

不同实验模型数据存储的分析与设计^①

崔超艳, 张嘉惠

(陕西省微生物研究所 计算机信息室, 西安 710043)

摘要: 科研实验中不同的实验模型产生结构各异的数据, 带来数据库庞杂、不便于系统分析和扩展等问题。通过实例对存在问题进行了分析, 详细讨论了通过面向对象编程的思想用异构、迁移、整理等技术实现综合数据库对不同实验模型数据的存储。

关键词: 实验模型; 数据迁移; 存储分析; 面向对象编程

Analysis and Design of the Data Storage of Different Experimental Model

CUI Chao-Yan, ZHANG Jia-Hui

(Dept. of Information Management, Shaanxi Institute of Microbiology, Xi'an 710043, China)

Abstract: Different experimental models generate heterogeneous data in experimental research. It is lead to database cumbersome, inconvenient to data analysis and system expanding. Through specific example, this paper analyzes on the existing problems. According to the thought of OOP is constructed, use technologies about data heterogeneous, data transfer and data cleansing, realise the data storage of different experimental model by integrated database is discussed in detail.

Key words: experimental model; data migration; storage analysis; object-oriented programming

1 引言

科学研究中进行的科学实验, 由于实验模型不一样, 产生的数据结构各异。有二维、三维、甚至多维, 由此导致数据库庞杂冗繁, 给数据分析和系统的扩展带来了许多潜在问题。如何解决数据库数据逻辑清晰结构稳定, 维护系统的可扩展性, 是我们面临的主要问题。面向对象的核心思想就是高内聚低耦合, 通过这一思想, 我们将不同实验模型的表抽象出来, 创建综合数据库。不管是什么结构的实验数据, 都可以抽象为二维表的形式。二维数据, 一行两列代表一个空间点; 三维数据, 一行三列代表一个空间点; 多维数据, 一行多列代表一个空间点。由此, 我们就可以把不同实验模型所产生的实验数据, 利用数据仓库和数据挖掘技术^[1]迁移或录入到一个综合数据库中^[2], 数据的存储分析也变得简单。

2 分析与设计

我们在设计《科研项目管理和实验数据分析统计系统》时, 针对实验模型多样化的特点, 使用综合数据库来存储各种实验模型所产生的数据。将具体实验抽象为实验对象, 在记录实验数据时, 设计一个实验模型, 指定具体的实验名称、实验数据名称、实验数据的值。即可以用一种模型制作一个表示实验结果的对象来对实验数据进行分析统计, 也可以用一种模型产生的数据制作多个对象来对实验数据进行分析统计, 还可以用两种或两种以上模型产生的数据制作一个对象来对实验数据进行分析统计^[3]。

有关实验的信息材料用主题数据库的方式解决, 用可扩展的系统结构满足数据的不断扩充^[4]。而实验数据和数据分析用综合数据库解决。有关实验数据的存储分析如下: 设计出(为了说明方便仅留下能说明关

^① 收稿时间:2012-03-07;收到修改稿时间:2012-05-01

联关系的主要字段)下列二维表及综合数据库关系图^[5]如图 1 所示.

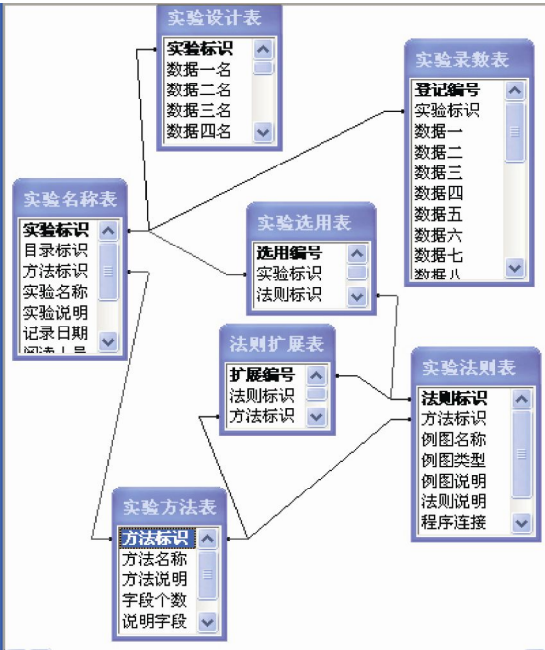


图 1 综合数据库关联图

实验标识	方法标识	实验名称
------	------	------

表 1 实验名称表

登记编号	实验标识	数据 1	数据 2	...	数据 n
------	------	------	------	-----	------

表 2 实验记录表

实验标识	数据 1 名	数据 2 名	...	数据 n 名
------	--------	--------	-----	--------

表 3 实验设计表

选用编号	实验标识	法则标识
------	------	------

表 4 实验选用表

方法标识	方法名称	字段个数
------	------	------

表 5 实验方法表

法则标识	方法标识	例图名称	例图类型	程序连接
------	------	------	------	------

表 6 实验法则表

扩展编号	法则标识	方法标识
------	------	------

表 7 法则扩展表

通过以上 7 个表的数据记载,可达到各种实验模型的实验数据的存储.并用数据中间层^[6]面向对象的系统结构对数据进行维护^[7].统计分析实验数据时,能方便、正确的提取相关数据,显示各种分析统计结果^[8].

实验模型在这里指实验所确定的数据模型及分析方法,即实验所用的可变参数以及实验的结果.例如:某一微生物的生长量实验.微生物的生长和时间、温度、培养基等一些环境变量有关,我们可将一些环境因素作为常量(恒定在某一个值),另一些作为变量(如温度、时间),来考察这些环境变化对这一微生物的生长量的影响.

我们通过三个步骤来确定实验模型:

首先,添加数据到实验方法表(表 5).实验方法表的每行记录记载着一个实验模型,字段个数表示该实验模型的数据结构是几维,也就是有几个实验参数.例如:不同温度、不同时间观察微生物生长量实验,温度、时间和生长量是三个变化中的量.我们就可将这种同类型的实验抽象为一个方法,作为一条记录记载在实验方法表(表 5)中.

其次,确定实验法则表(表 6)中数据.也就是确定实验模型中分析统计的对象.实验方法表(表 5)和实验法则表(表 6)是一对多关系.上述观察微生物生长量实验的分析结果就可以有多种显示效果,如趋势图、直方图、正态分布等.在实验法则表(表 6)中记载三条记录,分别表示三种对象.

最后,法则扩展表(表 7).表示两种以上的数据结构用一个对象分析统计并表示出结果的实验.例如:如果定义了时间考察生长量和温度考察生长量这两个单独的实验数据结构,又需要对这两个实验的数据进行综合分析,我们就可以在法则扩展表(表 7)中记载两条记录,来指定实验法则表(表 6)中某个对象可同时使用实验方法表(表 5)的两种方法的数据.

表 5、表 6、表 7 数据由“分析统计实验模型的数据对象”设计者来维护.

实验记录表(表 2)是综合数据库中基本的实验数据存储表,所有的实验数据都存储在此表中.通过和实验设计表(表 3)多对一的关联,得到该实验数据的每一列的名称.通过和实验名称表(表 1)多对一的关联,得到该实验数据来自哪一个具体实验,而实验名称表(表 1)又通过和实验选用表(表 4)一对多的关联,得到

该实验数据统计分析时所用的实验法则,从而推出使用某个模型。

3 实现

3.1 实验设计

要实现对科学数据进行存储,首先用程序对实验进行实验设计。如图2在“实验设计”下拉菜单中选择“新增实验设计”,如要修改则选择以前进行实验设计时所填的实验名称进行修改(数据来源实验名称表);在“实验名称”输入框中实验者填写实验的实验名称;在“实验说明”输入框中注解该实验;在“实验方法”下拉菜单中选择适用的实验方法(实验方法的数据来源应该是实验方法表中记录);在“待选实验图表”列表中选择实验法则右移到“选入实验图表”列表中(待选实验图表列表中数据来源应该是根据所选的实验方法关联出实验法则表记录)右移时可以多选;定义实验数据字段的名称时在右下数据表中填写相应的名称即可。如新增实验设计,字段列的内容为实验方法表中字段个数值推出字段的个数,名称列的内容为空。若是修改,名称列的内容为根据实验标识取实验设计表的记录。

以上内容填写好后,击“保存”按钮,将生成一个实验标识(新增时)并把用户填写或选择的实验名称、实验说明、方法标识等内容写入实验名称表中;把用户选择的实验法则写入实验选用表中;把用户填写的实验字段名称内容写入实验设计表中。修改时根据实验标识找到相应记录将所填内容改写到相应表中。对已填入实验数据的实验设计不能进行修改。

例如:做某细菌生长量实验。选择“新增实验设计”后;填写实验的实验名称:某细菌不同温度下生长量实验一;填写实验说明(将恒定环境因素及其他有关实验的信息都记载在实验说明里);选择适用的实验方法,因“分析统计实验模型的数据对象”的设计者已将实验方法表(表5)的数据加入,打开下拉菜单将列出所有的实验方法,选择适合此实验的实验方法;实验方法选定后,待选实验图表中就会根据表5、表6、表7的关系列出相应实验法则,“选入实验图表”列表中,选趋势图、直方图、正态分布分析三项;实验方法选定,实验数据结构的字段数量就可以确定,在“实验数据字段名称”列表中第一列就会列出数据一、数据二,我们在第二列填写温度、生长量,点击保存。一个实验的设计就做好了。而所填选的实验名称、实验说明、实

验方法、实验法则、实验数据字段名称将记载在表1、表3、表4中。

通过上述设计就完成一个实验的设计。随后即可进行向实验录数表中录入实验数据或将外部实验数据引入实验录数表中的工作。

图2

3.2 实验录数

做了实验设计以后,就可以对实验录数表(表2)录入数据,首先选择一个实验,也就是通过实验标识确定给哪一个实验录入数据。录入实验数据程序就是通常的增、删、改维护程序。实验录数表(表2)的数据也可以用程序取外部数据。有了数据就可以用以设计好的各种统计分析对象对各种实验数据进行统计分析。

其他外部数据(记载在纸上或电子文档中历史实验数据、其他外部实验处理系统的数据库中的数据)的处理通过异构、迁移、整理等技术将数据转录到综合数据库中^[2]。用户首先做实验设计,转录时由用户确定实验标识,通过特定程序自动转录到(表2)中。

例如:我们选定了“某细菌不同温度下生长量实验一”实验,程序就会调整为只允许录入温度、生长量这两列数据。保存后数据记载在表2中。产生一个登记编号,“某细菌不同温度下生长量实验一”的实验标识,数据一记录温度,数据二记录生长量,其他数据为空的这样一条记录。

3.3 实验结果输出

有了实验数据,就可以用各种方法对实验数据结果输出,可以用 Excel 现有的输出模型无缝整合到系统中。来展现实验数据的各种图表^[9]。从而让用户一目了然的看到实验结果。

例如:需要对“某细菌不同温度下生长量实验一”数据进行正态分布分析来观察温度和生长量是否符合正态分布,只需选择“某细菌不同温度下生长量实验一”,程序会根据用户进行实验设计时所选试验方法及实验法则列出所有可能的分析模型.我们选择“正态分布分析”,程序会根据实验法则表(表 6)的“程序连接”找出处理“正态分布分析”的程序对象进行处理,得出答案并显示出分布图来.

4 总结

通过《科研项目管理与实验系统分析系统》的分析设计,针对科研实验模型多样化的特点,利用建立综合数据库的方法实现对不同实验模型数据的存储分析,解决了由于数据库数据量大,关系复杂等原因给系统带来的分析统计和系统扩展方面的弊端^[10].运用面向对象的思想进行编程,使代码得到复用,提高了编程的灵活性,同时有利于系统的维护和扩展.

参考文献

- 1 夏火松.数据仓库与数据挖掘技术.北京:科学出版社,2009.
- 2 罗林球,孟琦,李晓,苏国平,张澄澈.异构数据库的迁移与实现.计算机应用研究,2006(12):233-235,238.
- 3 O'Docherty M. 面向对象分析与设计.余志翔译.北京:清华大学出版社,2006.
- 4 宗莲松,潘华,肖毅.可扩展式教务信息系统的设计与实现.实验技术与管理,2010(12):128-132.
- 5 王晴,邵冬华,朱敏,王艳红.关系数据库与 SQLServer 教程.北京:中国人民大学出版社,2009.
- 6 李启育.ASP 实现数据访问层 DAL.电脑编程技巧与维护,2010(9):74-76,84.
- 7 郑劲松.一种通用的表结构设计方法.计算机系统应用,2009,18(4):182-183.
- 8 赵玉珍,雷淑玲,龚江宏.显微组织数据库的建立和使用.实验技术与管理,2009(1):73-75.
- 9 张文博,余文芳.ASP.NET 编程中对 Excel 文档操作的探讨及应用.计算机系统应用,2010,19(3):1871-189.
- 10 孙凡.信息系统概念建模研究框架.计算机系统应用,2010,19(6):240-244.

(上接第 161 页)



图 7 菜单信息接收

服务器端的界面设计和客户端的类似,经过测试顾客能够将客户端的信息经过无线路由器传送到管理者的服务器端.图 6、图 7 所示为服务器端的界面,客户端将选定好的菜单发送过来对应的桌号颜色就会变化,单击一下信息就会显示在编辑栏中,并且客户可以发送自己的要求信息给餐厅管理者.

6 结语

很好体现了该设计的特色,综合了点菜、娱乐于一体的点菜系统,采用移植性且界面视觉好的 GTK+ 进行人机界面的开发,较好地适合现在电子发展的时代,GTK+ 使用最基本的语言(C 语言)进行的开发从而能够使更多的人员来开发使用它,该界面汉化显示,操作方便,开发费用低能够节省成本,该系统具有良好的实用价值,具有巨大的开发空间.

参考文献

- 1 孙建梅,曹新.基于 Qt4 的嵌入式点菜系统的设计与实现.微计算机信息,2009,(4):101-108.
- 2 罗贤全,尚朝轩,高勤.用 GTK+ 开发 linux 嵌入式 GUI 应用程序.微计算机信息,2004,20(6):67-68.
- 3 宗笋笋,郝卫东.基于 ARM 的无线点菜系统.计算机系统应用,2011,20(6):221-224.
- 4 杜华.编程技术详解.北京:人民邮电出版社,2007.