



数据库系统概述



1.1 认识数据库

■ 项目简介

数据库系统是为适应数据处理的需要而发展起来的一种较为理想的数据处理系统,是一个实际可运行的存储、维护和为应用系统提供数据的软件系统,是存储介质、处理对象和管理系统的集合体。

■ 学习目标

1. 了解数据管理技术的产生和发展。
2. 掌握数据库技术的基本概念。
3. 熟悉数据库系统的组成和特点。
4. 明确数据库系统中的用户角色。
5. 了解数据库系统的体系结构。

■ 任务分析

认识数据库分为以下两个工作任务。

【任务 1】 数据库技术基本概念

学习数据库技术应用之前,需首先明确数据、数据处理、数据管理、数据库以及数据库管理系统几个基本概念,并了解数据管理技术产生和发展的主要阶段及其特点。

【任务 2】 数据库系统简介

数据库系统是存储介质、处理对象和管理系统的集合体。需掌握数据库系统的组成、特点和体系结构,为以后的学习和工作打好基础。

1.1.1 数据库技术基本概念

数据库管理已经从一种专门的计算机应用发展为现代计算环境中的一个重要成分。数据库可以理解为存放数据的仓库,如同水库是存水的。如果程序需要具备对大量数据的存



储、整理、分析能力,这就需要涉及数据库。

数据库的应用非常广泛,被普遍应用于银行、金融、通信、航空、销售、生产制造、会计、人力资源等各行各业。数据库已经成为当今几乎所有企业默认的组成部分。

20世纪60年代后期以来,数据库的使用率在企业中有大幅增长。20世纪90年代末的互联网革命急剧地增加了用户对数据库的应用需求。数据库已经成为当今几乎每个人生活中不可或缺的组成部分。

1. 数据与信息

数据(data)是用来记录信息的可识别符号,是信息的具体表现形式。信息(information)是数据经过加工处理后所获取的有用知识,是对决策产生影响的数据,是对客观世界的认识,即知识。数据处理(data processing)是对数据进行加工的过程,即将数据转换成信息的过程。数据本身没有意义,数据只有对实体行为产生影响时才成为信息。

数据管理(data management)是指对数据进行分类、编码、存储、检索和维护,它是数据处理的基本环节,是任何数据处理业务中必不可少的共有部分。

2. 数据管理技术产生与发展

随着计算机技术的不断发展,在应用需求的推动下,数据管理技术经历了人工管理、文件系统管理、数据库管理三个阶段。每一阶段的发展以数据存储冗余不断减小、数据独立性不断增强、数据操作更加方便和简单为标志,各有特点。

1) 人工管理阶段

20世纪50年代中期,计算机主要用于科学计算。当时没有磁盘等直接存取设备,只有纸带、卡片、磁带等外存;也没有操作系统和管理数据的专门软件。数据处理的方式是批处理。数据的人工管理示意图如图1-1所示。

该阶段管理数据的特点为:数据不保存、应用程序管理数据、数据不共享、数据不具有独立性。



图 1-1 数据的人工管理示意图

2) 文件系统管理阶段

20世纪50年代后期到60年代中期,随着计算机硬件和软件的发展,磁盘、磁鼓等直接存取设备开始普及,这一时期的数据处理系统把计算机中的数据组织成相互独立的被命名的数据文件,可按文件的名字进行访问,并对文件中的记录进行存取。数据可以长期保存在计算机外存上,可以对数据进行反复处理,并支持文件的查询、修改、插入和删除等操作,这就是文件系统。文件系统实现了记录内的结构化,但从文件的整体来看却是无结构的。数据的文件系统管理示意图如图1-2所示。

该阶段管理数据的特点为:数据可以长期保存,由文件系统管理数据,数据冗余度大、共享性差、独立性差,管理和维护的代价大。

3) 数据库管理系统阶段

20世纪60年代后期以来,计算机性能得到进一步提高,更重要的是出现了大容量磁盘,存储容量大大增加且价格下降,使数据能为尽可能多的应用程序服务,这就出现了数据库这样的数据管理技术。数据库的特点是数据不再只针对某一个特定的应用,而是面向全组织,具有整体的结构性,共享性高,冗余度减小,具有一定的程序与数据之间的独立性,并

且对数据进行统一的控制。数据库管理示意图如图 1-3 所示。



图 1-2 数据的文件系统管理示意图

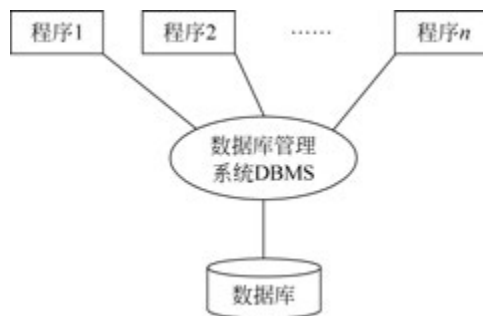


图 1-3 数据库管理示意图

此阶段管理数据的特点有以下 5 个方面：数据结构化，数据共享度高、冗余少且易扩充，数据独立性高，数据由 DBMS 统一管理和控制，DBMS 提供了独立的数据操作界面。

从文件系统到数据库系统，标志着数据管理技术质的飞跃。20 世纪 80 年代后不仅在大、中型计算机上实现并应用了数据管理的数据库技术，如 Oracle、Sybase、Informix、SQL Server 等，在微型计算机上也可使用数据库管理软件，如常见的 Access、FoxPro 等软件，使数据库技术得到广泛应用和普及。

数据库与学科技术的结合会建立一系列新数据库，如分布式数据库、并行数据库、非结构化数据库、知识库、多媒体数据库等，这将是今后数据库技术重要的发展方向。

3. 数据库技术的基本概念

1) 数据

数据是对现实世界的事物采用计算机能够识别、存储和处理的方式进行描述，其具体表现形式可以是数字、文本、图像、音频、视频等。

2) 数据库

数据库(database, DB)是“按照数据结构来组织、存储和管理数据的仓库”，即用来存放数据的仓库。DB 是一个长期存储在计算机内的、有组织的、可共享的、统一管理的大量数据的集合。

3) 数据库管理系统

数据库管理系统(database management system, DBMS)是操纵和管理数据库的软件系统，为用户或应用程序提供访问数据库的方法，包括数据的定义、数据操纵、数据库运行管理及数据库建立与维护等功能。目前市场上比较流行的数据库管理系统主要是 Oracle、MySQL、SQL Server 等产品。

4) 数据库系统

数据库系统(database system, DBS)可以有组织地、动态地存储大量关联数据，方便多用户访问，是由计算机硬件、软件和数据资源组成的系统，即采用数据库技术的计算机系统。通常由软件、数据库和数据管理员组成。

5) 数据库技术

数据库技术是研究、管理和应用数据库的一门软件科学。它通过研究数据库的结构、存储、设计、管理以及应用的基本理论和实现方法，并利用这些理论来实现对数据库中的数据进行处理、分析和理解。



笔记

【技能补强】

请借助网络调查,调研目前常用的数据库管理系统(DBMS)都有哪些产品,它们又是哪家公司生产的,以及这些产品当前最新版本是什么。请找出至少五种产品,越多越好。

1.1.2 数据库系统简介

1. 数据库系统结构

数据库系统结构如图 1-4 所示,下面介绍主要组成部分。



图 1-4 数据库系统结构

1) 数据库

数据库是指长期存储在计算机内的,有组织、可共享的数据的集合。数据库中的数据按一定的数学模型进行组织、描述和存储,具有较小的冗余、较高的数据独立性和易扩展性,并可为各种用户共享。

2) 硬件

硬件是指构成计算机系统的各种物理设备,包括存储所需的外部设备。

3) 软件

软件包括操作系统、数据库管理系统及应用程序。

2. 数据库系统特点

与人工管理和文件系统相比,数据库系统具有显著的特点和优势,主要体现在数据结构化强、共享度高、冗余度低、独立性高,以及数据由 DBMS 统一管理和控制。

3. 数据库的体系结构

数据库的体系结构如图 1-5 所示。数据库的模式是对现实世界的抽象,是对数据库中全部数据的逻辑结构和特征的描述。模式反映的是数据的结果及其联系,数据库在其内部具有三级模式和二级映像。三级模式分别是外模式、模式和内模式,二级映像则是外模式/模式映像、模式/内模式映像。



数据库管理系统简介及功能组成



数据库系统的显著特点和优势



数据库的体系结构

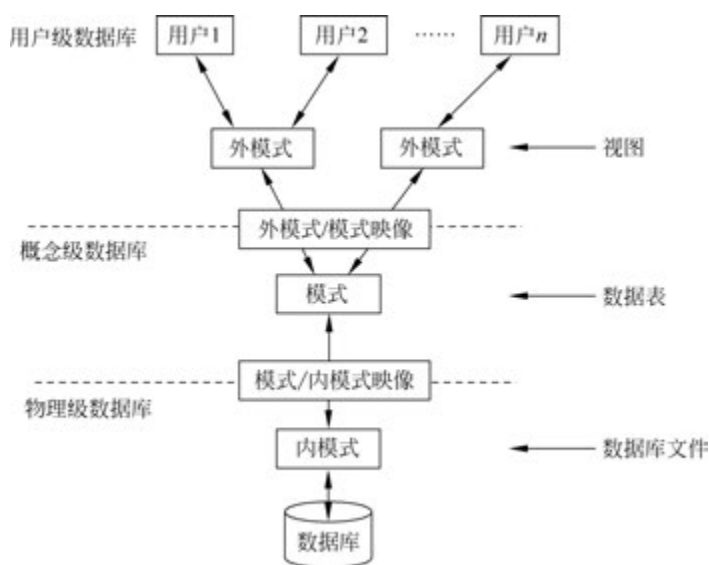


图 1-5 数据库的体系结构

【技能补强】

- (1) 了解数据库系统的体系结构。
- (2) 借助网络,了解 IT 职业岗位和数据库技术的关系,并在各个人才网上搜索计算机、信息技术等职位需求,观察数据库管理、软件工程师等岗位的职责和岗位要求。

1.2 设计“教学质量评价系统”数据模型

■ 项目简介

教学质量评价系统是用于教学评价的信息化平台,利用校园网、数据库和计算机软硬件资源,方便学生、教师和教学管理人员对教师的教学质量进行评价。

本项目将设计“教学质量评价系统”的概念模型和逻辑模型。

■ 学习目标

1. 了解数据模型的定义和分类。
2. 掌握数据库设计的方法与步骤。
3. 能够根据需求分析进行数据库的概念设计。
4. 能够运用关系模型知识将概念模型转换为关系模型。
5. 掌握对关系模型进行关系规范化的方法。

■ 任务分析

【任务 1】 数据模型

数据模型由数据结构、数据操作、数据完整性组成,根据数据结构将数据模型分为层次模型、网状模型、关系模型等。

【任务 2】 数据库设计简介

数据库设计是建立数据库及其应用系统的技术,由于数据库应用系统的复杂性,为了支



笔记

持相关程序运行,数据库设计就变得异常复杂,因此优秀设计不可能一蹴而就,而只能是一种“反复探寻,逐步求精”的过程,通过学习数据库设计的重要性、方法和步骤,掌握规划和结构化数据库中的数据对象以及这些数据对象之间关系。

【任务 3】 设计概念模型

将用户需求抽象为数据库的概念模型,用来描述世界的概念化结构。概念模型包括实体、联系、属性等基本要素,在进行概念设计时一般要依次确定实体、联系、主键、属性这四个步骤完成概念模型的设计。

【任务 4】 建立 E-R 模型

E-R 模型(entity-relationship model)是借助图形描述现实世界的概念模型,提供了表示实体类型、属性和联系的方法,通过 E-R 模型学习,建立起教学质量评价系统数据库的概念模型,完成概念设计。

【任务 5】 关系模型

信息世界的概念模型还不能被数据库管理系统直接使用,需要将概念模型进一步转换为逻辑数据模型,形成便于计算机处理的数据形式。逻辑数据模型是具体的 DBMS 所支持的数据模型,关系数据模型是目前较流行的数据库模型,它用关系表示实体和实体间联系,简单实用。关系模型包括关系数据结构、关系数据操作和关系数据完整性约束。

【任务 6】 建立逻辑模型

通过学习 E-R 模型向关系模型的转换规则,实现教学质量评价系统数据库从概念模型到逻辑模型的转换,完成逻辑设计。

【任务 7】 关系规范化

E-R 模型转换为关系模型时,为了解决在数据库中插入、删除、修改数据时出现异常问题,通过一些规则,使关系达到一定的规范化程度。通过第一范式、第二范式和第三范式,使教学质量评价系统数据库相对规范化。

1.2.1 数据模型

计算机不能直接处理现实世界中的具体事物,需要先将具体事物转换成计算机所能处理的数据,对现实世界数据特征的抽象就形成了数据库的数据模型。数据库系统的核心和基础是数据模型,现有的数据库系统均是基于某种数据模型的。

将客观事物抽象为数据模型,是一个逐步转化的过程,经历了现实世界、信息世界和机器世界这 3 个不同的世界。

1. 数据的 3 个世界

1) 现实世界

客观存在的世界即现实世界。数据描述的是现实世界的事物及其联系,事物都有一些特征或性质,人们总是选择感兴趣、最能表征该事物的基本特征来描述事物。事物既可以是抽象的,也可以是具体的。

2) 信息世界

人对现实世界的事物及其联系在头脑中分析、归纳、抽象形成信息,信息经过记录、整理、归类和格式化后构成了信息世界。经过数据处理,人们把事物的特征和联系通过符号记录下来,并用规范化的语言描述现实世界,从而构成基于现实世界的信息世界。



3) 机器世界

将信息世界中的信息数据化就是机器世界,也称计算机世界。将信息用字符或数字表示,用于计算机识别和处理。

2. 客观对象的抽象过程

客观对象的抽象过程就是建模的过程,将现实世界抽象到机器世界经历了两级抽象和转换。第一步,现实世界到信息世界的抽象,称为概念模型;第二步,信息世界到机器世界的抽象,称为逻辑模型;第三步,存储到计算机中,称为物理模型,如图 1-6 所示。

1) 概念模型

对现实世界的认识和抽象,用来描述世界的概念化结构,不考虑具体在什么计算机和 DBMS 上实现,所以称为概念数据模型,简称概念模型。

概念模型用于信息世界的建模:一方面应该具有较强的语义表达能力,能够方便直接地表达应用中的各种语义知识;另一方面应该简单、清晰,易于用户理解。在概念数据模型中最常用的是 E-R 模型、扩充的 E-R 模型、面向对象模型及谓词模型。

最常用的设计模型就是实体—联系模型。例如,教学质量评价系统中的学生、教师和选课可分别抽象为“学生”“教师”实体以及两个实体之间的“选课”联系,建立的 E-R 概念模型如图 1-7 所示。

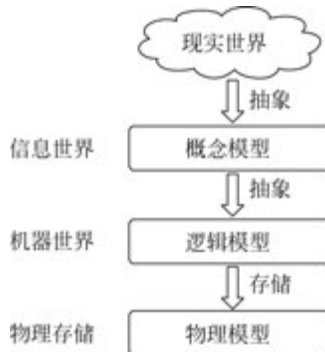


图 1-6 建模过程

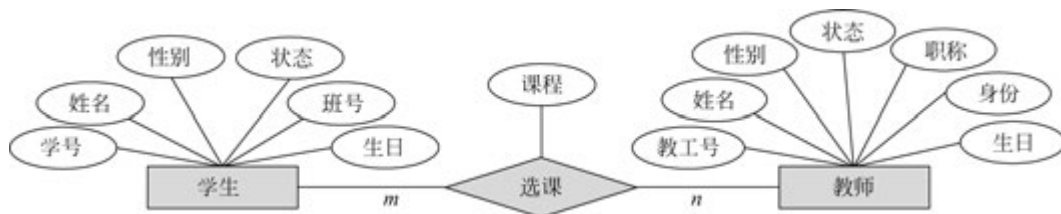


图 1-7 E-R 概念模型

2) 逻辑模型

逻辑模型是指按计算机系统的观点对数据建模,是基于某种(关系、层次、网状)逻辑数据结构的,用于 DBMS 的实现,所以称为逻辑数据模型,简称逻辑模型。

逻辑数据模型是具体的 DBMS 所支持的数据模型,主要有关系数据模型、层次数据模型和网状数据模型。关系数据模型是目前较流行的数据库模型,支持关系数据模型的数据库管理系统称为关系数据库管理系统。关系数据模型以二维表结构来表示事物与事物之间的联系,也可以称为实体与实体之间的联系。

例如,选课的关系逻辑模型如图 1-8 所示。

3) 物理模型

物理模型是指面向计算机物理表示的模型,描述了数据在存储介质上的组织结构,它不但与具体的 DBMS 有关,而且与操作系统和硬件有关,所以称为物理数据模型,简称物理模型。



笔记

TeacherCode	StudentCode	CourseName	AddDate
2015261621	31821160420	普通话	2025/4/21 14:06
2015261621	31821160423	普通话	2025/4/21 14:06
2015261621	31821160426	普通话	2025/4/21 14:06
2015261621	31821160429	普通话	2025/4/21 14:06
2015261621	31821160433	普通话	2025/4/21 14:06
2015261621	31821160447	普通话	2025/4/21 14:06
2017261651	31821160403	职业沟通讲座	2025/4/21 14:06
2017261651	31821160405	职业沟通讲座	2025/4/21 14:06
2017261651	31821160408	职业沟通讲座	2025/4/21 14:06
2017261651	31821160410	职业沟通讲座	2025/4/21 14:06

图 1-8 关系逻辑模型

3. 组成部分及联系

数据模型是严格定义的一组概念的结合,这些概念精确地描述了系统的静态特性、动态特性和完整性约束条件。因此,数据模型所描述的内容包括 3 个部分:数据结构、数据操作、数据完整性。

4. 重要模型

逻辑数据类型有 3 种基本的数据模型,分别是层次模型、网状模型和关系模型,分别如图 1-9~图 1-11 所示。这 3 种模型是按数据结构而命名的,前两种采用格式化的结构,关系模型为非格式化的结构,用单一的二维表的结构表示实体及实体之间的联系。

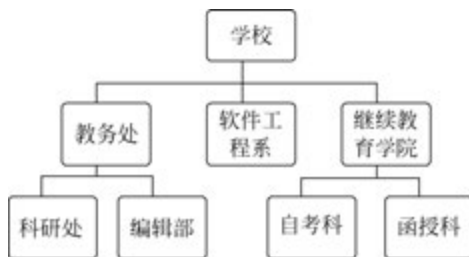


图 1-9 层次模型示例

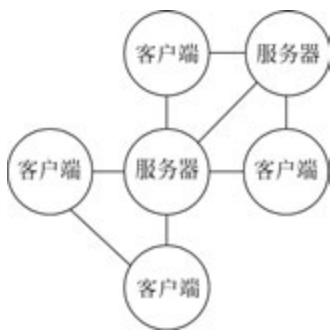


图 1-10 网状模型示例

学号	姓名	性别	状态
31821160401	巴雷静	女	正常
31821160402	毕晓帅	男	正常
31821160403	曹盛堂	男	正常
31821160404	桑晓涛	男	休学
31821160405	陈亚辉	男	正常
...			

图 1-11 关系模型示例

【技能补强】

了解关系模型的发展史。

1.2.2 数据库设计简介

数据库设计是指根据用户的需求,在某一具体的数据库管理系统上,设计数据库的结构



重要模型



和建立数据库的过程,目的是管理大量信息。

数据库设计为一个给定的应用环境,构造最优的数据库模式,建立数据库及其应用系统,使之能够有效地存储数据,满足各种用户的应用需求(信息要求和处理要求)。在数据库领域内,常常把使用数据库的各类系统统称为数据库应用系统。

数据库设计是建立数据库及其应用系统的技术,是信息系统开发和建设中的核心技术。

1. 数据库设计的重要性

数据库设计的重要性主要体现在有利于资源节约、有利于软件运行速度的提高和有利于软件故障的减少这3个方面。

2. 数据库设计的方法

1) 手工与经验相结合的方法

由于缺乏科学理论依据和工程方法的支持,在相当长的一段时期内数据库设计都依赖于设计人员的经验和水平,数据库设计成为一种艺术而不是工程技术。

2) 规范设计方法

随着软件工程理论的发展,设计人员总结出了运用软件工程思想设计数据库的方法,这些方法制定了各种设计准则和规程,因此被称为规范设计法。著名的新奥尔良(New Orleans)方法将数据库设计分为若干阶段,包括需求分析阶段、概念设计阶段、逻辑设计阶段、物理设计阶段。

此外还有基于E-R模型的概念设计方法、基于3NF(第三范式)的逻辑设计方法、基于抽象语法规则的设计方法等。

3. 数据库设计的步骤

按照规范设计的方法,同时考虑数据库及其应用系统开发的全过程,可以将数据库设计分为以下6个阶段:需求分析阶段、概念设计阶段、逻辑设计阶段、物理设计阶段、数据库实施阶段和数据库运行和维护阶段。

【技能补强】

了解数据库设计的注意事项。

1.2.3 设计概念模型

数据库系统的主要应用方面为管理信息系统。管理信息系统的开发类型一般有以下3种。

- 软件产品实施。通过软件参数设置对软件做少量的功能调整。
- 在开发型平台上研发。按用户需求来设计开发,组建研发团队研发合适的管理信息系统。
- 在应用设计平台上开发。按照用户需求通过二次开发进行个性化设计,可应对管理需求的变化,动态调整业务应用和管理流程。

本小节以涵盖数据库全部学习内容的“教学质量评价系统”数据库设计为案例,循序渐进地介绍数据库系统开发与维护所需的数据库知识与技术。

1. 需求分析

数据库设计的第一步是需求分析,重点是调查、收集与分析用户数据处理中的数据需



数据库设计的步骤



笔记

求、功能需求、完整性与安全性需求,应用分析方法与工具,经过用户的反复确认,形成需求分析报告。

1) 数据流图

数据流图(data flow diagram,DFD)是结构化分析方法中常用的工具,它以图形的方式描绘数据在系统中流动和处理的过程。在结构化开发方法中,数据流图是需求分析阶段产生的结果。

数据流图从数据传递和加工的角度,以图形的方式刻画数据流从输入到输出的移动变换过程。数据流图的基本图形元素如表 1-1 所示。

表 1-1 数据流图的基本图形元素

符 号	名 称	说 明
	数据源点或终点	软件系统外部环境中的实体(包括人员、组织或其他软件系统),一般只出现在数据流图的顶层图中
	数据处理	数据处理是对数据进行加工的单元,它接收一定的数据输入,对其进行处理并产生输出
	数据存储	又称数据文件,指临时保存的数据,它可以是数据库文件或任何形式的数据组织
	数据流	特定数据的流动方向,是数据在系统内传播的路径

2) 用例图

用例图(use case diagram)是由软件需求分析到最终实现的第一步,是用户与系统交互的最简表示形式,展现了用户及与其相关的用例之间的关系,它描述人们如何使用一个系统。用例图也经常和其他图表配合使用。

用例图包含 6 个元素,分别是参与者(actor)、用例(use case)、关联关系(association)、包含关系(include)、扩展关系(extend)以及泛化关系(generalization)。用例图基本图形元素如图 1-12 所示,参与者用人形图标来标识,用例用椭圆来表示,连线表示它们之间的关系。

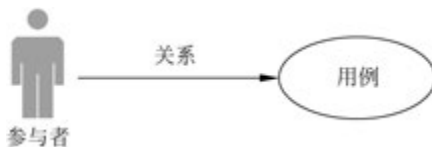


图 1-12 用例图基本图形元素

3) 数据字典

数据字典(data dictionary,DD)最重要的作用是作为分析阶段的工具,任何字典最重要的用途都是供人对不了解条目的解释进行查询。在结构化分析中,数据字典的作用是给数据流图上每个成分加以定义和说明。

数据字典用于描述系统中数据处理的数据需求。DD 是关于数据的信息集合,也就是对 DFD 中包含的所有元素定义的集合。值得注意的是,DD 是关于数据定义的描述,即元数据,而不是数据本身。数据字典通常包括数据项、数据结构、数据流、数据存储和处理过程 5 个部分。



用例图
的用法