web服务器传送回客户端。应用服务器主要负责应用逻辑的集中管理,即事务处理,应用服务器由可以根据其处理的具体业务不同而分为多个。

3) 数据层:数据库服务器

数据库服务器主要负责数据的存储和组织、数据库的分布式管理、数据库的备份和同步等。在数据库中包含系统的数据处理逻辑,位于数据库服务器端,它的任务是接受Web服务器对数据库操作的请求,实现对数据库查询、修改、更新等功能,把这行结果提交给Web服务器。在三层的B/S体系结构中,数据及应用可通过不同平台、不同网络存取,与平台无关,伸缩性大,而且,数据计算与处理集中在中间的功能层,而这一层的服务器性能易于提升,因而这种三层的B/S模式在企业级Intranet的构筑和改建中应用普遍而且满足了用户的需要。

随着网络应用不断发展,瘦客户机/服务器模式成为Intranet中的主流方式。在B/S结构下,客户机的规模被压缩,成为一个简单的浏览器;应用逻辑被集中起来,置于一个或多个服务器(应用服务器)上。

1.5 Web数据库访问技术概述

Web数据库系统的主要目的是要实现Web与数据库的连接以产生基于数据库的动态页面,这要通过Web访问数据库来实现,相关的技术我们称之为Web数据库访问技术(也称为动态页面技术)^[14]。目前这种技术种类繁多,大致可分为三类:

第一种是在Web服务器端提供中间件来连接Web服务器和数据库服务器,常用的中间件技术有通用网关接口(CGI)、应用服务器编程接口(WebAPI)、ASP、PHP、JavaServlet、JSP等。这种技术的特点是采用Web服务器作为通信中介,由Web服务器启动应用程序,并由应用程序完成数据库的访问,结果信息在经Web服务器返回客户端浏览器。通过中间件应用程序访问Web数据库的一般结构如图1.3所示。

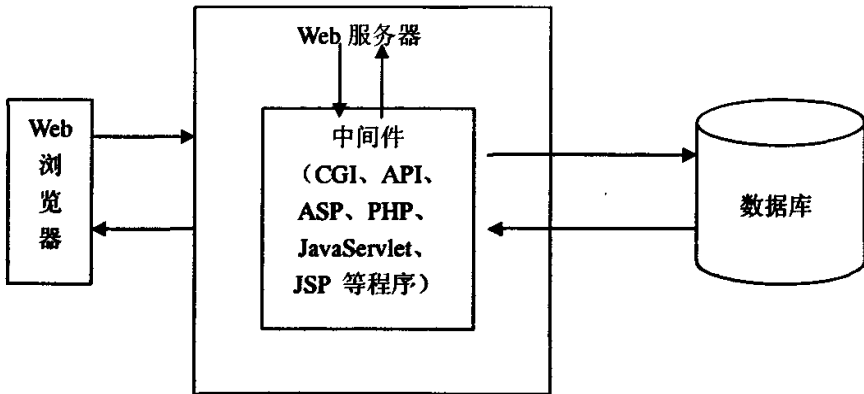


图1.3 基于中间件的Web数据库系统结构

Fig. 1.3 Web DBS architecture based on middleware

第二种是把应用程序下载到客户端运行，在客户端直接访问数据库服务器，例如：Java Applet等。通过客户端应用程序访问Web数据库的一般结构如图1.4所示。

第三种方式可看成是上述两种方法的组合，在服务器端提供中间件，同时将应用程序的一部分下载到客户端并在客户端通过WebServer及中间件访问数据库。

目前采用最多的解决方案是第一种方案，具体技术有：CGI、WebAPI、JSP、ASP、PHP等，其中JSP、ASP、PHP是功能相似的三种主流技术。这些方案各有其优缺点，用户可以根据具体需要，选择相应的技术。在这里我们选择JSP技术来实现第一种方案。

下面介绍有关Java技术并分析比较它们的优缺点。

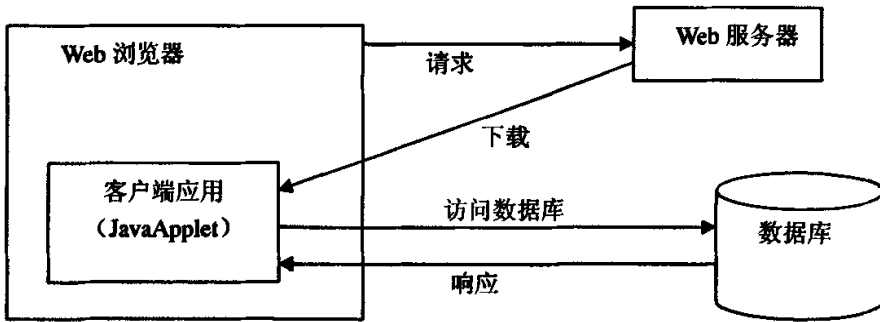


图1.3 基于客户端的Web数据库系统结构

Fig. 1.3 Web DBS architecture based on Client

1.6 JDBC与JavaBean

1.6.1 JDBC(Java Database connectivity)

JDBC是Javasoft公司设计的Java语言的数据库访问API^[16]，它由一组用Java语言编写的类和接口组成，通过它的一个程序，开发者可以在Java程序中建立与Database的链接，执行SQL语句，处理SQL语句返回的结果。

最初的Java语言并没有数据库访问的能力，JDBC是第一个支持Java语言的标准的数据库API，其目的在于使Java程序与数据库服务器的连接更加方便。JDBC的原理机制和ODBC相同，它给基于Java语言的应用程序提供了统一的数据库访问接口，只不过JDBC是用纯Java实现的，而ODBC是用纯C实现的。将Java和JDBC结合起来将使程序员只须写一遍程序就可以让它在任何平台上运行。

Java具有坚固、安全、易于使用、易于理解和可以从网络上下载等特性，JDBC支持三层模式，把对数据库的访问转交给专用服务器来完成，命令将被发送到中间件，中间件将SQL语句发送到数据库，然后数据库处理SQL语句并将结果返回中间件，最后由中间件返回给Web浏览器客户，即JDBC支持数据库系统的B/S结构。下面的Java Servlet和JSP技术都支持这种体系结构。

1.6.2 JavaBean的概念

JavaBeans是描述Java的软件组件模型，有些类似微软的COM组件概念。JavaBeans组件简单说就是Java类，这些类遵循一个接口格式，以便于使方法命名、底层行为以及继承或实现的行为能够把类看做标准的JavaBeans组件的方式进行构造。在Java模型中，通过JavaBeans可以无限扩充Java程序的功能，通过

JavaBeans的组件可以快速生成新的应用程序。

JavaBeans具有以下特性：

- 1) 可以实现代码的重复利用。
- 2) 易维护性、易使用性、易编写性。
- 3) 可以在支持Java的任何平台上工作，而不需要重新编译。
- 4) 可以在内部、网内或者网络之间进行传输。
- 5) 可以以其他部件的模式进行工作。

JavaBeans传统的应用在于可视化领域，如AWT(窗口抽象工具集)下的应用。现在，JavaBeans更多的应用在于不可可视化领域，它在服务器端应用方面表现出了越来越强的生命力。不可可视化的JavaBeans在JSP程序中常用来封装事务逻辑、数据库操作等，可以很好地实现业务逻辑与前台程序的分离，使得系统具有更好的健壮性和灵活性。

1.7 Servlet知识

1.7.1 Servlet的定义

Servlet是用Java语言编写的运行在服务器端的Java小程序，它在初始化时装入Web服务器的存储空间，并成为服务器的一个组成部分。Servlet通过建立在Web上提供请求和响应服务的机制来扩展服务器的功能，当客户端发送请求给Web服务器，服务器就将请求信息传递给一个Servlet，该Servlet通过JDBC访问数据库，构造响应结果信息^[16]。经由服务器返回给客户端，从而实现B/S体系结构Web数据库系统。

Servlet使用JavaServletAPI及相关类(Java类和软件包)编程，ServletAPI能融合在不同的Web服务器中。当启动Web服务器或Web浏览器第一次请求服务时，系统自动装入Servlet。装入后，Servlet保持运行状态直到其他Web浏览器发出请求。

1.7.2 Servlet的功能

Servlet在启动Java的Web服务器上运行并扩展了该服务器的功能，使用户浏览器访问到的不再是一成不变的静态内容，而是基于数据库的、实时交互的动态页面。Servlet功能涉及范围很广。包括：

- 1) 建立并返回根据客户请求产生的包含动态内容的整个页面；

- 2) 生成部分HTML页面，该部分页面可以插入已经存在的页面中；
- 3) 同服务器上的数据库或其他Java应用程序等资源进行通信；
- 4) 为多个客户机处理连接，接收多个客户机的输入，并将结果广播到多个客户机上。例如，Servlet可以是多参与者的游戏服务器；
- 5) 建立从服务器到浏览器Applet的连接，维持连接的通畅，允许在连接上进行多个数据传输；
- 6) 提供对服务器标准例程的定制处理，如一个Servlet能修改用户的身份鉴别方式。

1.7.3 Servlet的工作原理

Servlet通过JDBC访问数据库，采用B/S体系结构。

基于Java Servlet的数据库访问过程分为以下几步：

- 1) Servlet的调用
- 2) 获得客户端请求信息
- 3) 访问数据库
- 4) 返回处理结果

用户通过Web浏览器输入操作命令和数据，向Web服务器发出请求，调用Servlet。Web服务器收到请求后，加载相应Servlet。同时把用户请求信息传给该Servlet。Servlet分析用户请求信息，生成对应的SQL语句并执行，通过JDBC访问数据库服务器中的数据库，获得响应的结果集，生成HTML页面返回到客户端。

用户可以通过URL调用Servlet，<SERVLET>标签中指定Servlet，或者在JSP文件中调用Servlet，但最常见的是在<FORM>中通过GET()或POST()方式调用Servlet。当一个客户端的请求到达服务器，服务器生成一个请求对象和一个响应对象与该请求相对应，然后服务器调用Servlet的service()方法通过请求对象获得请求信息，处理请求，并利用响应对象把响应信息传递给客户端；service()方法可以调用其他方法处理请求，如doGET()，doPOST()及自己编写的各种方法等。

基于JDBC/JavaServlet的B/S结构的数据库访问过程如图1.5所示。

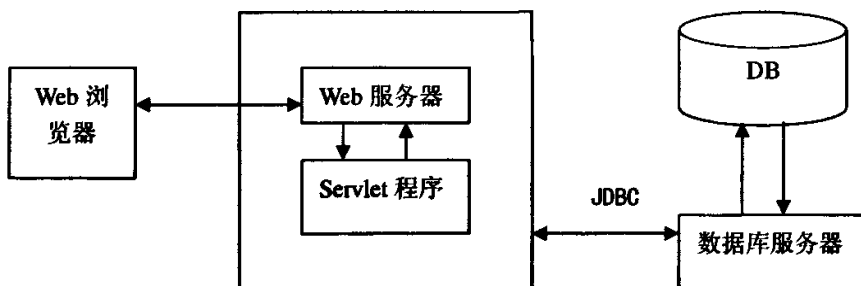


图1.5 Servlet的工作流程

Fig. 1.5 Servlet workflow structure

1.8 JSP技术

1.8.1 什么是JSP

JSP是由Sun Microsystems公司倡导、许多公司参与一起建立的一种动态网页技术标准，是一种很容易学习和使用的在服务器端编译执行的Web数据库系统应用程序编程语言，其脚本语言采用Java，完全继承了Java的所有优点。在传统的网页HTML文件(*.HTM, *.HTML)中加入Java程序片段(Scriptlet)和JSP标记(tag)，就构成JSP网页(*.jsp)^[10]。TAG绝大部分是以“<%”开始，以“%>”结束的。JSP文件的扩展名通常是.jsp，并且一般放在网页存放的地方。

Web 服务器在遇到访问JSP网页的请求时，首先执行其中的程序片段，然后将执行结果以HTML格式返回给客户。程序片段可以操作数据库、重新定向网页以及发送email等等，这就是建立动态网站所需要的功能。所有程序操作都在服务器端执行，网络上传送给客户端的仅是得到的结果，对客户浏览器的要求最低，可以实现无Plugin，无ActiveX，无Java Applet，甚至无Frame。

1.8.2 JSP的工作原理

JSP网页吸引人的地方之一就是能结合JavaBean技术来扩充网页中程序的功能。JavaBean是一种Java类(class)，通过封装属性和方法成为具有某种功能或者处理某个业务的对象。

使用 JSP 有两种技术方案：一种是JSP+JavaBean+JDBC，该技术方案通过JSP直接实现对客户端的应答。当客户的请求提交到JSP时，通过JavaBean将运行的结果再返回给客户端。

基于JSP和JavaBean和B/S结构的数据库访问过程如图1.6所示。

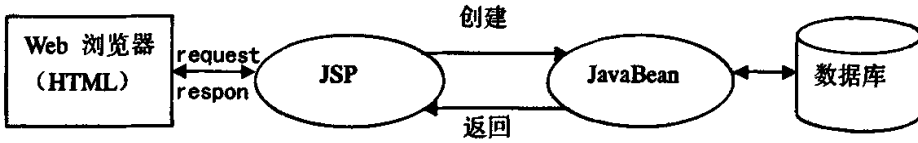


图1.6 JSP的工作流程I

Fig. 1.6 the work flow of jsp I

尽管这种技术方案十分适合简单应用的需要，它却不能满足复杂的大型应用程序的实现。不加选择地随意运用这种技术方案，会导致JSP页内被嵌入大量的脚本片段或Java代码，特别是当需要处理的请求量很大时，情况更为严重。另外一种JSP+Javaservlet+JavaBean+JDBC。它吸取了JSP和Javaservlet技术各自的突出优点，用JSP实现表示层，让Servlet实现深层次的远程教育系统的设计与实现处理任务，即实现应用层。在这里，Servlet充当控制者的角色，负责系统对请求的处理，创建JSP程序需要使用的bean和对象，同时根据用户的动作决定把哪个JSP页传给请求者。

基于JSP，Javaservlet和JavaBean的B/S结构的数据库访问过程如图1.7所示。

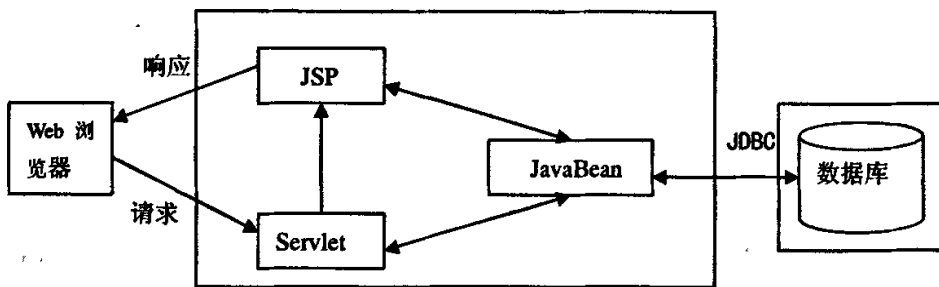


图 1.7 JSP的工作流程II

Fig. 1.7 the work flow of jsp II

这种技术方案的优点是：在JSP页内没有处理逻辑，它仅负责检索原先由servlet创建的对象或beans，从servlet中提取动态内容插入静态模板；它清晰

地分离了表达和内容,明确了角色的定义以及应用程序开发者与网页设计者的分工。

1.8.3 JSP技术优点

JSP可以通过SUN的JavaBean和EnterpriseJavaBeans (EJB) 获得规模支持,通过EJB/CORBA以及其它厂商的应用服务器获得结构支持。下面是它具体的优点。

1) 将内容的生成和显示进行分离

使用JSP技术, Web页面开发人员可以使用HTML或者XML标识来设计和格式化最终页面。使用JSP标识或者小脚本来生成页面上的动态内容。生成内容的逻辑被封装在标识和JavaBean组件中,并且捆绑在小脚本中,所有的脚本在服务器端运行。如果核心逻辑被封装在标识和JavaBean中,那么其他人,如Web管理人员和页面设计者,能够编辑和修改JSP页面,而不影响动态内容的生成。

在服务器端JSP引擎解释JSP标识和小脚本,生成所请求的动态内容(例如,通过访问JavaBean组件,使用JDBC技术访问数据库,或者包含文件),并且将结果以HTML(或者XML)页面的形式发送回浏览器,这有助于作者保护自己的代码,而又保证任何基于HTML的Web浏览器的完全可用性。

2) 生成可重用的组件

绝大多数JSP页面依赖于可重用的,跨平台的组件(JavaBean或者EJB组件)来执行应用程序所要求的更为复杂的处理。开发人员能够共享和交换执行普通操作的组件,或者使得这些组件为更多的使用者或者客户团体所使用。基于组件的方法加速了总体开发过程,并且使得各种组织在他们现有的技能和优化结果的开发努力中得到平衡。

3) 用标识简化页面开发

Web页面开发人员不会都是熟悉脚本语言的编程人员。JSP技术封装了许多功能。这些功能是在易用的、与JSP相关的XML标识中进行动态内容生成所需要的。标准的JSP标识能够访问和实例化JavaBean组件,设置或者检索组件属性,下载Applet,以及执行用其他方法更难于编码和耗时的功能。

1.9 数据挖掘技术

在基于web的远程教育系统中,在个性化数字服务模型的实现中,应用到对于用户综合行为的数据挖掘,具体是针对用户的数据挖掘,用户在系统上几乎所

有可能被记录下来的行为,都要尽可能的挖掘,分析后的结果作为为用户提供服务的依据。下面介绍一些数据挖掘方面的知识。

1.9.1 数据挖掘技术的基本概念

随着计算机技术的发展,各行各业都开始采用计算机及相应的信息技术进行管理和运营,这使得企业生成、收集、存贮和处理数据的能力大大提高,数据量与日俱增。企业数据实际上是企业的经验积累,当其积累到一定程度时,必然会反映出规律性的东西;对企业来说,堆积如山的数据无异于一个巨大的宝库。在这样的背景下,人们迫切需要新一代的计算技术和工具来开采数据库中蕴藏的宝藏,使其成为有用的知识,指导企业的技术决策和经营决策,使企业在竞争中立于不败之地。另一方面,近十余年来,计算机和信息技术也有了长足的进展,产生了许多新概念和新技术,如更高性能的计算机和操作系统、因特网(Internet)、数据仓库(datawarehouse)、神经网络等等。在市场需求和技术基础这两个因素都具备的环境下,数据挖掘技术或称 KDD(Knowledge Discovery in Databases; 数据库知识发现)的概念和技术就应运而生了。

数据挖掘(Data Mining)旨在从大量的、不完全的、有噪声的、模糊的、随机的数据中,提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识^[22]。还有很多和这一术语相近似的术语,如从数据库中发现知识(KDD)、数据分析、数据融合(Data Fusion)以及决策支持等。

1.9.2 数据挖掘的基本任务

数据挖掘的任务主要是关联分析、聚类分析、分类、预测、时序模式和偏差分析等。

1) 关联分析(association analysis)

关联规则挖掘由 Rakesh Apwal 等人首先提出。两个或两个以上变量的取值之间存在的规律性称为关联。数据关联是数据库中存在的—类重要的、可被发现的知识。关联分为简单关联、时序关联和因果关联。关联分析的目的是找出数据库中隐藏的关联网。一般用支持度和可信度两个阈值来度量关联规则的相关性,还不断引入兴趣度、相关性等参数,使得所挖掘的规则更符合需求。

2) 聚类分析(clustering)

聚类是把数据按照相似性归纳成若干类别，同一类中的数据彼此相似，不同类中的数据相异。聚类分析可以建立宏观的概念，发现数据的分布模式，以及可能的数据属性之间的相互关系。

3) 分类(classification)

分类就是找出一个类别的概念描述，它代表了这类数据的整体信息，即该类的内涵描述，并用这种描述来构造模型，一般用规则或决策树模式表示。分类是利用训练数据集通过一定的算法而求得分类规则。分类可被用于规则描述和预测。

4) 预测(predication)

预测是利用历史数据找出变化规律，建立模型，并由此模型对未来数据的种类及特征进行预测。预测关心的是精度和不确定性，通常用预测方差来度量。

5) 时序模式(time-series pattern)

时序模式是指通过时间序列搜索出的重复发生概率较高的模式。与回归一样，它也是用已知的数据预测未来的值，但这些数据的区别是变量所处时间的不同。

6) 偏差分析(deviation)

在偏差中包括很多有用的知识，数据库中的数据存在很多异常情况，发现数据库中数据存在的异常情况是非常重要的。偏差检验的基本方法就是寻找观察结果与参照之间的差别。

1.9.3 数据挖掘常用的基本技术

1) 统计学

统计学虽然是一门“古老的”学科，但它依然是最基本的数据挖掘技术，特别是多元统计分析，如判别分析、主成分分析、因子分析、相关分析、多元回归分析等。

2) 聚类分析

聚类分析主要是根据事物的特征对其进行聚类或分类，即所谓物以类聚，以期从中发现规律和典型模式。这类技术是数据挖掘的最重要的技术之一。除传统的基于多元统计分析的聚类方法外，近些年来模糊聚类和神经网络聚类方法也有了长足的发展。

3) 决策树分类技术

决策树分类是根据不同的重要特征,以树型结构表示分类或决策集合,从而产生规则和发现规律。

4) 人工神经网络

人工神经网络是一个迅速发展的前沿研究领域,对计算机科学、人工智能、认知科学以及信息技术等产生了重要而深远的影响,而它在数据挖掘中也扮演着非常重要的角色。人工神经网络可通过示例学习,形成描述复杂非线性系统的非线性函数,这实际上是得到了客观规律的定量描述,有了这个基础,预测的难题就会迎刃而解。目前在数据挖掘中,最常使用的两种神经网络是 BP 网络和 RBF 网络。不过,由于人工神经网络还是一个新兴学科,一些重要的理论问题尚未解决。

5) 规则归纳

规则归纳相对来讲是数据挖掘特有的技术。它指的是在大型数据库或数据仓库中搜索和挖掘以往不知道的规则和规律,这大致包括以下几种形式: IF ... THEN ...

6) 可视化技术

可视化技术是数据挖掘不可忽视的辅助技术。数据挖掘通常会涉及较复杂的数学方法和信息技术,为了方便用户理解和使用这类技术,必须借助图形、图像、动画等手段形象地指导操作、引导挖掘和表达结果等,否则很难推广普及数据挖掘技术。

1.9.4 数据挖掘技术实施的步骤

数据挖掘的过程可以分为 6 个步骤:

- 1) 理解业务:从商业的角度理解项目目标 and 需求,将其转换成一种数据挖掘的问题定义,设计出达到目标的一个初步计划。
- 2) 理解数据:收集初步的数据,进行各种熟悉数据的活动。包括数据描述,数据探索和数据质量验证等。
- 3) 准备数据:将最初的原始数据构造成最终适合建模工具处理的数据集。包括表、记录和属性的选择,数据转换和数据清理等。
- 4) 建模:选择和应用各种建模技术,并对其参数进行优化。

5) 模型评估: 对模型进行较为彻底的评价, 并检查构建模型的每个步骤, 确认其是否真正实现了预定的商业目的。

6) 模型部署: 创建完模型并不意味着项目的结束, 即使模型的目的是为了增进对数据的了解, 所获得的知识也要用一种用户可以使用的方式来组织和表示。通常要将活动模型应用到决策制订的过程中去。该阶段可以简单到只生成一份报告, 也可以复杂到在企业内实施一个可重复的数据挖掘过程。

1.10 MVC 设计模式

1.10.1 MVC 模式简介

对于任何一个应用程序, 都需要有程序流程的控制, 将这些控制进行抽象就可以总结出某类应用常用的设计模型。传统的面向对象模型编程中, MVC模型是指模型—视图—控制器模型, 其结构如图所示。

在 MVC 模型中, 一个应用程序分成三个独立的组成部分:

模型: 即程序将要实现的业务逻辑。模型是整个MVC模式的核心, 采用面向对象的方法, 将应用领域中的对象抽象为应用程序对象, 在这些抽象的对象中封装了对象的属性和这些对象所隐含的逻辑。

视图: 视图是应用程序的用户界面部分。视图负责应用程序与人之间的接口: 一方面, 它为用户提供了输入手段, 并触发业务逻辑的运行; 另一方面, 它又将业务逻辑运行的结果以某种形式显示给用户。

控制器: 控制器用于管理模型和视图之间的数据交互以及应用程序的流程。一方面, 它解释来自视图的输入, 并将其解释成系统能够理解的对象, 同时它也识别用户动作, 并将其解释为对模型的特定方法的调用; 另一方面, 它也处理来自模型的事件和模型逻辑执行的结果, 调用合适的视图为用户提供反馈。

MVC模型把应用程序的用户界面设计、流程控制和业务逻辑进行了分离。基于该模型进行应用程序的开发。各个组成部分之间的接口一旦制定, 开发人员可以根据各组成部分的主要功能选用适当的技术对该模块进行有针对性的并行开发, 且无须熟悉实现其他模块的具体技术细节, 最后根据所制定的接口进行系统集成即可。MVC模型适合面向对象的分析、设计和编程。

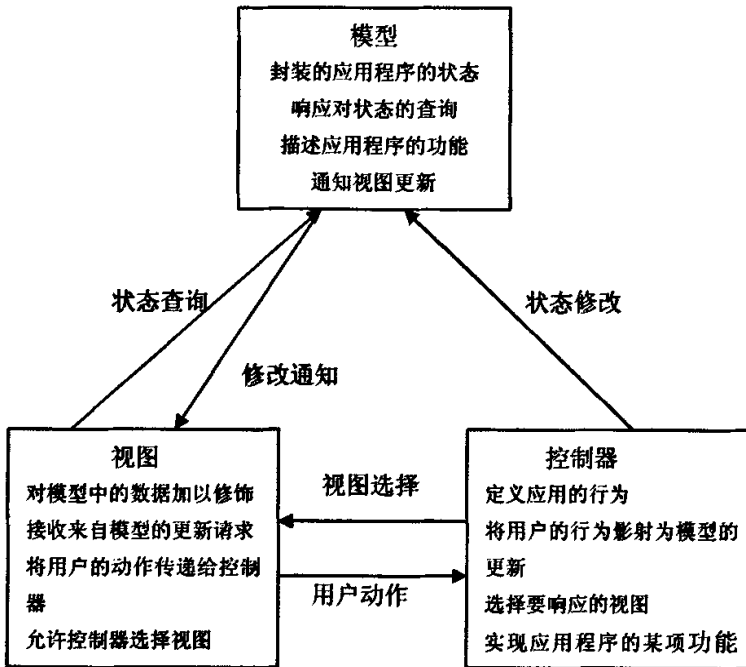


图 1.8 MVC 模型
Fig1.8 MVC model

1.10.2 Struts 框架简介

Struts 是一个开放源代码的 web 应用框架,基于 MVC 设计模式,使用 Servlet 和 JSP API 建立,可以用来构建复杂的 web 应用。它允许我们分解一个应用程序的商业逻辑、控制逻辑和表现逻辑的代码,使它的重用性和维护性更好。

第二章 需求分析

在上述方法和技术的基础上，我们可以对远程教育系统进行需求分析，系统设计和实现，来建立基于 web 的远程教育系统。需求分析阶段的任务不是具体地解决问题，而是准确地确定“为了解决这个问题，目标系统必须作什么”，主要是确定目标系统必须具备哪些功能。

2.1 基于 OO 的需求概述

OOA 强调直接针对问题域中客观存在的各项事物设立 OOA 模型中的对象。用对象的属性和服务分别描述事物的静态特征和动态行为。问题域有哪些值得考虑的事物，OOA 模型中就有哪些对象，而且对象及其服务的命名都强调与客观事物一致。另外，OOA 模型也保留了问题域中事物之间关系的原貌。这包括：把具有相同属性和相同服务的对象归结为类；用一般——特殊结构描述一般类与特殊类之间的关系；用整体——部分结构描述事物间的组成关系；用实例连接和消息连接表示事物之间的静态联系和动态联系^[6]。

总的概括，基于 web 的远程教育系统的研究目标是采用基于个性化的远程学习模型，运用面向对象的理论和技术，以及基于 web 的数据库技术，实现个性化的远程教育，以此来达到改进远程教育系统功能和性能的目的。

2.2 功能需求

根据用户需求，远程教育系统共包括以下七大功能模块：

(1) 入学系统

学生进入远程教育系统，首先要入学注册，选择要学习的课程，并交纳学费，这样教务管理系统中就有了学生的信息，学生凭借学号和密码才可以进行学习，答疑，作业，考试等行为。

(2) 学习系统

自主学习系统是学生利用远程教育系统中的教育资源进行自主学习的一种方式，也是远程教育系统区别于普通学校教育的一个重要方面。现在已经开通的远程教育系统几乎都包含了学生自主学习系统。自主学习系统通常是由教师或专业人员将教学内容精心制作为多媒体形式的电子教材，称为课件，然后放在专门的服务器上，学生随时随地可以通过网络访问这些电子教材，实现自我学习的目

的。

答疑子系统也是构成远程教育系统的一个重要部分。它的主要功能是对学生在学习过程中遇到的问题进行解答,同时对学生的学习效果进行跟踪检查。远程教育的一个重要特点是教师与学生是分离的,学生与学生也可能是分离的。这就要求远程教育系统中必须要有相应的答疑系统来解决学生的疑问,实现学生与学生、学生与教师之间的讨论。在现有的远程教育系统中答疑系统通常是利用Internet上的BBS、E-mail或聊天室来实现的。

(3) 作业 / 考试系统

作业 / 考试子系统的主要功能是负责学生作业的布置、提交和批改以及学习效果的测试。在现有的远程教育系统中,作业与考试系统基本上都是建立在Internet基础上的,比如用E-mail来完成作业的布置、提交,用WWW系统实现网上的在线测试等。

(4) 毕业系统

学生在完成必修课程的学习后,考试和作业都达到了要求,可以结业,进入毕业系统,学生可以得到远程教育系统颁发的电子毕业证书。

(5) 教师管理系统

负责远程教育系统中教师的管理,包括教师档案检索,教师档案新建,变更,教师详细资料等功能。

(6) 课件管理系统

课件是资料库中的内容,属于静态的数据内容,包括课件的添加,删除,更新,教师和管理员有这样的权限。

(7) 教务管理系统

教务管理系统的功能是对学生的注册、缴费、课程设置等进行综合管理。在现有的远程教育系统中很多都不是完全开放的,只有注册的学生才能自由使用教育资源,为此系统必须要有相应的系统来对学生的身份进行认证,并对他的学习活动过程进行跟踪。比如对学生在什么时间学习过什么内容,在线测试的成绩如何等信息进行统计。

因篇幅有限,下面主要详细介绍入学系统,学习系统和作业/考试系统这三大功能模块的需求以及整个系统的基本模型——类图。

2.3 入学系统需求

学生入学包括学生远程登陆系统,注册,缴纳学费和选择专业及课程。学生初次登陆系统,需要先进行注册,填写学生的基本信息,包括姓名,性别,出生年月,入学时间,地址,联系电话,Email等,并设定一个密码,注册成功后,系统给出提示,要求学生进行专业和课程选择,选择完成后,学生根据所选专业和课程缴纳一定的费用,即学费。学费的缴纳通过系统给出一个银行账户,用户汇学费到这个账户就可以了。收到学费后,系统给出提示,学费已缴纳。这样学生的信息在教务管理系统备案,具有了学生这个角色相应的权限:自主学习、答疑、作业、考试等行为。

如果一个学生在选择了一门功课之后,还想学习另外一门功课,这时,此学生还需要进行注册,另外申请一个用户名和密码,重复上面的内容。

2.4 学习系统需求

2.4.1 自主学习子系统需求

自主学习子系统主要指课件的学习,包括学生登陆到系统,通过身份确认后,不同的学生进入到的界面是完全不同的。学生根据用户的身份,在用户库里搜索对应用户的信息,指导学生进入上次的学习界面,查看上次的学习情况,然后学生可以根据自己的需要选择,继续学习后继部分,或者复习上次内容,或者学习上次课程的相关知识,学生也可以随意选择浏览自己感兴趣的知识。

2.4.2 答疑子系统需求分析

这里的答疑子系统指学生在学的过程中遇到疑问时,进行的答疑过程。学生在答疑板上写下疑问,进行答疑,系统根据关键字搜索答疑库,给出答案,如果搜索返回没有找到,这时学生需要把疑问通过Email的形式发送给这门课程的教师,教师通过Email给学生答复,并且教师把题目和答案放到答疑库中。充实答疑库,以供学生解决疑问。另外学生还可以把疑问放在交流板上,以便于同学和教师进行交流解答。

2.5 作业考试系统需求

所谓作业考试系统属于测验部分,学生在每节内容学习结束后,都要完成相应的作业,在每章结束后,完成这一章的考试。作业需要在规定的时间完成,教

师进行批改，考试需要在限定的时间完成上交，教师进行批卷，并公布成绩。关于作业情况和考试情况教师和学生都可以查看。

图 3-1 给出了远程教育系统类图。

2.6 权限管理

为了保证系统安全，特别是对数据库的修改操作，系统采用用户授权来限制用户对数据库的修改操作。权限机制实际上就是代码的使用策略和代码的基本控制权。

本系统有系统管理员，学生，教师，教务人员四种角色。下面是每种角色的权限分析：

系统管理员：负责学生，教师，教务人员的身份注册，根据他们身份的不同，赋予相应的访问权限。

学生：注册学生对课件库的访问，进行课件的浏览学习；通过答疑板对答疑库进行搜索，找到问题的正确答案；通过作业/考试系统完成作业和考试。

教师：教师对资源库——课件库、答疑库、作业库、试题库进行维护。

教务人员：在教务管理系统中对学生信息，教师信息进行维护，有更新资源库的权限。

2.7 基本性能

考虑到系统的功能需求和应用特点，该系统必须具有以下性能：

- 1) 具有安全可靠的网络保障；
- 2) 在平台和技术上的选择，应具有良好的开放性、互操作性和可扩展性；
- 3) 系统应易于管理和维护；
- 4) 具有统一、友好的系统界面，能适应不同用户的应用和要求。

2.8 配置要求

1) 数据库服务器端

操作系统：Microsoft 2000 Server

数据库管理系统：Microsoft SQL Server 2000

配置：TCP/IP

2) Web 服务器

Web 服务器和 JSP/Java Servlet 的容器: Tomcat

Web 服务器和数据库连接技术采用 JDBC

配置 TCP/IP

Microsoft SQL Server 2000 是最常用的大型数据库管理系统之一, 它采用了先进的数据存储技术, 能满足高性能的要求, 同时提供方便的数据备份和恢复机制^[24], 这些性能满足 Internet 环境下高频度和大数据量的信息查询与数据库一致性的要求。

本系统选用 Tomcat 作为 web Server 和 JSP/Java Servlet 的容器(也可以理解成为应用服务器)。Tomcat 是一个带有 JSP 运行环境的 Servlet 引擎, Servlet 引擎是一个管理并负责为用户激活的运行时刻的命令解释程序, 它提供了 Servlet/JSP 引擎的生命周期管理以及运行时刻的支持。

3) 客户端

操作系统: WindowsXP/NT

配置: TCP/IP

Web 浏览器: Internet Explorer 浏览器或 Netscape 浏览器

2.9 系统模型

图 3-表示的是系统组件及其关系的模型。

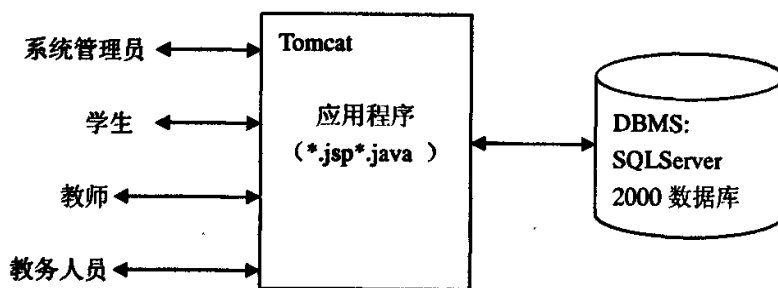


图3-2 系统模型

Fig.3-2 The system model

2.10 安全要求

身份认证一般包含两个方面的含义：一个是识别，一个是验证^[23]。所谓的识别是指对系统中的每个合法用户都有识别能力。要保证其有效性，必须保证任意两个不同的用户都不能具有相同的标识符。所谓验证是指系统对访问者自称的身份进行验证，以防假冒。标识信息一般是公开的，而验证信息必须是秘密的。下面着重研究远程教育系统对安全性的要求。

从上面的功能中可见，要保证远程教育的各种应用得以顺利实施，必须通过建立相应的安全基础设施来保证。远程教育管理系统的安全服务应提供以下安全机制功能：

- 信息的保密性机制

不向非授权用户泄露信息内容(如：不向非授权的学生泄露教学内容及试卷等内容)。

- 信息的源发性鉴别

可使授权的收信人对数字签名进行核实以可靠地确定发信人的身份。

- 信息的完整性机制

可使信息的授权收信人确信信息在传输过程中未受非法修改。

- 信息来源的不可否认性机制 -

可使通信双方对信息发送人的身份进行认证。

第三章 系统设计

3.1 基于 OOA 的 OOD

OOA 一般是方法的主体部分,解决软件生命周期中分析阶段的建模问题(建立系统的分析模型),OOD 是基于 OOA 的,是 OOA 的延续,用于解决设计阶段的问题(在分析模型基础上建立设计模型)^[7],称之为基于 OOA 的 OOD。

基于 OOA 的 OOD 有如下一些特点:

(1)它是在面向对象的分析基础上,继续运用面向对象方法进行系统设计,一般不依赖结构化分析的结果。

(2)与相应的 OOA 方法共同构成一个完整的 OOA&D 方法体系。在各种方法中,OOA 和 OOD 采用一致的概念与原则,但二者分别解决软件生命周期分析阶段和设计阶段的问题,有不同的目标及策略。

(3)较全面地体现面向对象方法的概念与原则,例如类和对象、它们的内部特征及相互关系,以及封装、继承等。

(4)大多数方法是独立于编程语言的,即,通过面向对象的分析与设计所得到的系统模型可以由不同的编程语言实现

3.2 个性化的远程学习系统的设计

3.2.1 个性化服务

什么是个性化服务?个性化服务的根本就是尊重用户,研究用户的行为习惯与兴趣,为用户选择更需要的资源,提供更好的服务。

传统行业中的个性化服务,如某国际连锁大酒店,为每一位客户建立服务档案,记录用户的习惯与喜好,然后有针对性的为客户选择需要的资源,提供更好的服务。无论顾客走到世界任何一个国家的这一国际酒店,都会享受到最贴切的服务。

互连网产业中开始提供个性化服务有 DELL 与 Amazon.com,虽然提供得并不是十分彻底。在 DELL.com 一旦用户购买了一种计算机产品,在相当长的一段时间里,网站会根据用户的需要与产品的特点,体贴地推荐最合适的配件与产品。在 Amazon.com,如果用户对旅游感兴趣,经常购买了旅游的书籍,Amazon.com 会非常体贴用户的需求,适时的推荐一些这方面的热门书籍。

数字式个性化服务将成为互联网服务的一种新模式和发展方向。但是,目前绝大多数的远程教育网站都缺乏个性化特征,因此如何把个性化服务概念进一步引入到远程教育中来,将成为远程教育发展的一个热点。

3.2.2 个性化的远程教育

尽管足不出户的学习方式给用户带来了极大的便利,但是目前的远程教育却忽略了学习本身是一种个性化的过程,没有做到因人而异、因材施教、兼顾到个体的差异,将个性化的概念引入远程教育中是非常必要的。引入个性化后的远程教育是一个智能化的系统,可以更多且及时地了解到学生的状况、需求、能力差异、学习进度、兴趣爱好,并动态地根据这诸多因素调整学习计划及进度,让用户得到针对其个性化的教育。传统的远程教育网站对每个用户来说都是一样的,浏览的课件、考试的题目与整个学习的流程都是单一固定的,在应用了个性化模型的远程教育中,每个学生都仿佛找到了一位针对自身特点进行教学的“老师”,充分发挥学习的主动性。同样一个远程教育系统,但是对于不同的受教育的对象,表现出的不同的页面和教育信息资源,这就是个性化的远程教育^[25]。

对远程教育网站来说,也可以统计学生的学习情况,更直接与科学地了解到用户的需求与自己提供的教育资源信息的冗余和不足,适时地调整自身的策略、方案来满足受教育对象的需求。

3.2.3 传统的基于 Web 的远程学习模型

基于 www 的远程教育指的是将普通教育的环节(如入学、上课、答疑、作业、考试等)放到网上来,学生足不出户便可享受到网上学校提供的服务,在家里就可以浏览课件、进行提问、和其他同学交流、参加考试等。基于 Web 的远程教育是一种非实时的远程教育,它包括以下几大支撑系统:课件及生成工具、网上答疑系统、作业发布系统、远程考试系统、远程交流系统、学习管理系统。在远程教育支撑系统的后端,有专人进行课件的制作、考试题库的整理、作业的批改,将它们放在 Web 上供学生下载浏览。

传统的远程教育,以系统自身为中心,并没有充分考虑用户需求与习惯,要求人来适应系统而不是系统去适应用户,没有充分按照人的学习规律来进行教育,所以造成了交互性差,教学效果不明显等诸多问题。

如图 3-1 所示,所谓的“远程教育学习系统”在大多数远程教育站点中只

是一个美观的界面而已，它并没有把下面五个子系统有机地结合起来，使之成为一个真正的网络学院，各个子系统有各自独立的数据库（包括用户信息和各自所用的基本信息），界面所做的工作无非是提供指向各个子系统的链接。

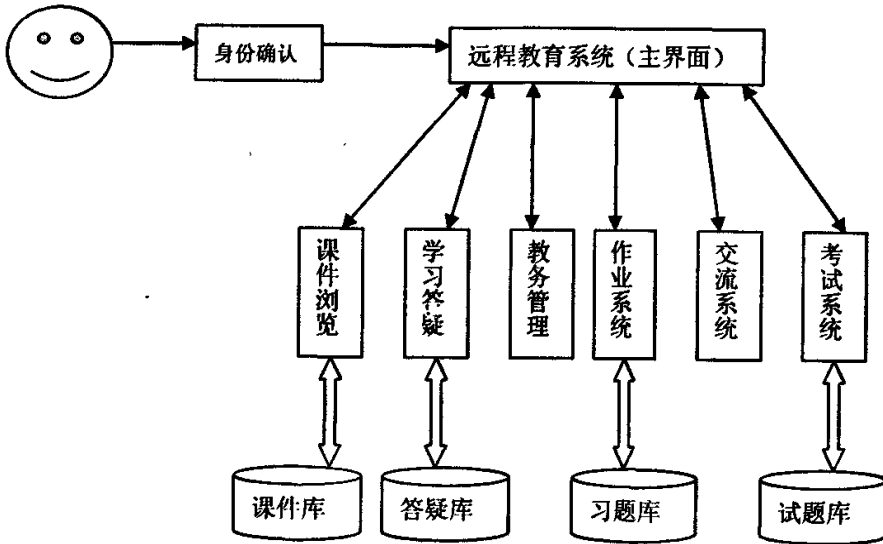


图 3-1 传统的远程学习系统

Figure 3-1 traditional remote education system

对于接受远程教育用户的行为不进行研究，忽视用户行为这一重要信息。这样，传统的远程教育系统显得不太具有人性化，智能化的特点。

3.2.4 个性化的远程学习系统

3.2.4.1 个性化处理的原理图

把个性化的思想应用于互联网上，就引出了个性化数字服务的模型，简单的结构图及描述如图 3-2。

个性化处理中信息传递过程：在图中我们可以看出，当用户登录到系统，通过身份确认后。系统的信息收集模块开始收集用户请求，跟踪用户的行为，这里一般收集的信息有用户点击的 URL、Search 的关键字、表单填写的情况、访问信息的类型、学习课件的内容及时间，以及做作业和考试的过程和结果等。然后送到原始信息库进行信息的预处理和归一化。然后个性分析处理中心根据最新采集的原始数据结合用户个性数据库进行个性分析、产生结果，更新用户个性数

数据库，并把结果送往个性调度中心，调度中心根据个性化信息和用户请求对原有数据库群发出调度命令，把用户最需要的信息反馈给用户。

得益于高速发展的互连网技术，人工智能和数据库技术，知识发现 (Knowledge Discovery in Database,简称 KDD)这种数据处理过程进入到 Web 领域，它通过机器学习，统计分析或其他方法，从大量数据中进行数据挖掘，提取有用的信息。这些使个性化数字服务成为可能，这将为目前的远程学习系统带来一次新的革命与机遇。也将使用户更多更高效的获取所需要的教育信息资源。

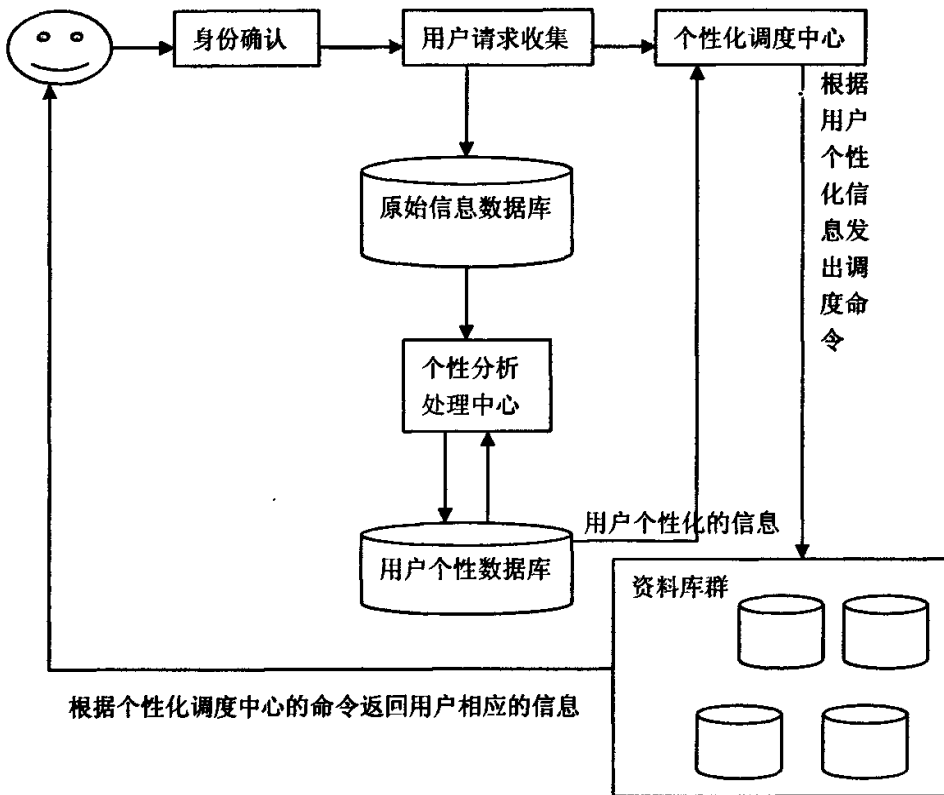


图 3-2 个性化数字服务模型结构图

Figure 3-2 the structure of personalization numeral service model

3.2.4.2 个性化远程学习模型

在此学习模型中，学生登陆系统后看到的界面不再是千篇一律的，而是与其个人密切相关的个性化的学习环境：课件学习进度、作业完成及批改情况、提问解答情况、同学列表、同学交谈记录、考试情况（做过的试卷、常错的题目及知

识点)、教师评价等都是针对该学生个人特点的, 不同学生看到的系统界面会有很大的差别。 某个学生从登陆到退出系统的整个学习过程中, 所有的动作及产生的结果都被及时记录到其本人的用户库中, 供下次登陆时系统产生用户界面作参考。

比较图 3-1 与图 3-3, 我们可以很明显的发现, 个性化的远程教育系统的模型比传统的模型增加了个性化处理引擎及一个更为详细和周密的用户信息库。从控制学的角度来看, 就是在系统中, 引入了一个反馈环节, 让系统根据不同的信息自动调整来适应用户学习需求^[25]。个性化引擎包括三个主要组成部分, 分别是用户信息收集模块、个性分析处理中心、信息调度中心。

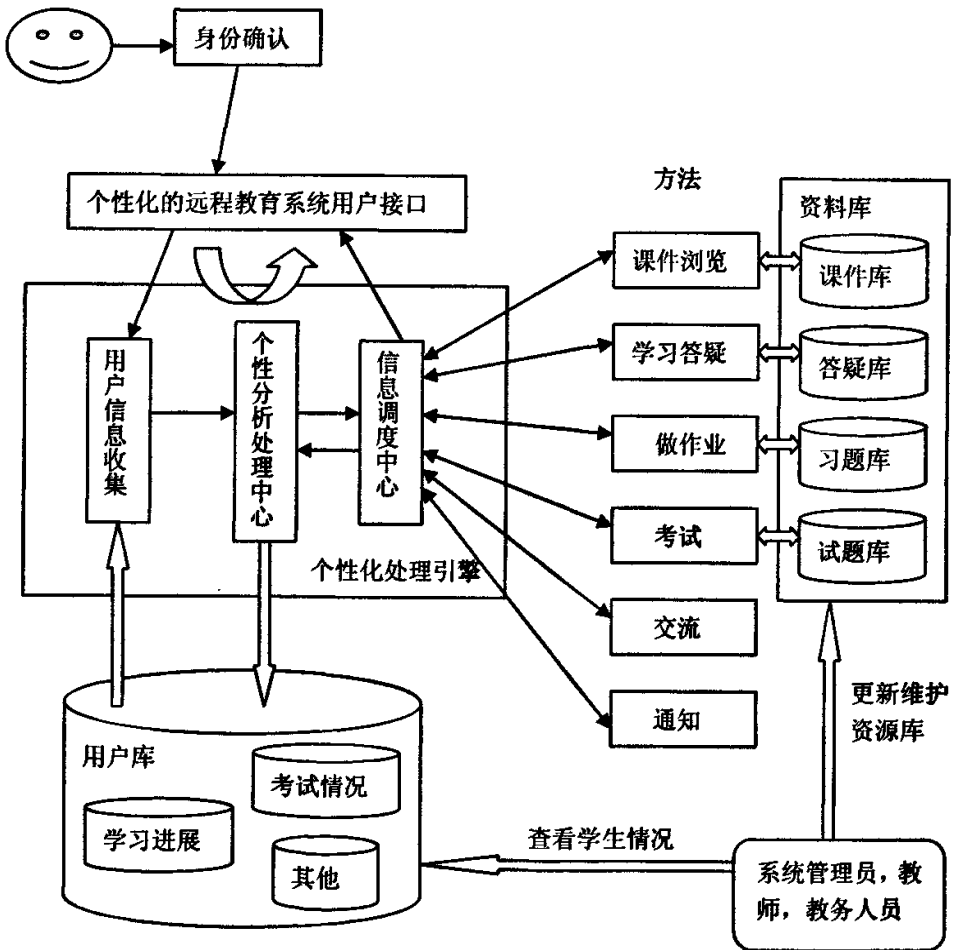


图 3-3 个性化的远程学习模型

Figure 3-3 personal remote education system model

学生登录系统后,只与远程学习系统用户接口(简称PUI, personalized user interface)打交道,学生发出的请求经用户接口传递给个性化处理引擎(简称PPE, personalized process engine),PPE再访问用户库,调度各个方法访问资料库,最后返回给PUI,整个学习过程就是在PUI与PPE之间不断地循环着。

该模型的特点如下:

首先,用户库与系统资源库分开。把所有与用户“有关”的信息(即与个性化有关的内容:如学习进度、出错知识点)单独提取出来放入用户库中,每个学生的个人数据集就组成了整个用户库,而与用户“无关”的信息(即所有学生共用的内容:如课件、问题及答案、试卷、作业等)存放到的信息库中,互相独立。比较而言,用户库是动态的、个性化的,需要经常进行读写,而课件库、答疑库、试卷库、作业库等教学资源库是静态的、共享的,大多数情况是读,极少数情况是写(发生在教师、管理员、教务人员对教学资源库进行更新的情况下)。

其次,强大的个性化处理引擎。这是整个远程学习系统的枢纽,也是体现个性化的关键。它包括三个部分:信息采集、个性分析处理中心、信息调度中心。如图所示,信息采集部分收集学生学习的动作(即对web页面发出的请求)如点击某个链接等,再从用户库里提出该学生的相关信息,递交到个性化分析中心;分析中心取得采集后的数据进行个性化分析,然后向信息调度模块发出命令;信息调度模块再根据发来的命令调度学习方法访问教学资源库,返回相应的信息(如某个课件知识点、某张试卷)到用户界面,同时返回个性化分析中心,由分析中心将结果数据分析后再写入用户库。这样学生发出的请求得到的是与其相关的个性化的内容,同时随着学习过程的变化,用户库也在不断更新,及时地保存和反映学生当前的个性信息。

3.3 知识混合结构和学习过程指导

3.3.1 总体描述

图3-4描述了远程学习的知识结构和学习过程的总体框架,它包括知识结构、学习指导和学习答疑三部分。其中,知识结构是基础,学习指导和学习答疑是在知识结构基础上实现的功能子系统。

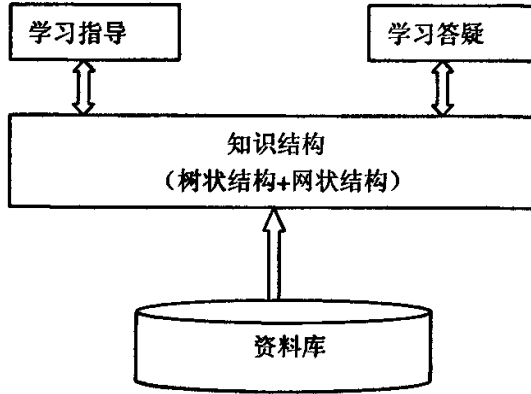


图3-4 远程学习知识结构与学习过程总体框架

Figure 3-4 knowledge structure of long-distance study and study process

3.3.2 知识的混合结构模型

出于个性化学习的要求,用于远程教育的教材内容的组织不是简单的将书本电子化,而是通过对学习内容的模块化来确定知识的内在联系和认知顺序。

在本文中,把组成教材的内部各部分称为“模块”。由于教材具有章、节等分层结构,所以模块是可嵌套的。模块间的基本关系可以分为3种:

1) 包含关系。指通常的章节之间的包含嵌套关系,形成知识的树状结构。

2) 前提关系。如果学习模块B需要以具备模块A的知识为前提条件,则称模块A是模块B的前提模块。模块之间以箭头线表示前提关系,模块之间的前提关系形成用有向图连接的网状结构。

3) 相关关系。模块之间有一定的关联,但是相互之间没有包含关系,也没有前提后续关系。例如,一个模块是另一模块的摘要,或数个模块在一起说明同一问题的不同方面。模块之间连线无箭头。

通常情况下,教材的结构主体是包含关系,由教材的分层目录,构成一棵完整的知识树。每个模块的知识属于模块的知识构成,不可再分的模块为叶子模块。在前提关系上,教材的编写一般按先基础后深入的顺序,隐含了前面模块是后面模块的前提。但是,教材知识模块间的关系并非这么简单,大量的情况是前面模块和后面模块仅仅是并列关系,或模块与其前提模块在章节顺序上相离很远。为此,在教材的组织中,以树状结构为基础,加入模块之间的前提关系和相关关系,

构成了知识的网状结构。这种表达知识之间复杂关系的结构称为知识混合结构，它包含树状结构和网状结构，如图3-5所示。

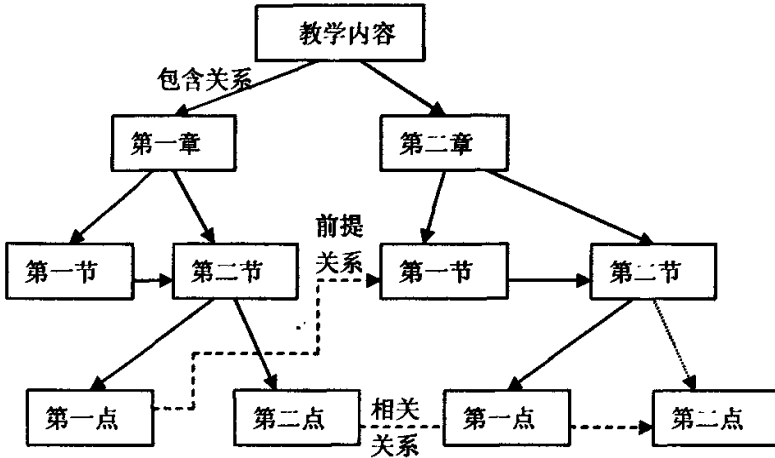


图3-5 知识的混合结构框图

Figure 3-5 composite structure diagram of knowledge

3.4 学习指导

1) 模块知识掌握估计

学习的过程是一个动态反馈的过程，学生在学习一阶段后要将学习的效果(掌握程度)反馈给教师或学习系统，教师或学习系统据此对学生的进行学习过程进行指导。为此，本文提出了一个度量参数“模块知识掌握估计(know well estimate, KWE)，取值范围在0~1之间，用以表示学生对本模块知识的掌握程度。对每一模块而言，设一个阈值KWET，如果学生的 $KWE > KWET$ ，则表示知识被掌握。

KWE有3种获取方法：

a) 客观测试

通过附在每个模块后的系列测试题获得。如果一个模块有n道测试题，则定义每道题的权值分别为 $Q[i]$ ， $0 < i < n$ ，且 $\sum Q[i] = 1$ ， $i = 1, \dots, n$ 。如果每道题只有对和错两种情况，则 $KWE = \sum Q[k]$ ，其中k为做对题的编号。

b) 教师评定

教师可以通过与学生的交流、查看学生对特定问题的回答来直接评价学生对

模块知识的掌握程度，并设置其KWE值。

c) 学生自评

2) 学习过程指导

系统按照知识结构模型建立针对个体的学习计划，对学生进行智能化学习指导：

a) 在一个模块的学习通过后，提示学生下一个应学模块编号；

b) 在模块学习中出现难以掌握的模块时(KWE值太小)，向学生显示当前模块的包含、前提和相关关系，以帮助学生确定复习的方法；

c) 搜索当前模块的整个前提模块链，并向学生提示其中KWE值较低的模块，以便学生针对自己的弱点进行复习；

d) 问题仍然无法解决时，转向教师参与过程。

3.5 学习答疑

3.5.1 设计目标

由于在远程教育过程中学生不能直接向教师请教，并受简单答疑系统的影响，学生在自学过程中遇到的问题不能得到及时的解答。因此，建立智能性和自适应性的答疑系统就成了目前开发网上教育平台的重要任务之一^[26]。目前，有许多远程教育平台都加进了智能答疑系统，但总存在这样或那样的不足。现就目前远程教学系统中有关答疑部分的缺点简要总结如下：

(1) 系统的智能性、自适应性不够；

(2) 学生用户提问的操作过程繁杂；

(3) 答疑系统同学生学习用的答疑材料库和课件库相脱离，没有充分体现授道与解惑的一体性；

(4) 答疑系统大多只给出参考答案，没有给出学生应巩固学习的知识点或学习策略的建议，个别的系统仅在答疑之后给出一些巩固练习题，这是不够的。

鉴于以上情况，这里提出一种智能性、自适应性较强的智能答疑系统的设计模型，系统为学生用户提供图形化的简单操作界面，学生可在学习过程中随时以简单的鼠标点击选择和输入自然语言问题来使用自动答疑系统。

总之，智能答疑子系统的设计目标是：

(1) 学员用户能在学习过程中随时中止课件演示进程，运用自然语言进行提

问，并可对提问方式进行选择(同步提问(系统自动回答、Web实时讨论)、异步提问(邮件答疑、留言板答疑))，也可选择系统自动回答的相关参数(如答案大小等)。

(2)系统既能通过在常见问题库或知识点库和课件素材库中搜索答案并自动回答用户所提问题，又能通过电子邮件、实时讨论或留言板方式使学员用户的问题得以解答。

(3)系统呈现给用户浏览器上的答案应包括两方面：一是多媒体形式的图文并茂的解答；二是用户应巩固复习的知识点建议及少量的巩固练习题。

(4)系统自动生成和维护常见问题库，并保持其结构的良好性，同时也支持专任教师用户的人工维护(增加、删除和修改)等。

3.5.2 基于 web 的智能答疑系统的总体结构

智能答疑系统的用户应有三类：一类是使用系统请求解答自己疑问的学生用户，他们是系统的服务对象；二是某课程的专任教师，他们负责构建课件库、知识点库，并提供对课件库、知识点库、常见问题库的不定期维护，还负责回答系统无法自动回答的问题；三是系统管理员，负责整个系统的安装和软、硬件维护，维持整个系统的良性运行。

根据智能答疑系统的设计目标和用户分析，构建如图3-6 所示的系统总体结构。

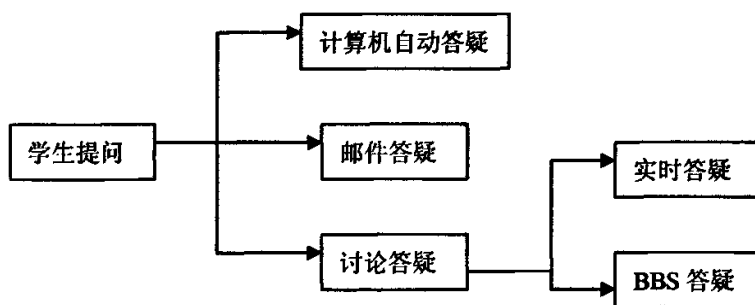


图3-6 基于web的答疑系统的总体结构模型

Figure 3-6 structural model of Q/A system based on web

3.5.3 搜索方法和常见问题库的建立与维护

结合包含关系导致的树状结构和前提关系导致的网状结构，这里给出一种比

树状结构优越的信息表达方式,即用逐渐展开目标集方式^[20]缩小提交给学生的信息量,并把问题按与学生问题接近程度的估计值排列。

1) 搜索方法

a) 模块内搜索

将产生问题的模块设置为起始搜索范围(高可能性范围),利用关键字搜索及逐条人工查看。

b) 按知识结构扩展搜索范围

在当前模块答疑失败后,系统搜索进入中可能性范围,按关系类型(包含、前提、相关)列出附近模块,并在附近模块内搜索,得到常见问题库。

c) 全局搜索

如果经过逐渐展开范围搜索,学生仍未找到答案,则抛开学生选取的初始模块,系统搜索进入低可能性范围,利用全局关键字搜索。

2) 常见问题库的建立与维护

如果智能答疑不能解决学生的疑问,问题将提交给教师,由教师参与回答,有两种可能:

a) 一个新问题。教师作出回答,并将问题与答案作为一个新的问题加到相应的模块中,从而不断丰富答疑知识库;

b) 一个问题库中已经解答的问题,但是放置的位置并不是学生搜索范围可能性最高的地方。学生的搜索记录向教师提供了一个较客观的参考和提示,通过不断地反馈和修正,把它放在学生最可能首先查找的模块中。

3.6 开发模式的选择及设计

本系统采用三层结构,即 B/S 模式,所谓的三层结构是指客户端应用程序 / 应用服务器 / 数据库服务器这三层结构,该体系将业务逻辑放在应用服务层,应用服务层接受客户机即表示层的业务请求,根据请求访问数据库,作相关处理,将处理结果返回客户机。这种“瘦”客户机,“胖”服务器的特点使维护变得很容易。

该结构各层次主要功能为:① 数据层:在这一层上主要存放与教育系统相关的数据,如课件、试题、答案等相关数据库以及用户信息管理数据库。数据以表的形式存放于数据库中,为上层的功能实现提供信息基础。②业务逻辑层:业

务逻辑层的主要职责是接受上层发来的用户请求，对数据层的数据进行存取，并根据服务要求用相应的服务程序完成用户请求功能。③表示层：表示层设计用户接口，按照用户类型产生基本用户格式表示和用户请求类型。

第四章 系统实现

4.1 面向对象的实现

面向对象的实现又称面向对象的编程，在OOA—OOD—OOP这一软件工程的过程系列中，OOP的分工比较简单：认识问题域与设计系统成分的工作已经在OOA和OOD阶段完成，OOP工作就是用同一种面向对象的编程语言把OOD模型中的每个成分书写出来^[6]。OOP阶段产生的程序能够紧密地对应OOD模型；OOD模型中一部分对象类对应OOA模型，其余部分的对象类对应与实现有关的因素；OOA模型中全部类及对象都对应问题域中的事物。

4.2 个性化的处理引擎

4.2.1 用户信息采集模块

这一模块是个性化远程教育的基础，其采集数据的类型由个性化远程教育系统的用户分析模型所确立，它是个性分析引擎的数据来源，其收集信息的质量与数量将直接影响个性系统分析的质量，采集的基本信息有用户经常访问的地址与栏目、感兴趣的话题、用户的IP地址、停留时间、访问的时间频率、经常提问搜索的关键字、作业与考试的完成过程与结果、进行交流讨论的话题、访问的时间段、用户的必要个人信息。信息采集模块负责收集这些信息，并把这些采集的信息顺利的送往下一步的个性分析中心。

对于用户来说，一旦登陆这个网站，其在上边发生的大多数的行为都可能被信息采集模块所收集，用来进行用户的个性分析。

4.2.2 个性分析处理中心

个性分析处理中心有两部分，第一就是个性信息源，由第一部分的信息采集模块负责。第二就是个性分析的模型。个性分析的模型是一个非常复杂的模型，是对远程教育系统的教和学两个过程和对象特点都做了充分研究的基础上建立的。个性分析处理中心是整个个性化数字服务的关键。

个性分析处理中心有两大主要组成部分：一是用户的信息库，二是分析处理模块。它的工作方式是分析处理模块结合最新采集的信息与用户库中的数据资源进行新的分析，并把分析结果传递到信息返回模块，同时更新用户信息库中的数据。

4.2.3 信息调度中心

信息返回模块根据个性分析引擎处理后的个性参数来调度几大子系统,并选择合适的教育信息资料并将页面反馈给用户以满足其需求,同时,把用户在子系统个性活动信息反馈给个性分析处理中心,然后由个性分析处理中心进行再加工,来更新用户信息库。对于用户来说,其发出一个请求得到一个响应,中间的过程是透明的,其个性处理的内部过程并不为用户所察觉。

4.2.4 各子系统的工作协调

这些子系统之间的联系是由个性化引擎和用户个性信息库来完成。比如说,一个学生在答疑系统中对于某方面的问题提问次数比较多,那么相应地这一信息将会被个性化引擎及时捕获,并记录到用户个性信息库中,在下次用户进入作业系统中,个性化引擎将根据这一信息调度作业,系统对此学生着重进行这一方面问题的训练。

4.3 个性化远程教育模型关键技术及实现

4.3.1 利用基于 Web 的 Data Mining 技术(Web Mining)设计个性化处理引擎

数据挖掘(Data Mining)就是指从大量数据(如数据库)中提取抽象的、潜在的有用信息的过程,它是数据库中知识发现 KDD(Knowledge Discovery in Databases)的核心,为大量数据的利用提供了有效的工具。数据挖掘的成功使研究人员将数据挖掘技术用于 Internet,但是由于 WWW 上的数据不同于传统数据,故产生了一个新的研究方向--Web 挖掘(Web Mining)。

但在这里,在个性化数字服务模型的实现中,应用到对于用户综合行为的数据挖掘,这不同于日志挖掘,日志挖掘是针对于一个站点综合信息的挖掘,而在这里我们需要实现的是针对用户的数据挖掘,用户在网站上几乎所有可能被记录下来的行为,都要尽可能的挖掘,分析后的结果作为为用户提供信息服务的依据。

4.3.2 利用 XML 技术设计个性化用户接口

XML 即 Extensible Markup Language(可扩展的标记语言),它是由 Worldwide Web Consortium(W3C)组织开发的基于 SGML(Standard generalized Markup Language)的一种标记语言,是 SGML 的简化版本。它比现在流行的 HTML(超文本标记语言)更为灵活且功能更先进。不同于 HTML(单一的标记语言)的是,XML 是一种元语言,支持自定义的标记。

正是由于 XML 是一种自由形式的语言, 尤其支持个性化的描述, 将彻底改革原有网站的信息响应模式, 应用在个性化数字服务模型的信息返回和调度模块, 所以对于个性化远程教育模型而言是一个关键技术。

4.4 数据库的设计与实现

4.4.1 学生库

用户库的信息设计要充分综合各子系统的信息特点与用户个性特点, 充分考虑子系统间信号通讯的方式。

学生库中存放了针对每个学生的详实的个人资料, 它由学生表、学生当前状态表、课件学习进度表、答疑表、作业表、考试表、交流表等构成, 具体说明如下:

学生表 (一个学生对应一条记录)

学号	学生名	密码	性别	出生年月	入学时间	所属系	地址	联系电话	Email	上次登陆时间
----	-----	----	----	------	------	-----	----	------	-------	--------

学生当前状态表 (一个学生对应一条记录)

学生名	是否进行课件学习	章节号	是否进行答疑	问题号	是否进行考试	
-----	----------	-----	--------	-----	--------	--

试卷号	是否进行交流	交流编号	是否进行作业	作业号	登陆时间
-----	--------	------	--------	-----	------

课件学习进度表

学号	课程代号	所学章号	所学节号	浏览时间	难度系数	掌握程度
----	------	------	------	------	------	------

答疑表 (一个问题对应一条记录)

学号	问题号	问题内容	有否解答	提问时间	难度系数	解答内容
----	-----	------	------	------	------	------

作业表 (一个作业对应一条记录)

学号	作业号	回答内容	有否批改	正确与否	作业提交时间	难度系数
----	-----	------	------	------	--------	------

考试表 (一次考试对应一条记录)

学号	试卷号	有否批改	分数	考试结束时间	难度系数
----	-----	------	----	--------	------

阶段测验表

学号	阶段测验号	有否批改	分数	难度系数
----	-------	------	----	------

试题项（一次考试对应一张表，一道题对应一条记录）

大题号	小题号	分值	对或错	难度系数	得分
-----	-----	----	-----	------	----

交流表（包括同学列表，交谈记录，通信记录等表，这里只给出同学列表）

同学列表

学生名	同学名	友好等级	课程号
-----	-----	------	-----

教师意见表

学生名	教师名	课程号	教师意见
-----	-----	-----	------

易错知识点表

学生名	课程号	章号	节号	易错知识点
-----	-----	----	----	-------

总之，学生库就象是结合了学生的个人信息的对教学资源库的一个索引。

4.4.2 教学资源库

除了将用户信息部分抽取掉以外，该库与传统的教学资源库没什么差别。

课件库

课程代号	课程名	任课教师	开课时间	学分	存储位置	简介	课件类型
------	-----	------	------	----	------	----	------

作业库

作业号	所属课程	作业题目	标准答案	发布时间	提交期限	难度系数
-----	------	------	------	------	------	------

答疑库

问题号	问题内容	标准答案	关键字	提问者	回答人	提问时间	解答时间
-----	------	------	-----	-----	-----	------	------

试卷库

试卷号	所属课程	出卷时间	总分	出卷人	简单说明	标准答案	发布时间	时间限制
-----	------	------	----	-----	------	------	------	------

以后还可以扩充其他的库，如公式库、图片库。他们的共同特点是只能由学生读取，由教师、教务人员或管理员进行维护。

4.4.3 用户接口

学生登录系统后就一直与学习系统的界面打交道，向界面发出请求，由界面向学生返回请求的结果。由于对学生学习过程的分析 and 调度都由个性化处理引擎来完成，所以这里的用户接口的设计很简单：它只需将学生发出的请求传递给个性化处理引擎，再将其他返回的结果显示出来即可。它可由一个登录页面、一个主页面和每个学习工具的界面和其他页面(如查看通知、查看教师意见)组成。

4.4.4 教师和管理员端

教师和管理员的操作包括两个方面，一是对学生库(即用户库)的管理，二是对教学资源库的管理。这一部分不需要进行个性化处理，值得一提的是，由于教学资源库不涉及学生的个人信息，故答疑、批作业、批试卷是先从学生库中得出的信息(如题目序号、试卷号)，再调出教学资源库里的内容。然后进行操作，最后存入学生库中。教学资源库只起到一个内容参考的作用。

与传统的远程教育模型中课件库、答疑库、作业库、考试库所不同的是，在个性化的远程学习模型中，这些库中不必包括具体的用户信息，它们是各自独立的，功能也更加专一，将各数据库独立开来的好处在于学习模型的模块化设计，各子系统的开发、修改和升级不会破坏系统的整体性。

4.5 Web mining的实现

个性化的远程教育系统，采用数据挖掘技术，通过对站点上积累的大量信息进行分析，以发现学生感兴趣的模式和规则，为广大教师、管理人员提供更优工作效率、工作水平的信息。在远程教育模型中，增加了一个web mining模块，该模块位于服务器端，定期地分析积累在站点的大量信息^[30]。

4.5.1 仓库的建立

对现有数据经有效整合后，形成两个存储数据库，即：学习记录数据库(用于存取历史学习成绩，当前学习成绩和将来的学习目标，以提供学生学习的连续

性)、知识库(作为学习内容的资源库,可以根据查询索引进行学习内容的查找,匹配的学习内容以内容索引的形式送往指导系统。从内容索引中提取定位索引并把它送往传递进程)。由此形成的数据仓库将满足决策系统、远程教育对数据的需求。具体过程是:

将现有的远程教育学生的地区、年龄、所学专业、所学课程、学习成绩等资料加载到学校的信息库,以建立学校的数据库基础环境,在远程教育站点的服务器端,存放的是整个教育系统的学习资源,网上讲课和其他支持决策活动和远程教育的事务处理软件资源,事务处理中需要解决的各种问题是经常出现的,后果影响范围小,因此远程教育的事务处理完成日常的数据处理。

此远程教育模型的实现,在客户端,主要是一个Web浏览器,如IE4.0、Netscape4.0等,服务器端的解决方案很多,如NT Server4.0+IIS4.0 SQL Server6.5,该模型有很多技术可以实现,如HTML技术、CGI技术、ASP技术、ActiveX和Java技术等。

4.5.2 Web mining模块的实现

Web mining就是文档的内容,可利用的资源的使用及资源间的关系进行分析,以发现有效的、新型的、潜在有价值的,并且最终可理解的模式和规则,当前主要研究web mining技术有Web Usage mining和Web content mining。web Usage mining通过利用数据挖掘技术对站点上积累的大量的Web数据进行分析,以发现用户站点的浏览模式、页面的访问频率等,为站点的管理员提供各种利于Web站点改进信息,以及潜在的需求或可带来经济效益的信息^[31]。Web content mining就是利用挖掘算法自动搜索Internet上的信息资源,在Web mining中采用的算法,有关联规则发现算法,序列模式发现算法,以及聚类算法,Web mining模块的实现如下:

(1)前期处理包括数据分类、整理,将现有业务系统中有关学生的访问数据中与数据挖掘无关的数据清出,方法是以文件的后缀名进行数据清除。用户识别,采用基于日表/站点的方法,启发性规划来识别用户。会话识别,将用户访问的页面分成单个的会话,路径补充:用户访问的路径对用户会识别很重要,可以利用类似于用户识别的方法进行路径补充。补充识别事件:目前访问路径长度和最大目前访问路径2种方法,用于事件识别。

(2)挖掘过程。站点上积累的信息包括学生学习进程,一方面从传输进程接收以多媒体方式显示的学习内容进行学习,另一方面他的学习行为被评估进程接收。还包括评估、指导等信息。从数据库中大量数据的分类、整理中获取有用信息,利用一些挖掘算法来挖掘出模式、规则等,为了提高数据分析的有效性和准确性,根据待分析数据的不同特征使用不同的挖掘算法。

(3)模式分析。利用一些算法对挖掘出来的模式,规则进行分析,找出感兴趣的模式和规则。

(4)Web mining的集成。Web mining放在服务器端,每隔一段时间对服务器端的系统日志和用户与站点的交互数据进行分析,并向管理人员、教师报告各种信息。

4.6 系统安全

远程教育系统的安全,是每一个设计者面临的不容回避的问题,系统实现为:访问每一个站点的用户都必须提供许可证,经验证授权的用户才能访问。验证用户的身份依据资料的重要程度可采用不同技术,在IIS4.0中,可采用“基本验证”、“Windows NT口令/回送验证”和“用户端认证身份验证”。“基本验证”通过收集使用者名称和密码的方法来确定用户身份。在“Windows NT口令/回送验证”方式中,使用者的浏览器与网站服务器以特殊密码交换的方式验证身份,实际的密码并未通过网络传送。“用户端认证身份验证”让网络服务器自动辨识使用者的身份,此种验证方式实现较复杂。

4.7 本系统的特点

本系统特点之一是采用了面向对象的软件工程方法来进行系统的研究,在需求分析阶段运用UML标准建模语言和Rational Rose建模工具对远程教育系统进行建模,并给出了类图。

本系统的特点之二是采用三层B/S体系结构,整个系统由web浏览器,web服务器和数据库服务器三层组成,应用程序放在web服务器上,数据库存储在数据库服务器中,用户可以在Internet上任意客户端上的浏览器中通过web服务器管理和访问数据库中的教学信息。因此系统具有大用户量,良好的可伸缩性和可移植性。

本系统的特点之三是在服务器端开发使用Java技术,使得本系统具有跨平台性、可移植性、高效性、可重用性、可扩展性和安全性。

本系统的特点之四是在应用开发过程中使用了面向对象编程的MVC模型,将用户界面,流程控制和业务逻辑的应用程序分别设计和编程,使应用系统条理清晰、容易继承和便于维护。

第五章 系统测试

5.1 测试的方法

对程序进行测试的方法有两种：白盒测试和黑盒测试。白盒测试又称为结构测试，是在知道产品内部工作过程，完全了解程序的结构和处理过程的前提下，通过测试来检验产品内部动作是否按照规格说明书的规定正常进行。黑盒测试又称为功能测试，是在程序接口进行的测试，是已经知道了产品应该具有的功能，通过测试来检验每个功能是否都能正常运行，能正常地接受数据产生正确的输出信息，并且保证外部信息（数据库或文件）的完整性^[5]。

无论是黑盒测试还是白盒测试，都不可能达到穷尽测试。因为如果使用黑盒测试法，为达到穷尽测试，要至少对所有输入数据的各种可能值的排列组合都进行测试，由此得到的应测试的情况数往往达到实际上根本无法测试的程度；如果使用白盒测试法，为达到穷尽测试，程序中每条可能的通路都应该在每种可能的输入数据下执行一次，所以即使测试很小的数据都不能做到这一点。既然都不能达到穷尽测试，那么软件测试也就不可能发现程序中的所有错误，也就不能通过测试证明程序是正确的。但是，这不是我们的目的，我们的目的是要通过测试保证软件的可靠性，用尽可能少的测试发现尽可能多的错误。

5.2 测试的步骤

由于在面向对象的程序中，最基本的单元是类，所以面向对象软件的测试相对于传统的软件测试来说有所不同^[43]。传统的软件测试包括模块测试、子系统测试、系统测试、验收测试和安装测试五步，而面向对象的软件测试通常分为单元测试、集成测试和确认测试三步，其中单元测试和集成测试采用白盒和黑盒相结合的测试方法，确认测试则采用黑盒测试的方法。现分别说明如下：

1. 单元测试

单元测试是对程序中的各个类分别进行测试。因为类是一个封装体，封装了这个类的所有的属性和方法，因此对类的测试包含对类的所有服务的测试，包含对类中各个服务与属性间作用的测试。

单元测试的目的是发现并纠正各个类内部的错误。由于类中的服务规模小、功能单一、逻辑简单，因此有可能对服务中的所有执行路径都测试到，如果因路

径数太大而做不到的话，则设计若干测试用例测试一些重要的执行路径（以白盒法为主），然后设计一些合理的和不合理的测试用例测试服务的功能（以黑盒法为主）。

2. 集成测试

单元测试纠正了各个类中的大部分错误。但是各个类只有相互配合才能很好地实现软件系统的各项功能，因此第二步就用集成测试来发现类之间配合上的问题和验证系统的功能。

因为应用程序的作用是使各个类相互配合从而实现软件的各种功能，因此集成测试也就是对应用程序进行的测试。鉴于缩小差错范围的需要，在测试前可以从应用程序的功能或使用事件的观点来观察和划分系统中的类，这样当应用程序相应某个使用事件（用户命令）或完成某个功能时，应用程序中只有部分的类会被激活。

因此在测试应用程序时，可以对每一个使用事件设计若干组测试用例，使用每一组测试用例观察程序的运行结果，如果运行结果有错，则从参与相应该事件的类的相互配合上找原因。

3. 确认测试

确认测试是在集成测试完成之后进行，目的是检验所开发的软件是否能按照需求说明书中规定的要求进行。确认测试是在用户积极参与下进行的，而且应尽可能使用系统将来要处理的实际信息数据进行测试。

针对本例中所做的远程教育系统的自主学习子系统的学习模块进行的测试，应该按照上述三个步骤进行的，首先采用白盒测试法对每一个类及其服务设计测试用例测试一些重要的路径，然后再用黑盒测试法测试服务的功能；接着再采用白盒测试法测试类之间的配合连接；最后采用黑盒测试法进行用户的确认测试。经过测试，所做的系统基本符合自主学习模块功能和流程的要求。

第六章 总结与展望

6.1 总结

本课题是基于 web 的远程教育系统的研究，侧重于入学系统，学习系统和作业/考试系统的分析，设计与实现，特别是学习系统中的自主学习模型的提出与建立以及答疑系统的设计，在不同的开发研究阶段，运用自己所学的知识 and 实际经验，以及不断的学习新的知识和技术，逐一解决各个阶段出现的问题和困难。通过对出现问题的分析，研究直到解决，从中获得了很大的提高。

6.1.1 在理论上的收获

通过基于 web 的远程教育系统的研究，加深了对 web 数据库的体系结构，web 数据库访问技术，数据挖掘技术，MVC 设计模式，Servlet 技术，JSP 技术，JavaBean 技术的理论及应用的理理解，同时对软件工程，面向对象的分析和设计方法有了进一步的认识，为今后的学习和工作打下了坚实的理论基础。

6.1.2 在技术上的收获

通过基于 web 的远程教育系统得研究，对 Java 平台，统一建模语言 UML 及建模工具 Rational Rose，SQL Server 2000 数据库的应用有了一定程度的了解，为今后的学习和工作奠定了一定的技术基础。

6.2 不足及未来发展方向

通过对基于 web 的远程教育系统得研究，随着对系统理解的不断加深，自身认为系统分析以及设计有待完善，还请各位老师予以斧正。

6.2.1 不足

远程教育系统是一个庞大而复杂的系统，因此，在系统的分析、设计与实现等各个阶段都存在着一些不足。本文重点研究了远程教育系统的建模和设计方案。

6.2.2 未来发展方向

从长远的发展的观点来看，远程教育系统有着很大的市场，所以完成一个功能完善，界面简单，并且有个性化的学习指导，智能完善的远程教育系统是一件很有意义的工作，本课题建立了基本模型，但是基于此模型的实现要用到多方面的知识和技术，具体的实现方法还不是很成熟。但个性化和智能化的概念和模型

是远程教育未来的发展模式。

参考文献

- [1] 教育部. 现代远程教育标准体系 (Distance Learning Technology Standards, DLTS) 之学习资料相关部分(DLTS-1 DLTS-2 DLTS-3 DLTS-4 DLTS-9).
- [2] 朱百钢主编. 现代远程教育及校园网建设全书[M]. 北京: 中国工人出版社, 2001: 10-12.
- [3] 张剑平. 现代远程教育及其发展[J]. 现代信息技术, 2000年第5期: 25-28.
- [4] Ronald J.Norman 著, 周之英, 肖奔放译. 面向对象分析与设计[M]. 北京: 清华大学出版社, 2000: 30-32.
- [5] 张海藩. 软件工程导论(第三版)[M]. 北京: 清华大学出版社, 1998: 20-25.
- [6] 邵维忠, 杨芙清. 面向对象的系统分析[M]. 北京: 清华大学出版社, 2003: 12-15.
- [7] 邵维忠, 杨芙清. 面向对象的系统设计[M]. 北京: 清华大学出版社, 1998: 16-18.
- [8] 陈钧, 董宏, 何英杰. 面向对象分析与设计在 MIS 开发中的应用探讨立[J]. 计算机系统应用, 1995: 40-43.
- [9] 刘超, 张莉编著. 可视化面向对象建模技术—标准建模语言 UML 教程[M]. 北京: 北京航空航天大学出版社, 1999: 30-32.
- [10] 冀振燕编著. UML 系统分析设计与应用案例[M]. 北京人民邮电出版社, 2003: 50-52.
- [11] 周达仁, 张昱, 陈意云. 用 UML 和 Rational Rose 实现面向对象的三层 C/S 结构设计[J]. 计算机工程, 2000: 23-25.
- [12] 王瑞卜. 基于 B/S 模式的远程数据采集系统的研究[J]. 计算机应用, 2003: 35-37.
- [13] 李文生, 潘世兵. 基于 WEB 的多层客户/服务器数据库应用程序[J]. 计算机应用研究, 2001年第2期: 21-23.
- [14] 樊胜. WEB 数据库的访问方式[J]. 情报科学, 2002: 16-18.
- [15] 张仁平, 周庆忠. 开发基于 web 的数据应用程序[J]. 计算机系统应用, 2004(10): 30-34.
- [16] Time 研究室, 萧仁惠. JDBC 数据库程序设计[M]. 北京: 中国铁道出版社, 2004: 82-84.
- [17] 刘丽珏, 张龙祥编著. JDBC 与 Java 数据库程序设计[M]. 北京: 人民邮电出版社, 2001: 81-82, 110-112.
- [18] Hall, M.Brown., 赵学良译. Servlets 与 JSP 核心编程[M]. 北京: 清华大学, 2004.

- [19] Ashton Hobbs 著 , 丁一青, 吴静等译. JSP 高级编程[M]. 北京: 清华大学出版社, 2001: 73-75.
- [20] 戴英侠, 连一峰, 王航. 系统安全与入侵检测[M]. 北京: 清华大学出版社, 2002 : 22-24.
- [21] 余晨, 李文炬. SQL Server 2000 培训教程[M]. 北京: 清华大学出版社, 2003: 11-12.
- [22] Jiawei Han Micheline Kamber 数据挖掘概念与技术[M]. 北京: 机械工业出版社, 2001 : 9-10.
- [23] [美]Merike Kaeo. 网络安全设计[J]. 计算机工程与应用, 2003: 21-23.
- [24] 刘青云, 章剑林, 商玮. 远程教育管理系统的身份认证研究与设计[J]. 计算机应用研究, 2001 年第 8 期 : 44-46.
- [25] 舒蓓, 申瑞民, 王加俊. 个性化的远程学习模型[J]. 计算机工程与应用, 2001.9 : 90-92.
- [26] 赵成龙, 薛欣. 基于 web 的智能答疑系统的设计与实现[J]. 计算机工程与科学, 2005.8 : 31-34.
- [27] 王涛, 陈怀义. 基于 web 的远程教学平台的设计与实现[J]. 计算机应用研究, 2003.20 (7): 103-104.
- [28] 龚婕, 王普, 周丽萍. 现代远程教学系统的设计和实现[J]. 计算机工程, 2002 Vol:28 No.5: 275-277.
- [29] 李豪荣, 张轩, 陈军, 耿先骏 . 智能化远程学习模型[J]. 清华大学学报(自然科学版), 2002.Vol.42.No.1: 100-103.
- [30] 曲红亭, 申瑞民. 基于数据挖掘的个性化学习导航系统的设计与实现[J]. 计算机工程, 2003.5: 59-61.
- [31] 陈恩云. 基于 web 的远程教育系统模型的研究[J]. 武汉理工大学学报, 2001. Vol.23.No.9: 84-86.
- [32] 刘连庆. 大型国企薪酬管理信息系统设计与实现[D]. 大连: 大连理工大学, 2005.
- [33] 王秀坤. 基于web的伦理法人会MIS系统的设计与实现[D]. 大连: 大连理工大学, 2003.
- [34] Grady Booch, James Rumbaugh, and Ivar Jacobson. The Unified Modeling Language User Guide.Assison-Wesley. 1999.
- [35] Sami,Akihiro;Syonura,Tsutomu;Suganuma,Hiroshi;Ishida,Itaru;Presentation layer

- framework of web application systems with server-side Java technology, Proceedings-IEEE Computer Society's International Computer Software and Applications Conference 2000, IEEE, Los Alamitos, CA, USA, 000B37156:473-478.
- [36] ISO/IEC 9126 "Information Technology Software Quality Characteristics and Metrics"; Part 1: Quality model, Part 2: External metrics, part 3: Internal metrics.
- [37] Tse, T.H. Comparative Review of Object-Oriented Programming Texts "Computing Review 35,4(April 1994):187-190.
- [38] Sharble, Robert C & Cohen, Samuel S. "The Object-Oriented Brewery: A Comparison of Two Object-Oriented Development Methods." SIGSOFT Software Engineering Notes 18,4(April 1993):60-73.
- [39] Booch, Grady. Object-Oriented Analysis and Design, Second Edition. Redwood City, Calif: Benjamin/Cummings, 1994. ISBN 0-8053-5340-2.
- [40] Gehringer, Edward F.; Louca, Tony M.; Web-based computer architecture course database; Proceeding-Frontiers in Education Conference V2200, IEEE, Piscataway, NJ, USA, OCB37135:S 2B-12-S2B-16.
- [41] Kenneth S. Rubin and Adele Goldberg, Object Behavior Analysis Communications of ACM Vol.36 No.9 SEPT 1992.
- [42] Peter Coad: Object-Oriented Patterns Communications of ACM Vol.36 No.9 SEPT 1992.
- [43] Grady BOOCH. Object-Oriented Analysis and Design with Application. USA: Addison Wesley Longman Inc, 1994, 137~162.
- [44] Roger S. Pressman., Software Engineering, A Practitioner's Approach, McGraw-Hill, 1977.
- [45] Ronald J. Norman, "Object-Oriented Analysis and Design", 清华大学出版社, 1997.
- [46] IA16JURS Hirschberg, Gerhard Schmitt, David Kumann, Branko Kolarevic, Brian Johnson, and Dirk Donath. THE 24 HOUR DESIGN CYCLE-An Experiment in Design Collaboration Over the Internet. CAADRIA'99-Proceedings of the Fourth Conference on Computer Aided Architectural Design Research in Asia. Shanghai, China. 1999.05.
- [47] A17] Mitsuo Morozumi, Yasuhiro Shounai, Riken Homma, Kazuhisa Iki, and Yuji Murakami. A GROUPWARE FOR ASYNCHRONOUS DESIGN COMMUNICATION

AND PROJECT MANAGEMENT .CAADRIA'99-Proceedings of the Fourth Conference on Computer Aided Architectural Design Research in Asia, Shanghai, China.1999.05.

[48] A18]Ms. Neo Mai&Mr. Ken Neo TK. Multimedia Learning: A strategic approach in using instructional technology in higher education. The First International Conference on Learning and Teaching On-line.Guangzhou.China.2001.01.

[49] [美]Bertrand Meyer. The Future of Object Technology[J].Computer,2001,(3):263-294

[50] [-UFurcas R Giua. Hybrid Petri Net Modeling of Inventory Management Systems [J], IEEE, 2000, 29(6):541-547.

[51] Jiang, Jun-jie; Li, chun-Bo; Huang, Li -Ping; Li ,Jian-Ming, Web-based 3D Dynamic browsing and markup system of product model, Jisuanji Jicheng zhizao xitong/Computer manufacturing Systems, CIMS v7 n 9 September 2001: 36-39.

致 谢

本文的完成是在导师庄晋林副教授的精心指导和悉心关怀下完成的。衷心感谢导师三年来给予我的谆谆教导。导师严谨的治学态度、渊博的知识和无私奉献的精神是我终生学习的楷模。在此我向导师表示衷心的感谢和深深的敬意。

论文的完成离不开华北水利水电学院信息工程系和研究生处的各位老师的热情帮助，离不开同学们的热情协助，离不开我的师姐、师弟和师妹的关怀和帮助，离不开朋友们的真诚相助，在此对帮助过我的老师、同学和朋友表示衷心的感谢。

特别感谢亲人们的无私亲情奉献，亲人们的关心、理解和支持是我不断克服困难、勇往直前的动力，并帮助我顺利完成学业。

同时衷心感谢在百忙之中评阅论文和参加答辩工作的各位专家和教授！

无论今后走到哪里，我将铭记曾经关心和帮助过我的人们的友谊！

任晓霞

2007年5月20日

攻读学位期间发表的学术论文及参加的实践活动

1. 发表论文:

庄晋林, 任晓霞. 基于 OOA 的远程教育系统开发研究[J]. 华北水利水电学院学报 Vol.28 No.3

2. 实践活动:

2006.08—目前 研究基于 web 的远程教育系统, 主要工作是远程教育系统的
需求分析和系统设计

2005.03—2005.12 讲授《计算机网络基础》课程

2005.05—2005.09 参与了“网络课程”《数据结构》在线学习网站的制作, 这个
项目包括课程的在线学习、在线考试、教学动画、教学
视频、留言系统