

高等院校计算机应用系列教材

Access 2016

数据库应用教程

彭毅弘 程 丽 主 编
刘永芳 李盼盼 副主编

清华大学出版社
北 京

内 容 简 介

本书以 Microsoft Access 2016 中文版为实践平台,以小型超市管理系统的创建、管理和应用为主线,通过大量实例介绍了数据库的基本概念、原理和开发技术。全书共分为 9 章,主要包括数据库技术基础、数据库和表、查询、结构化查询语言(SQL)、窗体、报表、宏、VBA 程序设计和 VBA 数据库访问技术等内容。

本书注重案例的实际应用价值,强调理论与实践结合,同时还引入了程序流程介绍,注重提升学生的逻辑分析能力。本书配套的《Access 2016 数据库应用教程实验指导》(ISBN 978-7-302-60854-7),为各章提供多种类型的实验案例,每个案例都附有操作步骤。

本书既可作为高等学校数据库基础与应用教材,又可作为全国计算机等级考试(二级 Access)的自学和培训教材,还可供从事数据库应用、设计和开发的技术人员参考。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。举报:010-62782989, beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

Access 2016 数据库应用教程 / 彭毅弘, 程丽主编. —北京: 清华大学出版社, 2022.8
高等院校计算机应用系列教材
ISBN 978-7-302-60883-7

I. ①A… II. ①彭… ②程… III. ①关系数据库系统—高等学校—教材 IV. ①TP311.132.3

中国版本图书馆 CIP 数据核字(2022)第 083114 号

责任编辑:王 定

封面设计:高娟妮

版式设计:孔祥峰

责任校对:马遥遥

责任印制:杨 艳

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-83470000 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:北京嘉实印刷有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:17.75 字 数:455 千字

版 次:2022 年 8 月第 1 版 印 次:2022 年 8 月第 1 次印刷

定 价:59.80 元

产品编号:097107-01

前 言

随着大数据时代的到来，数据已经成为极具价值的资源，如何分析、管理和利用这些数据对于很多企业来说是未来赢得竞争的关键。数据库技术研究和解决了大量数据如何有效地组织和存储的问题，实现了数据共享，能高效地检索和处理数据，并保障数据安全。

数据库基础知识是当今大学生信息素养的重要组成部分，数据库应用课程是高等学校一门重要的计算机基础课程，不仅普及数据管理知识和数据库操作技术，还涉及面向对象的编程基础。在物联网和人工智能技术迅速发展，程序设计能力培养已经深入基础教育的今天，数据库应用课程中的程序开发技能显得尤为重要。

本书共分为9章，主要内容有数据库技术基础、数据库和表、查询、结构化查询语言(SQL)、窗体、报表、宏、VBA程序设计和VBA数据库访问技术等，通过本书的学习可对Access数据库系统设计有一个清晰完整的认识。

本书具有如下特点：

(1) 案例更具应用价值。选用具有实际使用价值的、生活中普遍接触的、知识点全覆盖的数据库系统实例作为全书的主线，方便读者快速理解知识背景及掌握数据库开发技能。

(2) 强调理论与实践结合。注重将数据库中各种复杂抽象的原理实例化，通过大量的图例、实例、程序片段等手段，将抽象的数据库运行机制以易于理解的形式呈现。

(3) 引入程序流程介绍，培养系统化设计思维。本书从程序设计的本质出发，引入流程介绍，强调系统化分析方法，从整体到局部逐步细化问题，引导开发者使用程序语句来解决实际问题，进而提升逻辑分析能力。

(4) 配套实验教程，为各章提供多种类型的实验案例，每个案例都附有操作步骤。

(5) 适用对象广泛。本书既可作为高等学校数据库基础与应用教材，也可作为全国计算机等级考试(二级 Access)的自学和培训教材，还可供从事数据库应用、设计和开发的技术人员参考。

本书由彭毅弘组织统稿、主编和定稿，编写具体分工如下：彭毅弘编写第2~5章、第7章和第9章，程丽编写第6章，刘永芬编写第1章，李盼盼编写第8章。感谢林燕、赖晓燕、程铃钊、陈心瑜、赵浩等各位老师和专家的支持及对本书提出的宝贵意见。同时，感谢清华大学出版社在本书策划、编辑和出版过程中给予的大力支持。

限于作者水平，书中难免存在疏漏或不妥之处，恳请广大读者批评指正。

本书提供丰富的配套资源，包括教学课件、教学大纲、混合式教学设计、电子教案、数据库文件及资源、思考与练习参考答案以及微课视频，下载地址如下：



教学课件



教学大纲



混合式教学设计



电子教案



数据库文件及资源



思考与练习参考答案

编者
2022年4月

目 录

第 1 章 数据库技术基础	1
1.1 数据管理技术.....	3
1.1.1 数据与数据处理.....	3
1.1.2 数据管理技术的发展.....	3
1.2 数据库系统.....	6
1.2.1 数据库系统的组成.....	6
1.2.2 数据库系统的特点.....	8
1.2.3 数据库系统的三级模式.....	9
1.3 数据模型.....	9
1.3.1 数据抽象过程.....	10
1.3.2 概念模型(E-R模型).....	10
1.3.3 常见的数据模型.....	13
1.4 关系数据库.....	14
1.4.1 关系的术语.....	14
1.4.2 关系的运算.....	15
1.4.3 关系的完整性规则.....	16
1.4.4 从概念模型到关系模型的转换.....	17
1.5 数据库应用系统开发流程.....	18
1.6 Access简介.....	19
1.6.1 Access的特点.....	19
1.6.2 Access数据库的系统结构.....	20
1.6.3 Access 2016主界面.....	23
1.7 思考与练习.....	25
1.7.1 思考题.....	25
1.7.2 选择题.....	25
第 2 章 数据库和表	27
2.1 创建和操作Access数据库.....	29
2.1.1 创建空白数据库.....	29
2.1.2 利用模板创建数据库.....	30
2.1.3 打开数据库.....	30
2.1.4 保存数据库.....	31
2.1.5 关闭数据库.....	32
2.2 表的组成.....	32
2.2.1 表名.....	32
2.2.2 表结构.....	33
2.2.3 表数据.....	33
2.3 表的创建.....	34
2.3.1 使用数据表视图创建表.....	34
2.3.2 使用设计视图创建表.....	35
2.3.3 通过导入外部数据创建表.....	36
2.3.4 表视图的切换.....	38
2.4 表的字段.....	40
2.4.1 字段的命名规则.....	40
2.4.2 字段的数据类型.....	40
2.4.3 字段的属性.....	47
2.5 表的主键.....	55
2.5.1 主键的作用.....	55
2.5.2 设置主键.....	55
2.5.3 删除主键.....	57
2.6 表的编辑与操作.....	57
2.6.1 编辑表的字段.....	57
2.6.2 操作表的记录.....	58
2.6.3 表的复制、删除和导出.....	61
2.7 表间关系.....	63
2.7.1 创建表间关系.....	63
2.7.2 插入子数据表.....	66
2.7.3 关系的参照完整性.....	67
2.8 表的规范.....	70
2.8.1 第一范式(1NF).....	70

2.8.2	第二范式(2NF)	71	4.2	SQL查询的创建	108
2.8.3	第三范式(3NF)	72	4.3	SQL数据查询语句	109
2.9	思考与练习	73	4.3.1	SQL查询语句的一般格式	109
2.9.1	思考题	73	4.3.2	单表查询	110
2.9.2	选择题	73	4.3.3	多表查询	114
			4.3.4	分组查询	115
第3章	查询	75	4.4	SQL数据操作语句	116
3.1	查询概述	76	4.4.1	在表中插入记录	116
3.1.1	查询的作用	76	4.4.2	在表中更新记录	117
3.1.2	查询的类型	76	4.4.3	在表中删除记录	117
3.1.3	查询的视图	77	4.5	SQL数据定义语句	118
3.1.4	查询的创建方法	77	4.5.1	创建表	118
3.2	选择查询	80	4.5.2	修改表结构	119
3.2.1	创建不带条件的查询	80	4.5.3	删除表	121
3.2.2	创建带条件的查询	81	4.6	思考与练习	121
3.2.3	查询条件的使用	82	4.6.1	思考题	121
3.2.4	在查询中使用计算	85	4.6.2	选择题	121
3.3	参数查询	89	第5章	窗体	124
3.3.1	单参数查询	89	5.1	窗体概述	125
3.3.2	多参数查询	90	5.1.1	窗体的功能	125
3.4	操作查询	90	5.1.2	窗体的类型	125
3.4.1	生成表查询	91	5.1.3	窗体的视图	127
3.4.2	追加查询	92	5.1.4	窗体的组成	128
3.4.3	更新查询	94	5.2	创建窗体	128
3.4.4	删除查询	95	5.2.1	使用“窗体”工具创建窗体	129
3.5	交叉表查询	96	5.2.2	使用“空白窗体”工具创建窗体	130
3.5.1	通过向导方式创建	97	5.2.3	使用“窗体向导”工具创建窗体	131
3.5.2	通过设计视图方式创建	99	5.2.4	使用“多个项目”工具创建窗体	132
3.6	重复项查询	100	5.2.5	使用“数据表”工具创建窗体	133
3.7	不匹配项查询	101	5.2.6	使用“分割窗体”工具创建窗体	133
3.8	思考与练习	103	5.2.7	使用“窗体设计”工具创建窗体	134
3.8.1	思考题	103	5.3	为窗体添加控件	134
3.8.2	选择题	103	5.3.1	认识控件	135
第4章	结构化查询语言(SQL)	105	5.3.2	使用文本框(Text)控件	136
4.1	SQL概述	106	5.3.3	使用标签(Label)控件	140
4.1.1	SQL的特点	106	5.3.4	使用按钮(Command)控件	141
4.1.2	SQL的功能	106	5.3.5	使用选项卡控件	143
4.1.3	SQL查询语句和Access查询文件的 关系	107	5.3.6	使用子窗体/子报表(Child)控件	145

5.3.7	使用列表框(List)控件	146	6.5.1	思考题	182
5.3.8	使用组合框(Combo)控件	148	6.5.2	选择题	183
5.3.9	使用选项组(Frame)控件	150	第7章 宏		184
5.3.10	使用复选框(Check)控件	151	7.1	宏的概述	185
5.3.11	使用选项按钮(Option)控件	152	7.1.1	宏的基本概念	185
5.4	设置窗体和控件的属性	152	7.1.2	宏的设计视图	185
5.4.1	“格式”选项	153	7.1.3	常用的宏操作	187
5.4.2	“数据”选项	154	7.2	宏的创建	188
5.4.3	“事件”选项	156	7.2.1	创建独立宏: 顺序操作宏	188
5.5	美化窗体	158	7.2.2	创建独立宏: 宏组	189
5.5.1	窗体的主题应用	158	7.2.3	创建独立宏: 条件宏	190
5.5.2	控件的布局调整	159	7.2.4	创建嵌入宏	193
5.6	窗体的高级设计	160	7.2.5	创建数据宏	195
5.6.1	设置启动窗体	160	7.2.6	创建自动运行宏	195
5.6.2	设置导航窗体	161	7.3	宏的运行与调试	197
5.7	思考与练习	162	7.3.1	宏的运行	197
5.7.1	思考题	162	7.3.2	宏的调试	198
5.7.2	选择题	162	7.4	思考与练习	199
第6章 报表		164	7.4.1	思考题	199
6.1	报表概述	165	7.4.2	选择题	199
6.1.1	报表的功能	165	第8章 VBA 程序设计		201
6.1.2	报表的类型	165	8.1	VBA与宏	202
6.1.3	报表的视图	166	8.2	VBA的容器: 模块	202
6.1.4	报表的组成	166	8.2.1	模块的类型	202
6.1.5	报表与窗体的异同	167	8.2.2	模块的创建	203
6.2	创建报表	167	8.2.3	模块的组成	204
6.2.1	使用“报表”工具创建报表	167	8.3	VBA的编辑器: VBE	207
6.2.2	使用“空报表”工具创建报表	168	8.4	VBA的编程思想: 面向对象	210
6.2.3	使用“报表向导”工具创建报表	169	8.4.1	对象	210
6.2.4	使用“标签”工具创建报表	170	8.4.2	DoCmd对象	212
6.2.5	使用“报表设计”工具创建报表	173	8.5	VBA的编程基础	213
6.3	在报表中使用分组、排序和 汇总功能	176	8.5.1	数据类型	213
6.4	在报表中使用计算控件	179	8.5.2	常量	214
6.4.1	添加日期和时间	179	8.5.3	变量	216
6.4.2	添加页码	179	8.5.4	数组	219
6.4.3	报表中常用的聚合函数	180	8.5.5	运算符和表达式	221
6.5	思考与练习	182	8.5.6	函数	224
			8.5.7	程序语句	229

8.6 VBA的流程控制语句	230
8.6.1 顺序结构	231
8.6.2 选择结构	232
8.6.3 循环结构	236
8.7 过程	240
8.7.1 过程的定义与调用	240
8.7.2 过程的作用范围	240
8.7.3 参数传递	240
8.8 思考与练习	242
8.8.1 思考题	242
8.8.2 选择题	242
第9章 VBA数据库访问技术	245
9.1 常用的数据库访问接口技术	246
9.2 数据访问接口ADO	247
9.2.1 Connection对象	248
9.2.2 Recordset对象	250
9.2.3 Command对象	255
9.3 ADO编程实例	256
9.4 思考与练习	259
9.4.1 思考题	259
9.4.2 选择题	259
附录	261
附录A “小型超市管理系统”数据库的表结构及其记录	261
附录B 常用的ASCII字符集	265
附录C 窗体和控件的常用属性及方法	266
附录D 全国计算机等级考试二级Access考试大纲	270

第 1 章

数据库技术基础

在当今信息社会，电子商务和社交网络已然全面普及，数据资源从各式各样的终端不断地涌现，数据渗透到每一个行业和职能领域，成为重要的生产因素。对海量数据的挖掘和应用，改变了人类原有的生活方式和发展模式，也改变了人类认识世界和价值判断的方式。所有与数据信息有关的业务及应用系统都需要数据库技术的支持。数据库技术是 20 世纪 60 年代后期发展起来的一项重要技术，主要研究如何安全高效地管理大量、持久且可共享的数据，使数据释放更多的价值。数据库知识是当今大学生信息素养的重要组成部分，是各类计算机信息系统的核心技术和重要基础，更是大数据时代的支撑技术之一。因此，了解数据库原理与掌握数据库应用和开发技术，对于科学地组织和存储数据，高效地获取和分析数据，从而充分地利用信息资源是十分重要的。本章主要介绍数据库、数据库系统、数据模型等基础理论知识，为后面各章的学习打下基础。

本章要点

- 数据和信息的概念与区别
- 数据库系统的概念、组成和特点
- 数据模型的概念，E-R 模型的使用
- 关系数据库的基础知识
- 数据库应用系统的开发流程

本章知识结构如图 1-1 所示。

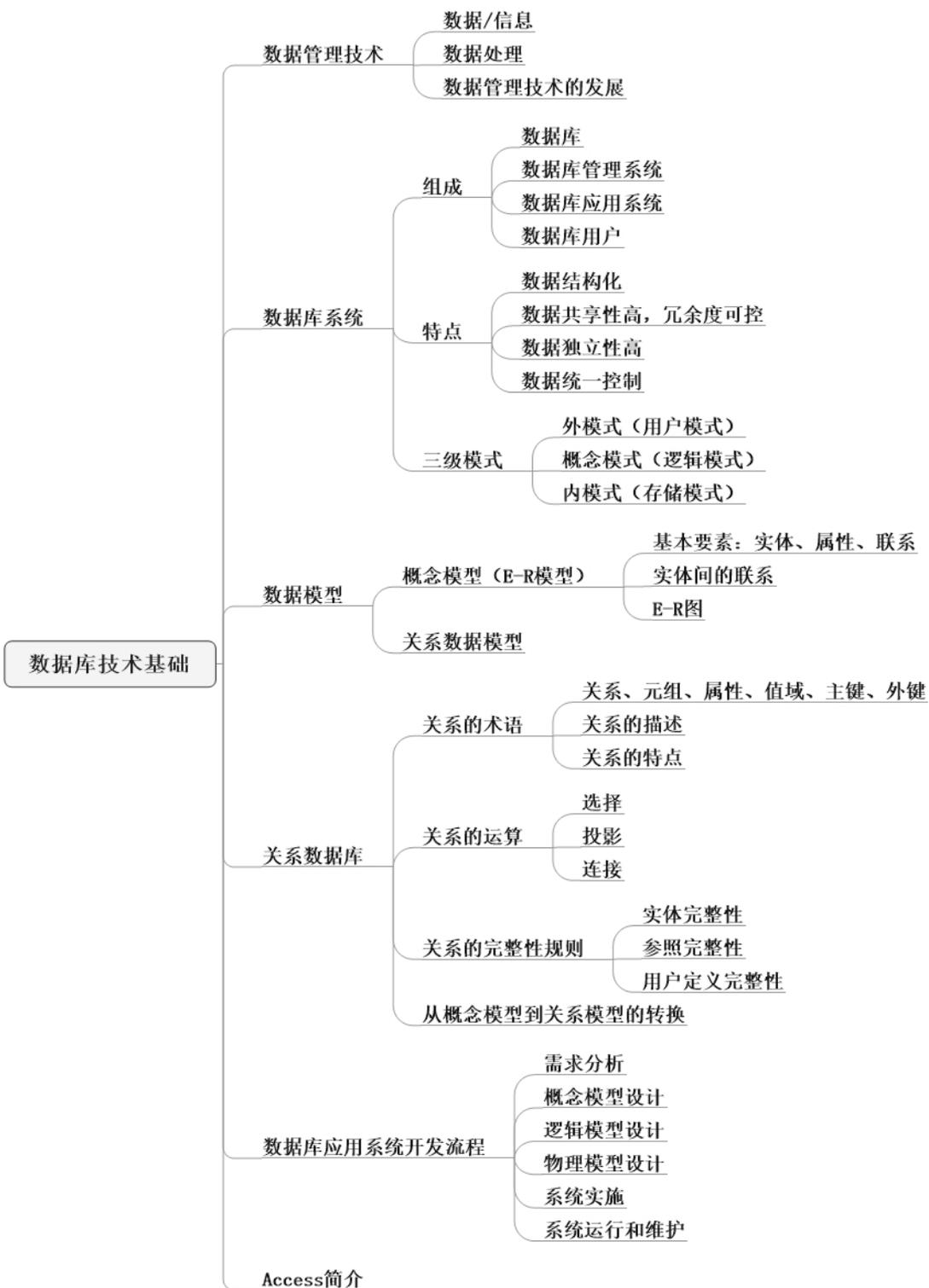


图 1-1 本章知识结构图

1.1 数据管理技术



视频 1-1 数据管理技术

数据库技术是管理数据的一种科学技术方法，专门研究如何组织和存储数据，如何高效地获取和处理数据，从而为人类生活的方方面面提供数据服务。

1.1.1 数据与数据处理

1. 数据和信息

数据(Data)和信息(Information)是数据处理中的两个基本概念，数据是信息的载体，但并非任何数据都能成为信息，只有经过加工处理后的数据才能成为信息。

(1) 数据。数据是人们用于记录事物情况的物理符号。为了描述客观事物而用到的数字、字符，以及所有能输入到计算机中并能被计算机处理的符号都可以看成数据。例如，张小明的年龄是 20 岁，籍贯福建，这里的“张小明”“20”“福建”就是数据。在实际应用中，数据可分为三种：第一种是可以参与数值运算的数值型数据，如年龄、成绩、价格等数据；第二种是由字符组成的、不能参与数值运算的字符型数据，如姓名、籍贯、性别等数据。第三种是图形、图像、声音等多媒体数据，如照片、歌曲、视频等数据。

(2) 信息。信息是经过加工处理并对人类社会实践和生产活动产生决策影响的数据。不经过加工处理的数据只是一种原始材料，它的价值只在于记录了客观世界的事实，对人类活动产生不了决策作用。只有经过提炼和加工，原始数据才会发生质的变化，给人们以新的知识和智慧。例如，收到一条淘宝通知“双 11 活动，商品全场 5 折”，根据这条通知获取数据“5 折”，然后根据商品原价格计算出打折后的商品价格，新的价格数据就是利用原始数据经过加工处理后得到的信息，这个信息可作为是否购买商品的依据。用户还可以进一步利用这个信息与前一年的数据进行比较分析，得到商品的价格走势、打折力度等有价值的信息。

2. 数据处理

数据处理是指将数据转换成信息的过程，包括对数据的采集、存储、分类、排序、检索、维护、计算、加工、统计和传输等一系列操作。其主要目的是从大量的、杂乱无章的、难以理解的数据中，通过分析、归纳、推理等科学方法，利用计算机技术、数据库技术等技术手段，提取出有价值的、有意义的信息，从而作为决策的依据。如图 1-2 所示，使用计算机可以实现数据处理的自动化，完成各种数据处理任务。

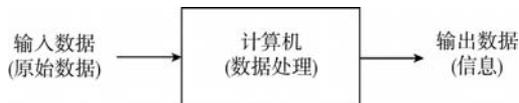


图 1-2 计算机数据处理

1.1.2 数据管理技术的发展

数据管理是指数据的收集、组织、存储、检索和维护等操作，这些操作是数据处理的中心环节，其主要目的是实现数据共享，降低数据冗余，提高数据的独立性、安全性和完整性，从而能更加有效地管理和使用数据资源。计算机技术的发展促使数据管理技术得到了很大发展，

计算机数据管理技术经历了人工管理、文件管理和数据库管理三个发展阶段。

1. 人工管理阶段

20 世纪 50 年代中期以前, 计算机主要用于科学计算。在硬件方面, 外存储器只有磁带、卡片和纸带等, 没有磁盘等直接存取的外存储器; 在软件方面, 只有汇编语言, 没有操作系统。当时的数据管理是以人工管理方式进行的, 有以下特点: 数据量少, 数据不需要长期保存; 没有专门对数据进行管理的软件, 数据由应用程序自行管理, 每个应用程序都要设计数据的存储结构和输入输出方法; 数据无法实现共享, 不同应用程序之间存在大量的重复数据; 数据对应用程序不具有独立性, 进一步加重了程序设计的负担。

以一个公司的信息管理为例, 在人工管理阶段, 应用程序和数据之间的关系如图 1-3 所示。图中不同应用程序产生各类数据, 并产生许多重复的数据, 例如工资数据包含部分员工数据。

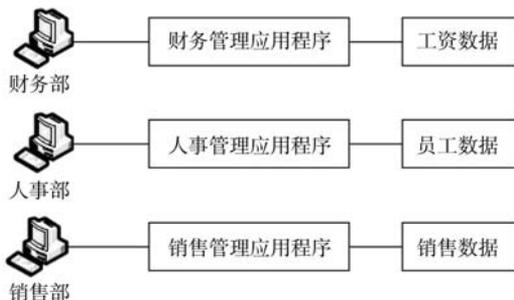


图 1-3 人工管理阶段应用程序和数据之间的关系

2. 文件管理阶段

20 世纪 50 年代后期至 20 世纪 60 年代中期, 计算机开始大量用于数据管理。在硬件方面, 出现了外存, 如磁盘、磁鼓等; 在软件方面, 出现了高级语言和操作系统, 应用程序利用操作系统的文件管理功能可实现数据的文件管理方式。文件管理阶段有以下特点: 数据可以组织成文件, 能够长期保存和反复使用; 数据和应用程序之间有一定的独立性, 通过文件系统把数据组织成一个独立的数据文件, 大大减少了应用程序维护的工作量; 不同应用程序的数据不能共享, 数据独立性差, 冗余度大。

在文件管理阶段, 公司信息管理中应用程序和数据文件之间的关系如图 1-4 所示。各应用程序通过文件系统对相应的数据文件进行存取和处理, 但各数据文件之间是孤立的, 缺乏对数据统一管理和控制的能力。例如, 因某个员工离职而在员工数据文件中删除了其数据, 但无法在工资数据文件中自动删除该员工的相关数据。

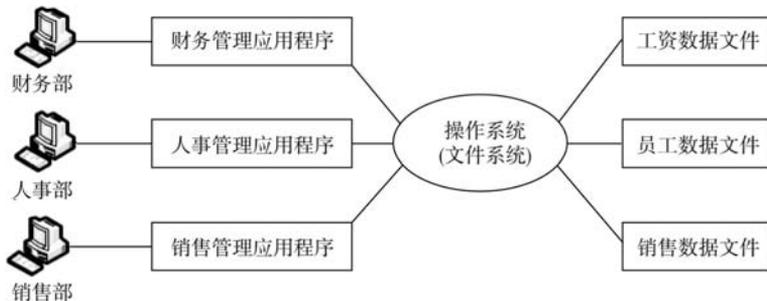


图 1-4 文件管理阶段应用程序和数据文件之间的关系

3. 数据库管理阶段

20 世纪 60 年代后期, 数据量急剧增加, 对数据共享性要求更加强烈。同时, 计算机硬件价格下降, 而编写和维护软件的成本相对增加, 文件系统已经无法满足多应用、多用户的数据共享需求, 于是出现了统一管理数据的数据库管理系统(Database Management System, DBMS)。

数据库管理系统把所有应用程序中使用的数据整合起来, 按统一的数据模型存储在数据库中, 提供给各个应用程序使用。数据与应用程序之间完全独立, 数据具有完整性、一致性和安全性等特点, 并具有充分的共享性, 有效地减少了数据冗余。

在数据库管理阶段, 企业信息管理中应用程序和数据库之间的关系如图 1-5 所示。企业信息管理的相关数据都存放在数据库中, 数据库面向整个应用系统, 实现了数据共享, 并使得数据和应用程序之间保持较高的独立性。

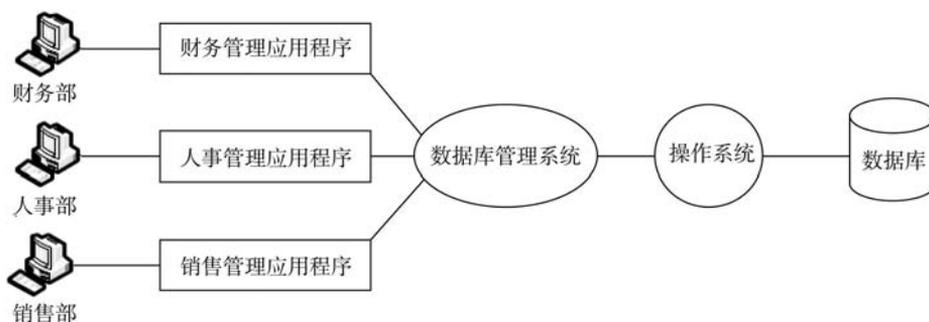


图 1-5 数据库管理阶段应用程序和数据库之间的关系

4. 数据库管理技术的新发展

数据库技术的发展先后经历了第一代数据库系统(层次数据库和网状数据库)和第二代数据库系统(关系数据库)。自 20 世纪 70 年代使用关系数据库后, 数据库技术得到了蓬勃发展, 但随着新需求的不断提出, 占主导地位的关系数据库系统已不能满足新的应用领域的需求。例如, 在实际应用中, 需要存储并检索多媒体数据、计算机辅助设计绘制的工程图纸、地理信息系统提供的空间数据和各种复合数据(如集合、数组、结构)等, 关系数据库无法实现对复杂数据的管理, 因此出现了许多不同类型的新型数据管理技术。下面对这些技术进行简要介绍。

(1) 分布式数据库系统。分布式数据库系统是数据库技术与计算机网络技术、分布式处理技术相结合的产物。一个分布式数据库在逻辑上是一个统一的整体, 在物理上则分别存储在不同的物理节点上。分布式数据库系统的特点主要有: 数据库中的数据分布在计算机网络的不同物理节点上; 分布在不同节点的数据在逻辑上属于同一个数据库系统, 数据间存在相互关联; 每个节点都有自己的计算机软硬件资源, 包括数据库、数据库管理系统等, 既能仅供本结点用户存取使用, 又能供其他结点上的用户存取使用。

(2) 面向对象数据库系统。面向对象数据库系统是面向对象的程序设计技术与数据库技术相结合的产物, 主要特点是具有面向对象技术的封装性和继承性, 提高了软件的可重用性。面向对象数据库系统包含了关系数据库管理系统的全部功能, 只是在面向对象环境中增加了一些新内容, 其中有些是关系数据库管理系统所没有的。面向对象数据库系统的基本设计思想是: 一方面把面向对象的程序设计语言向数据库方向扩展, 使应用程序能够存取并处理对象; 另一方面扩展数据库系统, 使其具有面向对象的特征, 以便对现实世界中复杂应用的实体和联系进

行建模。

(3) 多媒体数据库系统。多媒体数据库系统是数据库技术与多媒体技术相结合的产物。能够直接管理文本、图形、音频和视频等多媒体数据的数据库可称为多媒体数据库。多媒体数据库的结构和操作与传统格式化数据库的结构和操作有很大差别,在多媒体信息管理环境中,不仅数据本身的结构和存储形式各不相同,不同领域对数据处理的要求也比一般事务管理复杂得多,因此对数据库管理系统提出了更高的功能要求。综合程序设计语言、人工智能和数据库领域的研究成果,设计支持多媒体数据管理的数据库管理系统,已成为数据库领域中一个新的研究方向。

(4) 数据仓库技术。数据仓库技术是基于信息系统业务发展的需要,基于数据库系统技术发展而来,并逐步独立的一系列新的应用技术。数据仓库涉及三方面的技术内容:数据仓库技术、联机分析处理技术和数据挖掘技术。数据仓库用于数据的存储和组织,联机分析处理集中于数据的分析,数据挖掘则致力于知识的自动发现。它们都可以分别应用到信息系统的设计和实现中,以提高相应部分的处理能力。由于这三种技术内在的联系性和互补性,将它们结合起来就是一种新的决策支持系统架构。数据仓库最根本的特点是物理地存放数据,而且这些数据并不是最新的、专有的,而是来源于其他数据库的。数据仓库的建立并不是要取代数据库,它要建立一个较全面和完善的信息服务应用的基础上,用于支持高层决策分析,而事务处理数据库在企业的信息环境中承担的是日常操作性的任务。数据仓库是数据库技术的一种新的应用,到目前为止,数据仓库还是用关系数据库管理系统来管理其中的数据。

(5) 大数据技术。大数据(Big Data)是一种在获取、存储、管理和分析方面大大超出传统数据库软件工具能力范围的数据集合,具有数据规模大、数据种类多、要求数据处理速度快和数据价值密度低四大特征。大数据的概念与海量数据不同,后者只强调数据的量,而大数据不仅用来描述大量的数据,还进一步指出数据的复杂形式、数据的快速处理特性,以及对数据分析处理后最终获得有价值信息的能力。

目前,大数据已经广泛应用于包括金融、汽车、餐饮、电信、能源、体育和娱乐等在内的社会各行各业。例如,在互联网行业,借助于大数据技术可以分析客户行为,从而进行商品推荐和针对性广告投放;在汽车行业,利用大数据和物联网技术的无人驾驶汽车即将走入我们的日常生活;在城市管理行业,可以利用大数据实现智能交通、环保监测、城市规划和智能安防等。

1.2 数据库系统



视频 1-2 数据库系统

数据库系统(Database System, DBS)是指基于数据库的计算机应用系统,是实现有组织地、动态地存储大量相关数据,提供数据处理和信息资源共享的便利手段。

1.2.1 数据库系统的组成

一个完整的数据库系统,主要包括数据库(Database, DB)、数据库管理系统(Database Management System, DBMS)、数据库应用系统和数据库用户四大部分,各部分的关系如图 1-6 所示。

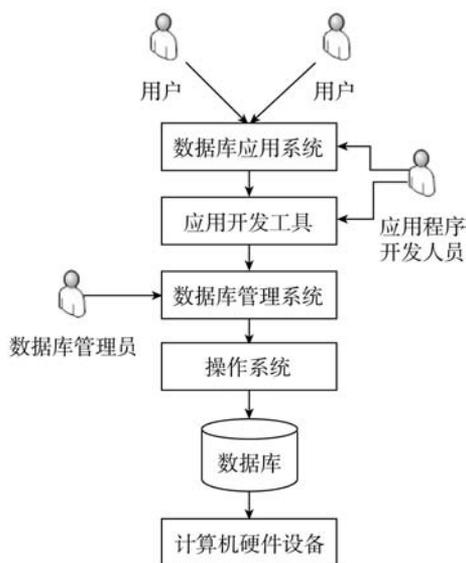


图 1-6 数据库系统

1. 数据库

数据库(DB)是指以一定结构存储在外部存储设备上的、能为多个用户共享的、与应用程序相互独立的、相互关联的结构化数据集合。数据库不仅存储了数据，还存储了数据与数据之间的关系。一个数据库由若干张表(Table)组成，例如，要创建一个超市管理系统的数据库，就需要建立员工表、部门表、工资表、商品表和销售表等，每个表都具有特定的结构，表与表之间有某种关联。在数据库的物理组织中，表以文件形式存储。

2. 数据库管理系统

数据库管理系统(DBMS)是用于描述、管理和维护数据库的软件系统，是数据库系统的核心组成部分。具有代表性的数据库管理系统有 Oracle、Microsoft SQL Server、MySQL 及 Microsoft Access 等。本书介绍的 Microsoft Access 软件是一种被广泛应用的小型数据库管理系统。

DBMS 在操作系统的基础上工作，它接收用户的操作命令并予以实施，从而完成用户对数据库的管理操作。无论是数据库管理员还是终端用户，都不能直接对数据库进行访问或操作，而必须利用 DBMS 提供的操作语言来使用或维护数据库中的数据。

数据库管理系统具有以下几个方面的功能。

- (1) 数据定义功能。使用数据定义语言(Data Definition Language, DDL)定义数据库的结构，刻画数据库框架等。
- (2) 数据操纵功能。使用数据操纵语言(Data Manipulation Language, DML)实现数据的检索、插入、删除、修改等操作。
- (3) 数据库运行管理功能。控制整个数据库系统的运行，控制用户的并发性访问，检验数据的安全、保密与完整性等。
- (4) 数据库的建立和维护功能。控制数据库初始数据的输入与数据转换，记录工作日志，监视数据库性能，修改更新数据库，恢复出现故障的数据库等。
- (5) 数据通信功能。与操作系统协调完成数据的传输，实现用户程序与 DBMS 之间数据的

通信。

3. 数据库应用系统

数据库应用系统就是我们常用的应用程序，是指开发人员利用某种应用开发工具开发出来的、面向某一类实际应用的软件系统，例如各种常用的手机软件，如微信、QQ、淘宝等。数据库应用系统需要通过数据库接口技术，在数据库管理系统的支持下才能获取或修改数据库中的数据。

4. 数据库用户

数据库用户主要有以下四类。

(1) 终端用户。终端用户是指通过应用程序界面使用数据库的人，他们不必了解数据库原理和实现细节，数据库对于他们而言是透明的。当他们使用应用程序访问数据库时，实质上是利用系统的接口或查询语言访问数据库。

(2) 系统分析员和数据库设计人员。系统分析员负责应用系统的需求分析和规范说明，确定系统的硬件配置和参与数据库系统的概要设计。数据库设计人员负责确定数据库中的数据和设计数据库各级模式。

(3) 应用程序开发人员。应用程序开发人员负责开发使用数据库的应用程序，这些应用程序可对数据进行检索、建立、删除或修改等。

(4) 数据库管理员。数据库管理员负责数据库的总体信息控制，具体职责包括：决定数据库的存储结构和存取策略，定义数据库的安全性要求和完整性约束条件，监控数据库的使用和运行等。

1.2.2 数据库系统的特点

1. 数据结构化

在数据库系统中，每一个数据库都是为某一应用领域服务的，因此不仅要考虑某个应用的数据结构，还要考虑整个组织(多个应用)的数据结构。这种数据组织方式使数据结构化，在描述数据时不仅要描述数据本身，还要描述数据之间的联系。数据库系统实现整体数据的结构化是数据库的主要特点之一，也是数据库系统与文件系统的本质区别。

2. 数据共享性高，冗余度可控

数据库技术的根本目标之一是解决数据共享问题。数据共享是指多个用户或应用程序可以访问同一个数据库中的数据，而且数据库管理系统提供并发和协调机制，保证在多个应用程序同时访问、存取和操作数据库数据时不产生任何冲突。数据冗余是指数据之间的重复，也可以说是同一数据存储在不同数据文件中的现象。数据冗余既浪费存储空间，又容易产生数据不一致等问题。数据库的数据已经根据特定的数据模型结构化，有效地节省了存储资源，减少了数据冗余，保证了数据的一致性。

3. 数据独立性高

数据独立性是指应用程序与数据库的数据结构之间相互独立。在数据库系统的数据存储结构发生改变时，不会影响数据的全局逻辑结构，保证了数据的物理独立性；在全局逻辑结

构发生改变时，不影响用户的局部逻辑结构及应用程序，保证了数据的逻辑独立性。

4. 数据统一控制

为保证多个用户能同时正确地使用同一个数据库，数据库管理系统提供了一套有效的数据控制手段，包括数据安全性控制、数据完整性控制、数据库的并发控制和数据库的备份恢复等，增强了多用户环境下数据的安全性和一致性保护。

1.2.3 数据库系统的三级模式

数据库领域公认的标准结构是三级模式结构，它包括外模式、概念模式和内模式，如图 1-7 所示。三级模式对应三个抽象级别，使用户能够逻辑地、抽象地处理数据，而不必关心数据在计算机中的物理表示和存储方式，把数据的具体组织交给数据库管理系统去完成。



图 1-7 数据库系统的三级模式

1. 外模式

外模式又称用户模式，是数据库用户看到的视图模式。“视图”是数据库用户所看到的数据库的数据视图，是与某一应用有关的数据的逻辑表示。用户可以通过外模式描述语言来描述与定义对应于用户的数据记录，也可以利用数据操纵语言对这些数据记录进行操作。外模式反映了数据库系统的用户观。

2. 概念模式

概念模式又称逻辑模式，是对数据库中全部数据的逻辑结构和特征的总体描述，是由数据库管理系统提供的数据库模式描述语言来描述、定义的。概念模式反映了数据库系统的整体观。

3. 内模式

内模式又称存储模式，它描述了数据在存储介质上的存储方式和物理结构，对应着实际存储在外存储介质上的数据库。内模式由内模式描述语言来描述、定义的。内模式反映了数据库系统的存储观。

在一个数据库系统中，内模式是唯一的，但建立在数据库系统之上的应用则是非常广泛且多样的，所以对应的外模式不是唯一的，也不可能是唯一的。

1.3 数据模型

由于计算机不能直接处理现实世界的具体事物，所以必须将这些具体事物转换成计算机能够处理的数据。在数据库技术中，用数据模型(Data Model)对现实世界中的事物进行抽象和表示。数据(Data)是描述事物的符号记录，模型(Model)



视频 1-3 数据模型

是现实世界的抽象，数据模型从抽象层次上描述了系统的静态特征、动态行为和约束条件，为数据库系统的信息表示与操作提供了一个抽象的框架。

1.3.1 数据抽象过程

从现实世界中的客观事物到数据库中存储的数据是一个逐步抽象的过程，如图 1-8 所示，这个过程经历了现实世界、概念世界和计算机世界三个阶段。



图 1-8 数据抽象过程

1. 现实世界

现实世界是指客观存在的事物及其相互间的联系。现实世界中的事物有着众多的特征和千丝万缕的联系，计算机处理的对象是现实世界的客观事物，在实施处理的过程中，需要对事物进行整理、分类和规范，进而将规范化的事物数据化，最终实现由数据库系统存储和处理。

2. 概念世界

概念世界又称为信息世界，是人们把现实世界中事物的信息和联系，通过特定符号记录下来，然后用规范化的数据库定义语言来定义描述而构成的一个抽象世界。在概念世界中，不是简单地对现实世界进行符号化，而是要通过筛选、归纳、总结、命名等抽象过程产生出概念模型，用以表示对现实世界的抽象与描述。概念模型的表示方法很多，目前较为常用的是实体-联系模型(Entity Relationship Model)，简称 E-R 模型。

3. 计算机世界

计算机世界又称为数据世界，是将概念世界的内容数据化后的产物。计算机世界将概念世界中的概念模型，进一步转换成数据模型，形成计算机能够处理的数据表现形式。

1.3.2 概念模型(E-R 模型)

把现实世界抽象为概念世界，建立概念世界中的数据模型，该数据模型称为概念模型。概念模型是面向数据库用户的对现实世界的抽象与描述的数据模型，它使数据库的设计人员在设计的初始阶段摆脱计算机系统及数据库管理系统的具体技术问题，集中精力分析数据及数据之间的联系等。最常用的概念模型表示方法是 P. P. Chen 于 1976 年提出的“实体-联系模型”(即 E-R 模型)，使用 E-R 图来表示。

1. E-R 模型的基本要素

(1) 实体(Entity)。客观存在并可以相互区别的事物称为实体。实体可以是人、事、物(例如：一名员工，一个商品)，也可以是抽象的概念和联系(例如：员工和公司的关系)。同一类型实体的集合称为实体集(例如：全体员工就是一个实体集)。

(2) 属性(Attribute)。用来描述实体的特性称为属性。例如,员工具有姓名、年龄、性别等属性信息。不同的属性会有不同的取值范围,属性的取值范围称为该属性的值域。例如,“年龄”属性的值域是 0~150。

(3) 联系(Relationship)。实体之间的对应关系称为联系。例如,顾客和商品之间具有购买关系。

2. 实体间的联系

两个实体之间的联系可分为三种类型:一对一联系(1:1),一对多联系(1:n),多对多联系(m:n)。

(1) 一对一联系(1:1)。对于实体集 A 中的每一个实体,实体集 B 只有一个实体与之联系,反之亦然,则称实体集 A 和实体集 B 具有一对一的联系,记作 1:1。例如图 1-9 所示,一个乘客只能坐一个座位,而一个座位只能被一个乘客坐,乘客与座位之间的联系就是一对一的联系。

(2) 一对多联系(1:n)。对于实体集 A 中的每一个实体,实体集 B 中有 n 个实体与之联系;反之实体集 B 的每个实体,实体集 A 中只有一个实体与之联系,则称实体集 A 与实体集 B 具有一对多的联系,记作 1:n。例如图 1-10 所示,一个部门有许多个员工,但一个员工只能在一个部门任职,部门和员工之间的联系就是一对多的联系。

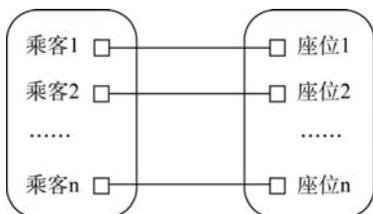


图 1-9 一对一联系

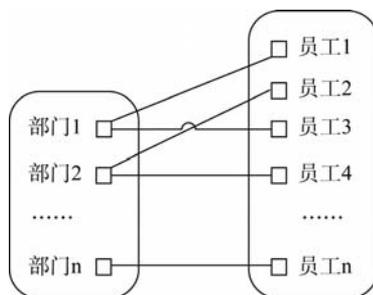


图 1-10 一对多联系

(3) 多对多联系(m:n)。对于实体集 A 中的每一个实体,实体集 B 中有 n 个实体与之联系;反之实体集 B 的每个实体,实体集 A 中有 m 个实体与之联系,则称实体集 A 与实体集 B 具有多对多的联系,记作 m:n。例如图 1-11 所示,一名员工可以销售多种商品,任何一种商品可以被多名员工销售,员工和商品之间具有多对多的联系。

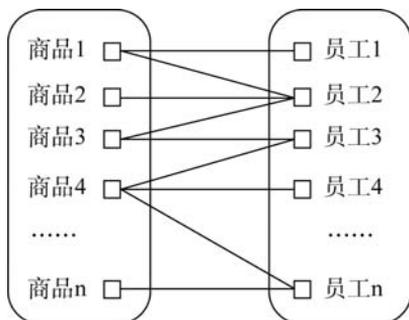


图 1-11 多对多联系

3. E-R 模型的表示方法：E-R 图

使用实体-联系图(即 E-R 图)可以直观地表达 E-R 模型。在 E-R 图中, 实体用矩形表示, 属性用椭圆表示, 联系用菱形表示, 在各自内部写明实体名、属性名和联系名, 并用连线连接起来, 同时在连线上标注联系的类型(1:1、1:n 或 m:n)。E-R 图用到的符号如图 1-12 所示。



图 1-12 E-R 图的表示符号

使用 E-R 图能够直观地表达数据库的信息组织情况。前述乘客和座位的 E-R 图如图 1-13 所示, 其中“乘客”实体有“身份证号”和“姓名”两个属性, “座位”实体有“座位号”和“舱位”两个属性, “乘坐”联系有“乘坐时间”一个属性, 乘客和座位之间是一对一的联系。

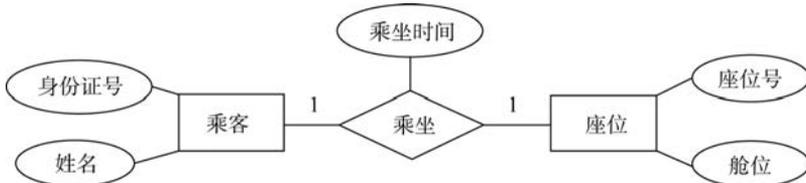


图 1-13 乘客和座位的 E-R 图

前述部门和员工的 E-R 图如图 1-14 所示, 其中“部门”实体有“部门编号”“部门名称”和“部门电话”三个属性, “员工”实体有“员工编号”“姓名”和“性别”三个属性, “聘请”联系有“是否在职”一个属性, 部门和员工之间是一对多的联系。

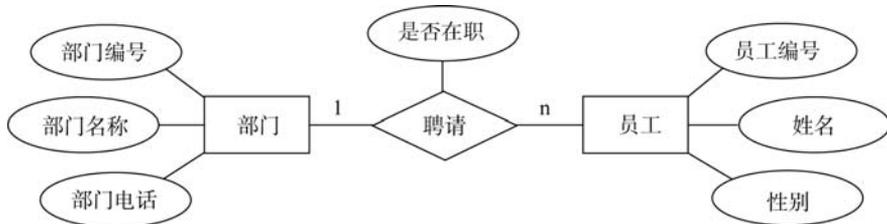


图 1-14 部门和员工的 E-R 图

前述员工和商品的 E-R 图如图 1-15 所示, 其中“员工”实体有“员工编号”“姓名”和“性别”三个属性, “商品”实体有“商品编号”“商品名称”和“零售价”三个属性, “销售”联系有“购买数量”一个属性, 员工和商品之间是多对多的联系。

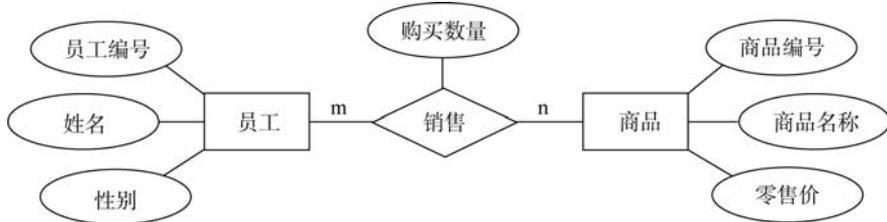


图 1-15 员工和商品的 E-R 图

1.3.3 常见的数据模型

数据库的类型是根据数据模型来划分的,任何一个 DBMS 也是根据数据模型有针对性地设计出来的。目前成熟地应用在数据库系统中的数据模型有层次数据模型、网状数据模型和关系数据模型。三者间的根本区别在于数据之间联系的表示方式不同,层次数据模型以“树结构”表示数据之间的联系;网状数据模型是以“图结构”来表示数据之间的联系;关系数据模型是用“二维表”(或称为关系)来表示数据之间的联系。

1. 层次数据模型

用树状结构表示实体及实体之间联系的数据模型称为层次数据模型。层次数据模型是数据库系统最早使用的一种模型,它的数据结构是一棵“有向树”,如图 1-16 所示,根结点在最上端,子结点在下,逐层排列。层次模型树中每一个结点表示一个实体,结点之间的连线表示实体之间的联系。这种模型适用于表达一对多的层次联系,但不能直接表达多对多的联系。

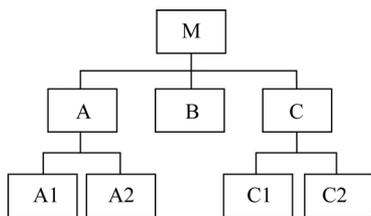


图 1-16 层次数据模型

2. 网状数据模型

用网状结构表示实体及实体之间联系的模型称为网状数据模型。网状数据模型和层次数据模型类似,用每个结点表示一个实体,结点之间的连线表示实体间的联系,但与层次数据模型不同的是,网状数据模型允许一个以上的结点无父结点,并且一个结点可以有多个父结点,如图 1-17 所示。网状数据模型能更直接地表示实体间的各种联系,但它的结构复杂,实现的算法也复杂。

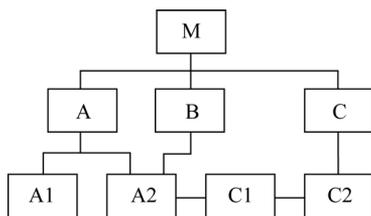


图 1-17 网状数据模型

3. 关系数据模型

用二维表的形式表示实体和实体之间联系的数据模型称为关系数据模型。在关系数据模型中,操作的对象和结果都是二维表,每个二维表又可称为关系,例如表 1-1 是一个商品关系。关系数据模型是目前最流行的数据库模型,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,Access 就是一种关系数据库管理系统。

表 1-1 “商品” 关系

商品编号	商品名称	规格	类别	库存	零售价
S2018010201	凉茶	250mL	饮品	810	¥2.40
S2018010202	可口可乐	355mL	饮品	91	¥2.50
S2018010203	雪碧	355mL	饮品	145	¥2.50
S2018010204	矿泉水	550mL	饮品	121	¥1.30
S2018010205	冰红茶	490mL	饮品	150	¥2.40

1.4 关系数据库



视频 1-4 关系数据库

关系数据库是采用关系数据模型作为数据组织方式的数据库。关系数据库的特点在于它将每个具有相同属性的数据独立地存储在一个表中。

1.4.1 关系的术语

1. 关系

关系就是一张二维表，由行和列组成。每个关系都有一个关系名，在 Access 数据库中，关系名就是数据库中表的名称。例如，图 1-18 的“部门”关系就是“部门”表。

2. 元组

在一个二维表中，表中的行称为元组，每一行是一个元组，也称为一条记录，它对应于实体集中的一个实体。例如，图 1-18 的“部门”关系里每一个元组代表一个部门。

3. 属性

二维表中的列称为属性，每一列有一个属性名，也称字段名。例如，图 1-18 的“部门”关系里的“部门编号”就是部门的一个属性。

部门编号	部门名称	部门主管	部门电话
D1	客服部	Y001	86828385
D2	人事部	Y006	86821222
D3	销售部	Y009	86820304
D4	财务处	Y013	86824511

一个属性
 一列
 一个字段
 一个元组
 一行
 一条记录

“部门”关系 → “部门”表

图 1-18 关系与表

4. 值域

属性的取值范围称为值域，关系的每个属性都必须对应一个值域。例如，图 1-19 中的员工表的“性别”字段，值域为“男”或“女”两个值。

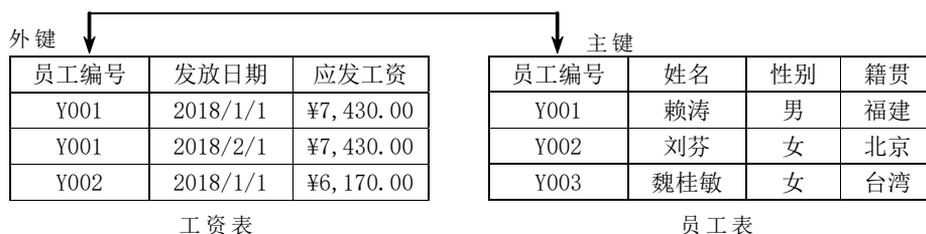


图 1-19 利用外键实现表与表的联系

5. 主键

主键又称关键字，或称为主码，是二维表中某个属性或属性的组合，其值能唯一地标识一个元组。例如，图 1-19 中的员工表的“员工编号”字段常常被设为主键，而不是用“姓名”字段作为主键，因为姓名可能重名，不能唯一地标识一个元组。

6. 外键

外键是外部关键字的简称。在关系模型中，为了实现表与表之间的联系，通常将一个表的主键作为数据之间的纽带放到另一个表中，这个起联系作用的属性就称为外键。例如，图 1-19 中员工表的“员工编号”属性是员工表的主键，工资表的“员工编号”属性是工资表的外键。通过“员工编号”这个公共属性，使得员工表和工资表产生了联系。

7. 对关系的描述

描述一个关系的格式为：关系名(属性名 1, 属性名 2, …, 属性名 n)

例如，表 1-1 的“商品”关系描述格式是：商品(商品编号, 商品名称, 规格, 类别, 库存, 零售价)

8. 关系的特点

关系是一个二维表，但并不是所有二维表都是关系。关系应具有以下特点。

- (1) 关系中的每个属性值是不可分解的。
- (2) 关系中的各列是同质的，即每一列的属性值必须是同一类型的数据，来自同一个值域。
- (3) 在同一个关系中不能出现相同的属性名。
- (4) 关系中不允许有完全相同的元组。
- (5) 在一个关系中元组和列的次序无关紧要，可以任意交换。

1.4.2 关系的运算

关系的基本运算有三种：选择、投影和连接。

1. 选择

选择运算是根据给定的条件，从一个关系中选出符合条件的元组(表中的行)，被选出的元组组成一个新的关系，这个新的关系是原关系的一个子集。选择是从行的角度进行的运算。例如，从员工关系中选出性别是男的员工信息，组成一个新的关系，如图 1-20 所示。

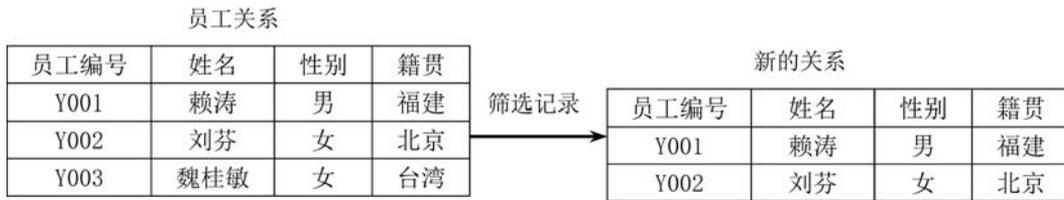


图 1-20 选择运算

2. 投影

投影就是从—个关系中选择指定的属性(表中的列), 被选中的属性重新排列组成一个新的关系。投影是从列的角度进行的运算。例如, 从员工关系中选取姓名、性别和籍贯属性, 组成一个新的关系, 如图 1-21 所示。

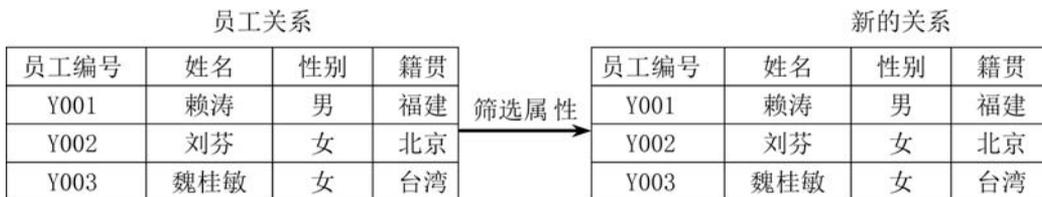


图 1-21 投影运算

3. 连接

连接运算是从两个或多个关系中选取属性间满足一定条件的元组, 组成一个新的关系。例如, 从员工关系和部门关系中选取员工编号、姓名、性别、部门名称和部门电话属性, 组成一个新的关系, 如图 1-22 所示。

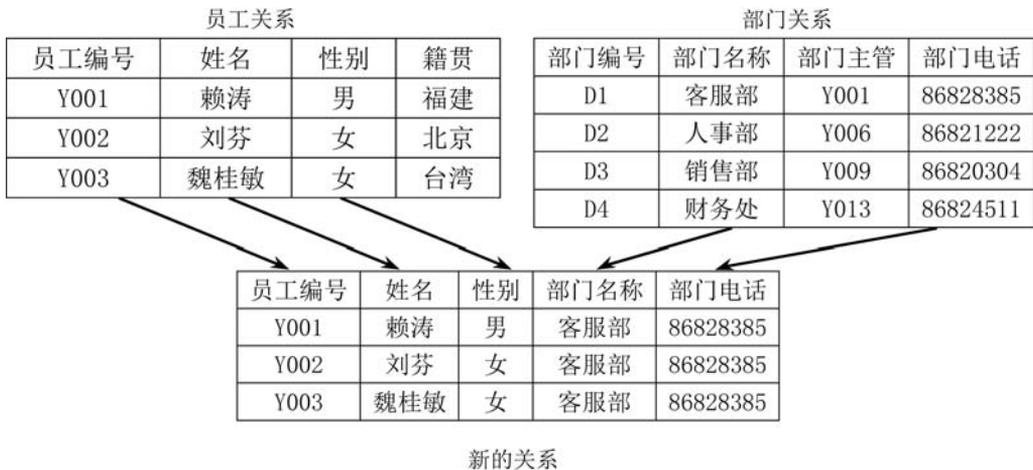


图 1-22 连接运算

1.4.3 关系的完整性规则

关系模型的数据完整性是指数据库中数据的正确性和一致性, 数据的完整性由数据完整性规则来维护。数据完整性规则有如下三种。

1. 实体完整性

实体完整性是指关系的主键不能取空值或重复的值。如果主键是多个属性的组合，则这些属性均不得取空值。例如，表 1-1 的“商品”关系，将“商品编号”属性作为主键，那么意味着该列不得有空值并且不得有重复的值，否则将无法对应某个具体的商品，这样的二维表是不完整的，该关系不符合实体完整性规则的约束条件。

2. 参照完整性

参照完整性反映了“主键”属性和“外键”属性之间的引用规则。外键要么取空值，要么等于相关关系中主键的某个值。例如，图 1-19 中的“员工”关系和“工资”关系，“工资”关系的外键“员工编号”属性的取值必须存在于“员工”关系中，而且是“员工”关系的主键。

如果实施了参照完整性，那么当主表(如“员工”关系)中没有相关记录时，就不能将记录添加到相关表中；也不能在相关表中存在匹配的记录时，删除主表中的记录；更不能在相关表中有相关记录时，更改主表中的主键值。

3. 用户定义完整性

实体完整性和参照完整性是关系模型中必须满足的完整性约束条件。除此之外，不同的关系数据库系统根据其应用环境的不同，或是为了满足应用方面的要求，往往还需要一些特殊的约束条件，这些完整性是由用户定义的，因此称为用户定义完整性。用户定义完整性比较常见的是设置属性的数据类型、取值范围、是否允许空值等。例如，对于表 1-1 的“商品”关系，可以对“库存”这个属性定义必须大于 0 的约束条件。

1.4.4 从概念模型到关系模型的转换

完成概念模型设计后，得到数据库的 E-R 模型，将 E-R 模型转换为关系模型，实际上就是将实体、实体的属性和实体间的联系转换为关系模式。对于不同的实体间联系，转换规则如下。

1. 1:1 联系到关系模型的转换

若实体间的联系是 1:1，可以在两个实体转换成两个关系模式后，在任意一个关系模式中增加另一关系模式的主键(作为外键处理)和联系的属性。

例如，图 1-13 所示的 E-R 图中有乘客和座位两个实体，两个实体是一对一联系，可以转换为如下两个关系：

乘客(身份证号，姓名，乘坐时间，座位号)

座位(座位号，舱位)

其中，“身份证号”是“乘客”关系的主键，“座位号”是“座位”关系的主键，在“乘客”关系中增加了“座位”关系的主键“座位号”作为外键，还增加了联系的属性“乘坐时间”。

2. 1:n 联系到关系模型的转换

若实体间的联系是 1:n，可以在两个实体转换成两个关系模式后，在 n 方实体的关系模式中增加 1 方实体的主键(作为外键处理)和联系的属性。

例如，图 1-14 所示的 E-R 图中有部门和员工两个实体，两个实体是一对多联系，可以转换为如下两个关系：

部门(部门编号, 部门名称, 部门电话)

员工(员工编号, 姓名, 性别, 是否在职, 部门编号)

其中,“部门编号”是“部门”关系的主键,“员工编号”是“员工”关系的主键,在“员工”关系中增加了“部门”关系的主键“部门编号”作为外键,还增加了联系的属性“是否在职”。

3. m:n 联系到关系模型的转换

若实体间的联系是 $m:n$,除了要将两个实体转换成两个关系模式,还要为联系单独建立一个关系模式,其属性是两个实体的主键加上联系的属性,其主键是两个实体主键的组合。

例如,图 1-15 所示的 E-R 图中有员工和商品两个实体,两个实体是多对多联系,可以转换为如下两个关系:

员工(员工编号, 姓名, 性别)

商品(商品编号, 商品名称, 零售价)

销售(员工编号, 商品编号, 购买数量)

其中,“员工编号”是“员工”关系的主键,“商品编号”是“商品”关系的主键,在“销售”关系中“员工编号”和“商品编号”组合起来成为主键。

1.5 数据库应用系统开发流程



视频 1-5 数据库应用系统开发流程

数据库应用系统的开发流程一般分为六个阶段:需求分析、概念模型设计、逻辑模型设计、物理模型设计、系统实施、系统运行和维护。在实际开发过程中可以根据应用系统的规模和复杂程度进行灵活调整,无须刻板地遵守整个开发流程,但总体上应当符合“分析→设计→实现”这个基本流程。表 1-2 简要列出了数据库应用系统开发各个阶段的主要任务和注意事项。

表 1-2 数据库应用系统开发流程

基本环节	阶段	主要任务和注意事项
分析	需求分析	(1) 需求分析就是了解和分析用户对系统的要求,这是设计数据库的起点。 (2) 需求分析的结果是否准确将直接影响到后面各个阶段的设计,并影响到设计结果是否合理和实用。 (3) 在收集需求时必须充分考虑今后可能的扩充和改变。 (4) 需要考虑数据库的安全性、完整性要求等
	概念模型设计	(1) 将需求分析得到的用户需求抽象为概念模型的过程。 (2) 概念模型是各种逻辑模型的基础。 (3) 概念模型的常用工具是 E-R 图
设计	逻辑模型设计	(1) 将概念模型转换为被数据库管理系统所支持的数据模型的过程,并对转换结果进行规范化处理。 (2) 如果采用的是关系数据库,就是将 E-R 图转换为关系模型的过程

(续表)

基本环节	阶段	主要任务和注意事项
设计	物理模型设计	(1) 为逻辑数据模型选取一个最适合应用要求的物理结构的过程。 (2) 物理模型设计一般分两步：一是确定数据库的存储结构和存取方法；二是对物理结构进行评价，重点评价时间和空间效率
实现	系统实施	(1) 根据数据库逻辑设计和物理设计结果建立数据库，创建各种数据库对象。 (2) 组织数据入库。 (3) 编码和调试
	系统运行和维护	(1) 在数据库系统的运行过程中，要不断地对数据库设计进行评价、调整和修改，这是一个长期的维护工作。 (2) 对数据库经常性的维护工作主要是由数据库管理员完成的，工作包括数据库的备份和恢复、数据库的安全性完整性控制、数据库性能的分析改造等

值得注意的是，设计一个完整的数据库应用系统是不可能一蹴而就的，它往往是上述六个阶段的不断反复。数据库不是独立存在的，它总是与具体的应用相关，前期需求分析阶段显得尤为重要，因此需要耐心地收集需求和分析数据，仔细梳理清楚数据间的关系，才能构造出最优的数据库模式，进而满足各种用户的应用要求。

1.6 Access 简介



视频 1-6 Access 简介

Microsoft Office Access 是由微软发布的关系数据库管理系统，本书采用的是 Access 2016 版本，是一个强大的、成熟的桌面关系数据库管理系统，包含在 Office 办公系列软件中，界面友好，易学易用且接口灵活。使用 Access 可以很高效地完成各种中小型数据库管理工作，可用于行政、财务、教育、审计等众多管理领域，尤其适合普通用户开发自己工作需要的各种小型数据库应用系统。

1.6.1 Access 的特点

Access 简单易用，通过 Web 数据库可以增强运用数据的能力，从而可以更轻松地跟踪、报告和共享数据。Access 主要的特点和增强功能有如下几个方面。

1. 应用模板实现专业设计

以 Access 中的数据库模板为基础，可以对其进行快速设置和修改数据外观，或进一步自定义以制作出美观的表格和报表。

2. 智能特性

Access 几乎为每一个对象都设有向导功能，利用向导功能可以迅速地建立一个基本对象，例如查询向导、窗体向导、报表向导等向导功能。同时，Access 采用的可视化设计工具，使得

用户基本不需要编写任何代码就可以完成数据库的大部分工作。另外，使用简化的表达式生成器，可以使用户更快速、更轻松地编写表达式。

3. 支持面向对象

Access 支持面向对象的开发方式，它将数据库管理的各种功能封装在表、查询、窗体等各类对象中，通过对对象的属性和方法来完成数据库的操作管理，极大地简化了用户的开发工作。

4. 功能强大的宏设计器

Access 具有一个改进的宏设计器，使用该设计器可以更快速地创建、编辑和自动化数据库逻辑，并轻松地整合更复杂的逻辑以创建功能强大的应用程序。

5. 通过 Web 网络共享数据库

Access 提供两种数据库类型的开发工具，一种是标准桌面数据库类型，另一种是 Web 数据库类型。使用 Web 数据库开发工具可以轻松方便地开发出网络数据库，从而使得没有 Access 客户端的用户也可以通过浏览器打开 Web 表格和报表，用户所做的更改也会自动同步到数据库中。

1.6.2 Access 数据库的系统结构

Access 是一个面向对象的可视化数据库管理工具，它提供了一个完整的对象类集合，在 Access 环境中进行的操作其实都是面向对象进行的。Access 数据库中包括六种数据对象，如图 1-23 所示，它们分别是表、查询、窗体、报表、宏和模块，通过“创建”选项卡提供的命令完成各种数据库对象的创建。



图 1-23 导航窗格中的六种数据库对象

1. 表

表是 Access 数据库中最基本的对象，它是实际存储数据的地方。一个 Access 数据库可以包含多个表，表与表之间可以相互独立，也可以相互联系。在创建数据库时，应先创建表，再创建其他数据库对象。

表由字段和记录组成，一个字段就是表中的一列，一条记录就是表中的一行。如图 1-24 所示的“商品”表，一条记录(即一行)对应一种商品，每个字段(即每列)描述商品的相关属性，如“商品编号”“商品名称”“规格”等。

商品编号	商品名称	规格	类别	库存	零售价
S2018010201	凉茶	250ml	饮品	810	¥2.40
S2018010202	可口可乐	355ml	饮品	91	¥2.50
S2018010203	雪碧	355ml	饮品	145	¥2.50
S2018010204	矿泉水	550ml	饮品	121	¥1.30
S2018010205	冰红茶	490ml	饮品	150	¥2.40
S2018010206	黑芝麻汤圆	800g	速冻品	147	¥14.90
S2018010207	猪肉水饺	千克	速冻品	100	¥18.20
S2018010208	山东馒头	280g	速冻品	86	¥39.90
S2018010209	肉粽	960g	速冻品	100	¥74.90
S2018010210	叉烧包	270g	速冻品	50	¥22.00

图 1-24 数据库对象“表”示例

2. 查询

查询是数据库处理和分析数据的工具，是基于表对象的基础上建立起来的。它是根据事先设置好的条件从表或其他查询中筛选出所需的数据，供用户查看、更改和分析使用。尽管从查询的数据表视图上看到的数据形式与从表的数据表视图上看到的数据形式完全一致，都是以二维表的形式显示数据，如图 1-25 所示。但查询与表不同，它并不是数据的物理集合，查询只记录该查询的操作方式，不会存储数据。

类别	平均价格	最高价格	最低价格
电器	¥1,515.00	¥4,299.00	¥89.00
日用品	¥28.74	¥45.00	¥12.90
速冻品	¥33.98	¥74.90	¥14.90
饮品	¥2.22	¥2.50	¥1.30

图 1-25 数据库对象“查询”示例

3. 窗体

窗体既是管理数据库的窗口，又是用户和数据库之间的桥梁。用户通过窗体可以方便地输入数据，编辑数据，查询、排序、筛选和显示数据。虽然“表”和“查询”也能展现和操作数据，但窗体的优点在于可以更人性化的方式呈现和操作数据，如图 1-26 所示。

人事管理
- □ ×

欢迎使用本数据库

员工信息
员工照片
添加新员工

员工编号	<input type="text" value="Y001"/>	部门名称	<input type="text" value="客服部"/>
姓名	<input type="text" value="赖涛"/>	部门主管	<input type="text" value="Y001"/>
性别	<input type="text" value="男"/>	上一条(P)	
籍贯	<input type="text" value="福建"/>	下一条(N)	
电话	<input type="text" value="13609876543"/>	关闭(C)	
年龄	<input type="text" value="57"/>		

人事管理的子窗体

员工编号	发放日期	应发工资	扣税	实发工资
Y001	2018/1/1	¥7,430.00	¥288.00	¥7,142.00
Y001	2018/2/1	¥7,430.00	¥288.00	¥7,142.00
* Y001				

记录: < 第 1 项(共 2 项) >
无筛选器

图 1-26 数据库对象“窗体”示例

4. 报表

报表是数据库中的数据通过打印机输出的特有形式。通过报表可以把用户需要的数据进行整理、计算或汇总统计，然后按照指定的样式打印输出，如图 1-27 所示。

订单详情							2022年1月12日 17:35:39
订单编号	顾客卡号	消费时间	实付款	商品名称	零售价	购买数量	
1	G201801	2018/5/1 9:30:00	¥23.50	矿泉水	¥1.30	5	
				凉茶	¥2.40	5	
				可口可乐	¥2.50	1	
				雪碧	¥2.50	1	
				商品总数量:		12	
2	G201802	2018/5/1 11:30:00	¥133.10	牙膏	¥16.90	1	
				猪肉水饺	¥18.20	2	
				山东馒头	¥39.90	2	
				商品总数量:		5	

图 1-27 数据库对象“报表”示例

5. 宏

宏是一个或多个操作命令的集合，其中每个命令实现特定的功能，如图 1-28 所示。某些普通的需要多个命令连续执行的任务可通过宏操作自动完成。因此，用户通过宏即可不用编写程序代码就能自动化地完成大量的工作。



图 1-28 数据库对象“宏”示例

6. 模块

模块是以 VBA(Visual Basic for Application)语言为基础编写的程序集合。模块中包含过程，每个过程实现特定功能，如图 1-29 所示。模块的主要作用是设计建立复杂的 VBA 程序以完成宏不能完成的任务。

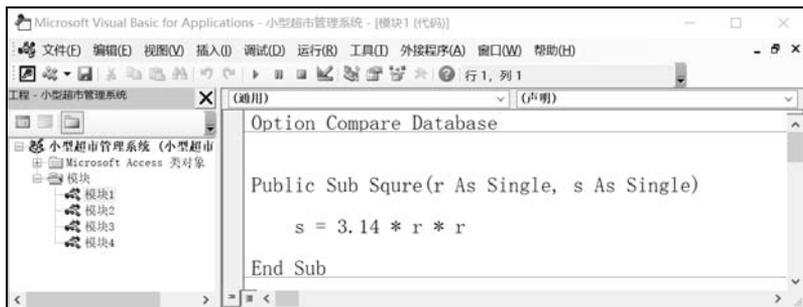


图 1-29 数据库对象“模块”示例

1.6.3 Access 2016 主界面

1. 后台视图

在使用数据库前需要先打开 Access 程序，然后打开需要使用的数据库文件。用户启动 Access 但还未打开数据库文件时，主界面自动进入后台视图，如图 1-30 所示。后台视图包含创建新数据库、打开现有数据库等命令。



图 1-30 Access 2016 的后台视图

2. 功能区

在 Access 中打开数据库文件后，功能区显示在 Access 主窗口的顶部，每一个功能区都由一个选项卡标签来标识。选项卡分为主选项卡和上下文选项卡。

(1) 主选项卡包括“文件”“开始”“创建”“外部数据”“数据库工具”“帮助”选项卡，如图 1-31 所示。



图 1-31 Access 2016 的功能区

功能区中的每个选项卡都包含多组相关命令，可以用来操作相应的数据对象。各选项卡包含的主要操作见表 1-3。

表 1-3 Access 2016 功能区选项卡包含的主要操作

选项卡	主要命令
文件	打开后台视图
开始	选择不同的视图
	从剪贴板复制和粘贴
	对记录进行排序和筛选
	使用记录(刷新、新建、保存、删除、合计等)
	查找记录
	设置字体格式
创建	创建表格
	创建查询
	创建窗体
	创建报表
	创建宏和模块
外部数据	导入或链接到外部数据
	导出数据
数据库工具	压缩和修复数据库
	宏和 VBA 模块
	创建和查看表关系
	运行数据库文档或分析性能
	将数据库移至 Microsoft SQL Server 或 Access 数据库

(2) 上下文选项卡是只有当用户执行了某种特定的操作后才会出现的。例如，当用户打开查询设计视图时，才会出现“查询工具”上下文选项卡——“设计”选项卡，主要用于对查询设计进行相关设置操作，如图 1-32 所示。



图 1-32 “查询工具”上下文选项卡

3. 导航窗格

在 Access 中打开数据库文件时，左侧的导航窗格区将显示当前数据库中各种数据库对象，如表、窗体、报表、查询等。通过选择对象类型可以罗列出所选对象，如图 1-33 所示。

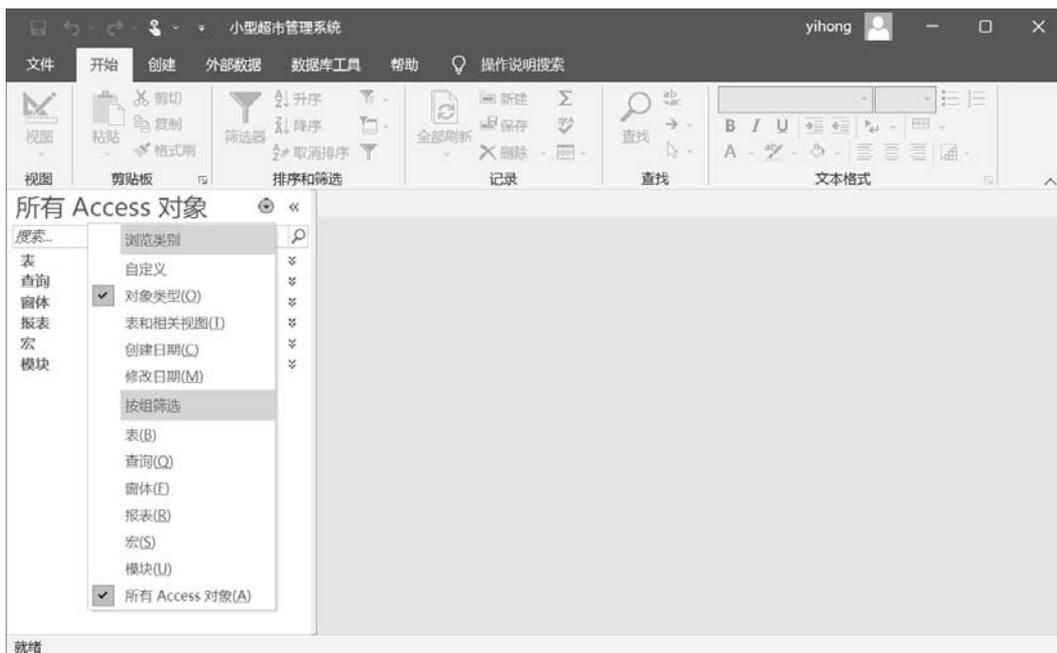


图 1-33 Access 2016 的导航窗格

1.7 思考与练习

1.7.1 思考题

1. 数据和信息有何区别？
2. 数据库系统由哪几部分组成？请解释各部分的作用和区别。
3. 数据库系统的特点有哪些？
4. 数据库系统的三级模式结构是什么？
5. E-R 模型的基本要素有哪些？
6. 实体之间的联系有哪三种类型？举例说明。
7. 常见的数据模型有哪些？什么是关系数据库？
8. 数据库应用系统的开发设计有哪些基本步骤？
9. Access 数据库的系统结构由哪些对象组成？其中最基本的对象是什么？

1.7.2 选择题

1. 有关信息与数据的概念，下面说法正确的是()。
 - A. 信息与数据是同义词
 - B. 数据是承载信息的物理符号
 - C. 信息和数据毫无关系
 - D. 固定不变的数据就是信息
2. 数据库(DB)、数据库系统(DBS)和数据库管理系统(DBMS)三者之间的关系是()。
 - A. DBS 包括 DB 和 DBMS
 - B. DBMS 包括 DB 和 DBS
 - C. DB 包括 DBS 和 DBMS
 - D. 三者不存在关系

3. 数据库系统中数据的特点是()。
- A. 共享度高, 无冗余, 独立性好 B. 共享度高, 冗余度低, 独立性好
C. 共享度高, 冗余度高, 独立性差 D. 共享度低, 冗余度低, 独立性好
4. 在数据库系统的三级模式结构中, 为用户描述整个数据库逻辑结构的是()
- A. 外模式 B. 概念模式 C. 内模式 D. 存储模式
5. 用二维表来表示实体及实体间联系的数据模型是()。
- A. 实体-联系模型 B. 层次模型 C. 关系模型 D. 网状模型
6. 关系数据库管理系统中的关系是指()。
- A. 不同元组间有一定的关系 B. 不同字段间有一定的关系
C. 不同数据库间有一定的关系 D. 满足一定条件的二维表格
7. 从教师表中找出女性讲师的记录, 属于()关系运算。
- A. 选择 B. 投影 C. 连接 D. 交叉
8. 在 E-R 图中, 用()来表示属性。
- A. 椭圆形 B. 矩形 C. 菱形 D. 三角形
9. 以下对关系模型的描述, 不正确的是()。
- A. 在一个关系中, 每个数据项是最基本的数据单位, 不可再分
B. 在一个关系中, 同一列数据具有相同的数据类型
C. 在一个关系中, 各列的顺序不可以任意排列
D. 在一个关系中, 不允许有相同的字段名
10. 开发超市管理系统过程中开展超市信息处理的调查, 属于数据库应用系统设计中() 阶段的任务。
- A. 物理设计 B. 概念设计 C. 逻辑设计 D. 需求分析
11. Access 数据库中最基本的对象是()。
- A. 表 B. 宏 C. 报表 D. 模块
12. Access 中表和数据库的关系是()。
- A. 一个数据库可以包含多个表 B. 一个数据库只能包含一个表
C. 一个表可以包含多个数据库 D. 数据库就是数据表
13. Access 是一种()。
- A. 操作系统 B. 数据库管理系统
C. 电子表格 D. 字处理软件