

数据库工程设计

钟 平, 续志军

(中国科学院长春光学精密机械与物理研究所, 吉林 长春 130021)

摘要: 为一个大型的数据处理系统建立数据库是一项艰巨而复杂的任务。在数据库的设计中实施软件工程是数据库设计的重要方法。根据软件工程的一般规律和方法并结合数据库系统本身的特点与要求, 分析在数据库工程设计中应考虑的问题, 提出建立数据模式的基本要求, 并依据软件工程生存周期各阶段的基本任务和标准, 提出数据库工程设计的主要内容、方法和步骤。

关键词: 数据库; 软件工程; 设计

中图分类号: TP311.5 文献标识码: A

1 引 言

数据库是用来满足多种类型终端用户需要的相关数据的共享集合。这种数据的存储不依赖使用它的程序, 加入新的数据, 修改和检索现有的数据都得到周密的控制, 数据检索可以由多个用户用不同的方法以适当的个人控制来进行。这种数据的结构为实际应用开发提供了一个扎实的数据基础。建立数据库管理系统的宗旨是减少维护并在需要时能灵活地导出信息。然而数据库管理系统的设计即软件的开发实际是一件非常复杂的事, 没有一套系统的方法, 软件的研制工作可能会陷入困境。利用软件工程化的方法是一种有效的方法, 它指出明确的工作步骤, 也给出了描述软件产品的文档格式, 这为建立高质量数据库系统提供了具体的“求解过程”, 有了这一方法, 开发人员就可以按照方法提供的“求解过程”, 有计划、有步骤、纪律化地开展工作的, 软件开发就可以像制造机器或建造楼房一样“工程化”地进行。从而达到以较少的投资获得易维护、可靠、高效率 and 易理解的软件产品。

2 数据处理系统主要考虑的问题

要建立一个数据库管理系统, 贯穿整个软件开发过程的主要问题是“数据”以及对“数据进行加工”。在分析阶段, 软件开发人员应该充分地分析用户提出的“数据要求”, 包括有哪些数据, 数据本身的属性及性质, 数据之间的关系等。同时, 又

要分析用户提出的“加工要求”, 包括对数据进行哪些加工, 加工的逻辑要求等, 从而从“数据”和“数据经受的加工”这两个相互补充的方面表达一个数据处理系统——数据流图。当然数据流图只描述了系统的“分解”, 它并没有表达出每个数据和加工的具体含义, 这些信息需要在“数据词典”或“说明书”表示出来。数据流图具有直观、容易理解及被开发小组人员同时进行审查的特点, 如图中存在错误, 一般比较明显, 易被人发现, 所以数据的这种描述方法易于软件的检测与维护。

3 数据库系统的组成

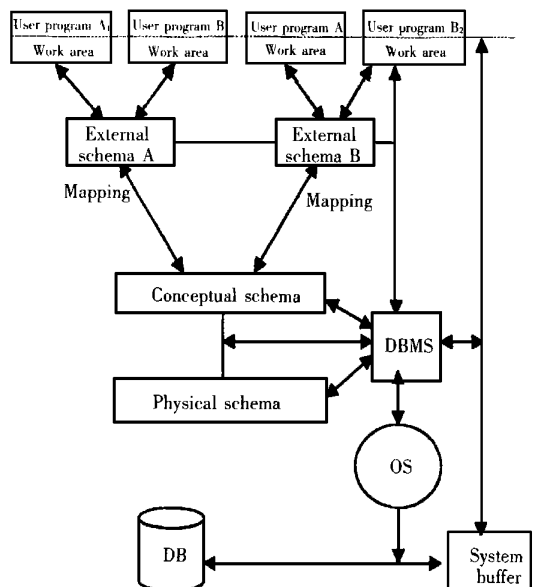


Fig. 1 Database system structure

数据库系统一般来说可由模式、子模式、应用程序数据库和数据库管理系统(DBMS)等组成,如图1所示。

其中DBMS可以从现有的产品中选择,而模式、子模式、应用程序等则必须据用户的共同要求进行分析和设计,即数据库设计。数据库设计的核心问题是如何建立一个数据模式,建立一个数据模式应满足下面几个条件:

(a) 符合用户的要求,即能正确反映用户的现时环境,它应包括用户需要处理的所有数据并能支持用户所需的所有操作。

(b) 能被某个现有的DBMS所确认。

(c) 具有较高的质量,即易于维护,可读性强,效率高等。

4 数据库的工程化设计

软件是程序及开发、使用和维护程序所需的所有文档,为了用工程化的方法研制设计和有效地管理与维护软件的全过程,可引出软件生命期这个概念,软件生命期可以分为以下五个阶段:

(a) 系统分析阶段: 系统分析阶段的主要工作是理解和表达用户的要求,并产生系统说明书和方案书。

(b) 系统设计阶段: 系统设计阶段的任务是建立系统硬件结构,并对系统的输入、输出、文件、数据库、程序模块以及处理算法等作出详细定义和描述,产生各种设计说明书。

(c) 程序设计阶段: 程序设计阶段的任务是编写程序。

(d) 测试阶段: 程序测试阶段的任务是对程序模块进行调试,以产生可工作的应用系统。

(e) 系统运行和维护阶段: 系统运行和维护阶段的任务是根据运行评价以及发现问题,进行维护和改进,并产生改进了的系统提交给用户使用。

从以上我们可以看出,每个阶段都有相应的确定任务,并产生一定规格的文档转交给下一阶段,下一阶段在前一阶段提供的基础上继续开展工作。这样软件的开发过程就不再是无形的思维过程,而是一个可见的、可管理的生产过程了。将软件生命期划分为这样几个阶段,也为有计划的检测及维护打下良好的基础。数据库的设计是指在现有数据库管理系统(DBMS)的基础上建立数据库的过程,数据库的设计尽管它是一个应用课题,但由于它涉及的面广,关系复杂,所以要设计

一个性能较好的数据库也不是一件容易的事。对于采用同一个DBMS,不同的设计者所设计出来的数据库,其性能可能差距较大。笔者认为数据库的设计全过程按软件工程化的观念可分为四个阶段:

(a) 系统分析

(b) 建立概念性数据模式

(c) 逻辑设计

(d) 物理设计

按照相应于软件生命期的观念来看待数据库设计的全过程,前面两个阶段是面向“问题”的,它相应于生命期中的分析阶段,而后面的两个阶段是面向“解答”的,它相应于设计阶段。

第一阶段: 系统分析是整个系统设计的开始,也是为以后的具体设计作准备。系统分析所涉及的方面很多,它包括收集和分析用户的要求,定义系统的内容和范围,确定系统功能和目的,提出系统的环境条件及各种限制因素,制定可行性分析,通过模型计算或预估,比较方案,并作出初步评价等。

第二阶段: 建立概念性的数据模型,是将用户的数据要求明确的表达出来,它相应于软件生命期中建立系统说明书。概念性数据模型是一种面向问题的数据模型,它描述从用户角度看到的数据库,它反映了用户的现时环境,但与数据库的实现无关,它是在用户和设计人员之间起到桥梁作用,因为它一方面是明确地表达用户的要求的模型,而另一方面这个模型又是开发软件进行工程化设计的数据结构的基础。建立概念性数据模型是数据库设计过程的一个关键,其中常用的适合做概念性数据模型的ER(实体联系法)是一个典型的代表。

第三阶段: 逻辑设计就是建立数据模型。由E-R方法表示的概念模型,是用户数据要求的形式化,它独立于任何一种数据模型,同时也不为任何一个具体的DBMS所支持。为了最终建立起用户要求的数据库,就必须把上述概念模型转换为某个具体的DBMS所支持的数据模型,这就是逻辑设计的主要内容。它同软件生命期中设计阶段的总体设计相对应,这一阶段首先要根据用户要求的特点,选合适的数据库管理系统(DBMS),然后据概念性数据模型及所选的DBMS的具体特点,设计出这个管理系统能够接受的数据模型。

第四阶段: 物理设计选择适合于应用环境的物理结构。它进一步设计数据模型的一些物理细

节,如数据文件的存储结构和存取方法,索引文件的建立,如何分配存储空间等问题。然后对该存储模式进行性能评价和反复修改,以便最终得到一个性能较好的数据存储模式。这一阶段可用软件生命期中设计阶段的“详细设计”相对应。这阶段考虑的主要问题是使数据库系统具有较高的效率,所采用的技术与选购的 DBMS 的具体特点密切相关。这里要指出的是有些数据库管理系统在操作系统的支持下,已能自动完成上述工作,对

设计者已经变得透明,需要设计者完成的只剩下对应关系模型中的各关系建立数据库。

最后要指出的是软件的开发过程是复杂的,涉及多方面的因素,软件工程化方法不可能考虑到所有的因素,软件工程化应有一定的灵活性,在实际问题中,我们不能机械地照搬某个方法的步骤或文档格式,而应掌握方法的精神实质,灵活地应用,据具体的情况对方法作适当的修改剪裁。

参考文献:

- [1] 郝忠孝.数据库系统原理[M].吉林:吉林科技出版社,1992.
- [2] 普雷斯曼 R S.软件工程[M].北京:国防工业出版社,1988.
- [3] 张海藩.软件工程导论[M].北京:清华大学出版社,1992.
- [4] 王选.软件设计方法[M].北京:清华大学出版社,1992.
- [5] 潘锦平.软件开发技术[M].上海:上海科学技术文献出版社,1985.

Database engineering design

ZHONG Ping, XU Zhi-jun

(*Changchun Institute of Optics, Fine Mechanics and Physics,*
Chinese Academy of Sciences, Changchun 130021, China)

Abstract: It is a hard and complicated task to build a database for a great data processing system. Executing software engineering in the database design is an important method of database design. According to the common rule and method of software engineering, the problems which should be considered in the database engineering design have been analyzed combined with the characteristics and requirement of database system. And then, the requirements for building data schema have been suggested. The main contents, methods and approaches for database engineering design have also been discussed and put forward based on the basic task and standard of life cycle of software engineering in every phase.

Key words: database; software engineering; design

作者简介:钟平(1963-),男,江西省南康市人,1996年毕业于吉林工业大学计算机及应用专业,并获工学学士学位。现就读于中国科学院光学精密机械与物理研究所研究生部,攻读机械电子工程专业硕士学位。