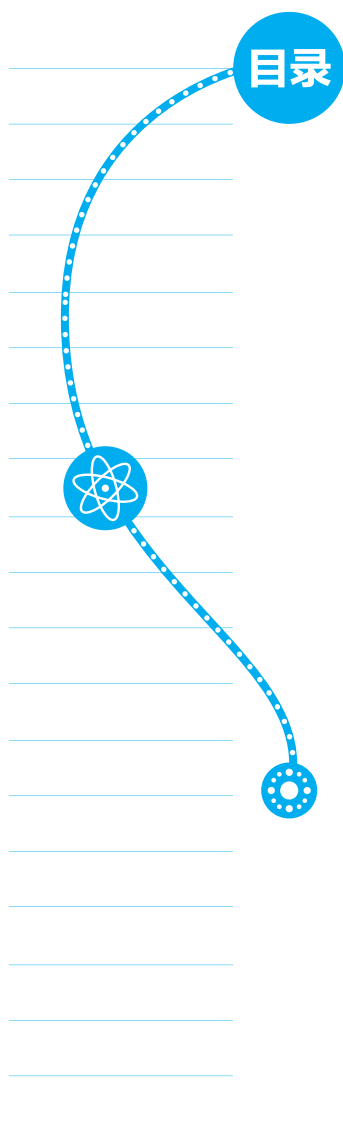




第1章 数据处理基础	1
1.1 数据与数据处理	2
1.1.1 数据与信息	2
1.1.2 数据处理	2
1.1.3 数据处理的发展阶段	3
1.2 数据处理与数据管理	3
1.2.1 人工管理阶段	3
1.2.2 文件系统阶段	3
1.2.3 数据库系统阶段	4
1.3 数据模型	4
1.3.1 概念模型与 E-R 图	5
1.3.2 逻辑数据模型	7
1.4 商品进销存系统的数据库设计	9
1.4.1 商品进销存系统背景介绍	9
1.4.2 商品进销存系统功能介绍	10
1.4.3 商品进销存系统数据模型设计	10
本章小结	11
习 题	12

第2章 宏与 VBA	15
2.1 宏的基础操作	16
2.1.1 录制宏	16
2.1.2 查看宏	17
2.1.3 执行宏	18
2.1.4 删除宏	18
2.1.5 保存启用宏的工作簿	18
2.2 宏的执行方式与安全设置	18
2.2.1 使用快捷键执行宏	19
2.2.2 使用对象执行宏	19
2.2.3 使用功能区按钮执行宏	20
2.2.4 宏的安全设置	21
2.3 宏的实际应用	23
2.3.1 制作简易工资条	23
2.3.2 利用宏一键排名	25

2.4 VBA 与 VBA 开发环境	25
2.4.1 VBA 简介	25
2.4.2 进入 VBA 开发环境	26
2.4.3 使用 VBA 开发环境	26
2.4.4 配置 VBA 开发环境	30
本章小结	31
习 题	31

第 3 章 VBA 编程基础	33
3.1 VBA 语言基础	34
3.1.1 关键字与标识符	34
3.1.2 数据类型	34
3.1.3 常量	36
3.1.4 变量	38
3.1.5 运算符和表达式	40
3.1.6 常用标准函数	43
3.2 VBA 语句	50
3.2.1 声明语句和执行语句	50
3.2.2 语句书写规则	50
3.2.3 语句格式规则	51
3.3 VBA 程序与程序调试	52
3.3.1 VBA 程序	52
3.3.2 调试程序与错误处理	53
3.3.3 错误处理	55
3.4 VBA 数据输入与输出	56
3.4.1 输入对话框	57
3.4.2 消息提示对话框	57
3.4.3 Debug.Print 方法	59
3.5 程序控制结构	60
3.5.1 结构化程序设计	60
3.5.2 顺序结构	61
3.5.3 选择结构	61
3.5.4 循环结构	66

3.6 数组	72
3.6.1 数组的声明	72
3.6.2 数组元素的引用和赋值	73
3.7 程序设计实例	75
3.7.1 顺序结构程序实例	75
3.7.2 选择结构程序实例	76
3.7.3 循环结构程序实例	77
本章小结	83
习 题	83

第 4 章 模块化编程	87
4.1 模块化编程相关概念	88
4.1.1 模块化设计概述	88
4.1.2 模块化编程概述	88
4.2 过程与函数的定义	89
4.2.1 过程	89
4.2.2 函数	90
4.2.3 过程与函数的区别	91
4.3 过程与函数的调用	92
4.3.1 调用过程	92
4.3.2 调用函数	93
4.3.2 递归调用	93
4.4 参数传递	94
4.4.1 形参和实参	94
4.4.2 按值传递和按地址传递	94
4.4.3 数组参数	96
4.4.4 可选参数与命名参数	96
4.5 变量的作用域与生存期	97
4.5.1 变量的作用域	97
4.5.2 静态变量	97
本章小结	97
习 题	97

第 5 章 Excel 数据对象与数据操作	101
5.1 面向对象程序设计基础	102
5.1.1 面向对象程序设计思想	102
5.1.2 Excel VBA 对象	103
5.1.3 Excel VBA 对象引用方法	105
5.1.4 设置对象属性值与调用对象方法	109
5.1.5 集合对象的属性和方法	110
5.2 Application 对象	111
5.2.1 Application 对象常用属性	112
5.2.2 Application 对象常用方法	116
5.2.3 Application 对象常用事件	117
5.3 Workbooks 对象与 Workbook 对象	119
5.3.1 Workbooks 对象及其方法	119
5.3.2 引用 Workbook 对象	120
5.3.3 Workbook 对象常用属性	120
5.3.4 Workbook 对象常用方法	122
5.3.5 Workbook 对象常用事件	124
5.4 Worksheets 对象与 Worksheet 对象	126
5.4.1 Worksheets 对象及其方法	126
5.4.2 引用 Worksheet 对象	127
5.4.3 Worksheet 对象常用属性	128
5.4.4 Worksheet 对象常用方法	130
5.4.5 Worksheet 对象常用事件	133
5.5 Range 对象	133
5.5.1 引用 Range 对象	133
5.5.2 Range 对象常用属性	137
5.5.3 Range 对象常用方法	142
5.6 Excel 中的数据管理	145
5.6.1 建立数据清单及数据管理工作表	145
5.6.2 记录数据的显示与浏览	147
5.6.3 记录数据的新增、删除和修改	149
本章小结	150
习 题	150

第 6 章 窗体设计	153
6.1 窗体设计器环境	154
6.1.1 插入窗体与窗体设计器	154
6.1.2 窗体工具箱	154
6.1.3 属性窗口	155
6.1.4 窗体代码窗口	156
6.1.5 控件对象的基本操作	157
6.1.6 窗体的设计步骤	158
6.2 窗体设计基础	160
6.2.1 控件对象的属性	160
6.2.2 控件对象的方法	161
6.2.3 控件对象的事件	161
6.3 窗体及常用控件	162
6.3.1 窗体	162
6.3.2 标签控件	163
6.3.3 文本框控件与编辑框控件	163
6.3.4 命令按钮控件	166
6.3.5 选项按钮控件	168
6.3.6 复选框控件	168
6.3.7 组合框控件和列表框控件	169
6.3.8 滚动条与旋转按钮控件	170
6.3.9 图像控件	171
6.3.10 多页控件	171
6.4 窗体设计综合实例	173
本章小结	176
习 题	176

第 7 章 数据透视表操作	179
7.1 数据透视表概述	180
7.2 建立数据透视表	181
7.3 查看数据透视表	184
7.4 数据透视表涉及的对象	185
7.4.1 PivotTable 数据透视表对象	185
7.4.2 PivotField 数据透视表字段对象	186

7.4.3 PivotCache 内存缓冲区对象	186
7.5 理解对象	187
7.5.1 录制宏	187
7.5.2 理解宏代码	188
7.6 实现数据透视表自动化	189
7.6.1 分析实际问题	189
7.6.2 创建保存数据透视表的工作表	189
7.6.3 指定数据透视表的源数据区域	190
7.6.4 设置数据透视表的行标签、列标签和汇总值	191
7.6.5 CreatePivotTable 过程	192
7.6.5 调用 CreatePivotTable 过程	194
本章小结	197
习 题	197

第 8 章 数据处理系统的开发	199
8.1 数据处理系统的开发步骤	200
8.2 Excel 应用程序	201
8.2.1 应用程序的组成	201
8.2.2 应用程序的组织与管理	202
8.3 商品进销存系统开发实例	202
8.3.1 商品进销存系统的组织	202
8.3.2 建立商品进销存系统项目	204
8.3.3 商品进销存系统的测试	218
本章小结	220
习 题	220

参考文献	222
-------------	-----



本章概述

随着信息技术的发展和移动互联网的快速普及,各种网络应用与人们的日常生活和企业的生产经营交汇融合,产生了大量的数据。数据的产生方式也从早期的企业内部运营系统收集数据发展到 Web2.0 时代的用户原创数据。在即将到来 5G 网络时代,随着物联网的普及,大量的传感器将产生巨量的数据。据 IDC 的估测,全球的数据以每年 50% 的速度增长,预计到 2020 年,全球将总共拥有 35ZB 的数据。面对如此巨量、繁杂的数据如果不经数据处理,将造成数据过载,从而使数据管理的代价超过从数据中发掘出信息的价值。数据处理技术正是解决这一矛盾的技术,它所研究的问题是如何科学地对数据进行采集、存储、加工处理,从而得到对人们有价值的信息。数据处理几乎涉及所有的应用领域,从小型的 Excel 数据处理系统到大数据处理系统,从电子商务到电子政务,数据处理技术都发挥着不可替代的作用。



学习目标

- ①了解什么是数据、信息,以及数据处理
- ②了解数据处理与数据管理的发展
- ③掌握常用的数据模型及其设计



知识结构



在数据处理这一计算机应用领域，人们首先遇到的基本概念是信息和数据，它们是两个不同的术语，却有着不可分割的联系。

1.1.1 数据与信息

数据是计算机存储的基本对象，它是记录下来描述事物状态或属性的物理符号。这种物理符号的表现形式不仅仅是数字（狭义的数据理解）、文字和其他特殊字符，还包括图形、图像、声音等多媒体数据，这些数据都可以经数字化后存入计算机。

日常生活中，为了了解、研究事物和相互交流，需要对事物进行描述。通常我们采用自然语言描述事物。然而非形式化的自然语言，不便于用计算机进行存储和处理。为此，人们常常用抽象的符号序列对其所研究事物的某些方面的特征或属性进行描述，并记录下来。这些记录下来的描述事物特征或属性的物理符号就是数据。例如，如下数据描述一个银行账户（张明，4367420280010532818，4350.00），这样一行数据我们称为一条记录。仅从这行数据是无法得知其确切意义的，但如果知道这行记录每个数据项的确切含义，我们就得到如下信息：张明的银行帐号是 4367420280010532818，余额是 4350 元。数据有一定的格式，例如，银行帐号的长度是 19 位数字，账户余额小数点位数是 2 位。这些格式的规定就是数据的语法，而数据的含义就是数据的语义。通过解释、分析归纳、演绎推导等手段从数据中抽取对人们有价值、有意义的数据，称为信息。因此，数据是信息的载体，是信息存在的形式，信息要依靠数据表达，数据是原始事实，它必须通过解释或处理之后才能成为有用的信息。

1.1.2 数据处理

简单地讲，数据处理就是将数据转换成信息的过程。广义地讲，数据处理是对各种类型的数据进行收集、整理、存储、分类、加工、检索、维护、统计和传播等一系列活动的总称。数据处理的目的是从大量的、原始的已知数据出发，根据事物之间的联系和运动规律，通过分析归纳、演绎推导等手段从中抽取对人们有价值、有意义的数据，即信息，并以此信息作为个人行为和企业决策的依据。由此可见，信息是被加工成特定形式的数据，这种数据对数据的接收者是有意义的。数据的加工可以比较简单，也可以相当复杂。简单加工包括分类、排序等；复杂加工甚至要用统计学、数学模型、人工智能对数据进行深层次的加工。据统计 80% 以上的计算机主要用于数据处理，可见数据处理是计算机的主要应用领域。

在数据处理过程中数据与信息的概念表现出了相对性。数据是数据处理“原料”，是输入，而信息则是“产品”，是输出。如果将每个数据处理过程作为数据处理流水线的一个加工工序，即两个或两个以上数据处理过程前后相继，前一数据处理过程的输出信息相对于后一数据处理则是二次数据。例如，某一学生的“出生日期”是该学生不可改变的基本数据，属于原始数据，而年龄则是经加工得到的数据，属于二次数据。

1.1.3 数据处理的发展阶段

与其他任何事物一样，计算机数据处理也经历了产生并由低级到高级的发展过程。使用计算机进行数据处理从 20 世纪 50 年代初开始，先后经历了电子数据处理系统（EDP），管理信息系统（MIS）和决策支持系统（DSS）三个阶段。

电子数据处理系统（Electronic Data Processing，EDP）是以文件系统为基础，对某一部门内的数据进行收集、分类、加工处理的信息系统。它一般没有管理决策功能，但却是管理信息系统和决策支持系统的共同基础。

管理信息系统（Management Information System，MIS）是以数据库技术与网络技术为工具，对生产、经营和管理全过程进行预测、管理、调节和控制的信息系统。它主要面向确定型、结构化的决策。

决策支持系统（Decision Support System，DSS）是以数据库、模型库和方法库为基础，面向半结构化或非结构化决策问题的信息系统。通过充分的人机交互，帮助管理决策者进行多方案综合比较分析，提高决策水平，改善运营策略的正确性与有效性。

1.2

数据处理与数据管理

数据处理是以大量原始数据为基础得到信息的，因此如何有效地管理大量的数据成为数据处理的基本问题。可以说，数据管理是数据处理的核心问题。数据管理是指对数据进行组织、分类、编码、存储、检索、维护。它是数据处理的基本环节，而且是任何数据处理过程中必不可少的共有部分。随着计算机硬件、软件技术的发展及计算机应用范围的拓展，计算机数据管理技术经历了人工管理阶段、文件系统阶段和数据库系统阶段。

1.2.1 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。当时在硬件方面，外存储器只有纸带、卡片、磁带，没有磁盘这样的直接存储设备。软件方面，没有专门管理数据的软件，数据由处理它的程序自行携带，数据处理方式是批处理。在人工管理方式下数据与程序不具有独立性，一组数据对应一组程序。当数据的类型、格式、输入输出方式改变时，应用程序必须做相应修改；同时由于数据与程序直接相关，在一个程序中使用的数据，无法被其他程序共享使用，程序之间存在大量重复数据。

1.2.2 文件系统阶段

从 20 世纪 50 年代后期到 60 年代中期，操作系统有了专门管理数据的软件，称为文件系统。在文件系统管理方式下，数据以文件形式长期保存在外存上，通过文件系统提供的数据管理功能和程序与数据之间的存取方法，使数据与程序之间具有一定的独立性，因此比人工管理阶段前进了一步。然而，文件仍然是面向应用程序的，在文件系统中数据存取的基本单位是记录，记录内各个数据项的逻辑结构必须由应用程序处理后，才能访问其中的数据项。这一点使得各应用程序之间即使有大部分的相同数据项时，也必须建立各自的文件，而

不能共享数据，因此，程序之间仍存在大量重复数据，即有数据冗余。同时文件系统对文件的并发访问支持并不完善，因此，无法满足多用户联机实时处理的要求。

1.2.3 数据库系统阶段

从 20 世纪 60 年代后期开始，计算机的应用规模更加庞大，数据量急剧增长，联机实时处理对共享数据的要求越来越高。此时，文件系统已不能完全满足用户在数据管理上的要求。在这种背景下，数据库作为新的数据管理技术应运而生。数据库系统有效地克服了文件系统的缺陷，提供了对数据更高级、更有效的管理，提高了数据的一致性、完整性，减少了数据冗余。

1.3

数据模型

数据是对事物的描述，现实世界中的事物不但有其本身的属性和状态，而且事物之间也是相互联系的。同样，数据处理过程中需要对其所处理的数据建模，以便使数据能比较真实地反映现实世界中的事物的属性、状态及其联系。然而由于计算机不具有对现实世界中事物及事物之间相互联系的自主认识能力，因此，计算机不可能直接感知、认识和反映现实世界的具体事物。这一矛盾的解决必须借助于人们建立数据模型，即人们把自身对现实世界的理解和认识通过数据模型这一工具“传授”给计算机。基于上述原因，数据模型一般应满足三个要求。第一是数据模型要能比较真实地模拟现实世界；第二是数据模型要容易被人们理解；第三是数据模型要能够很方便地在计算机上实现。由于一种模型很难同时满足这三方面的要求，所以在数据建模中根据模型应用的不同目的，将模型分为不同层次的两大类。即概念层数据模型和组织层数据模型。

1. 概念模型

概念层数据模型称为概念模型或信息模型，它是从数据的应用语义的角度来抽取模型并按用户的观点对数据和信息建模。由于这类模型容易被人们理解，因此，它主要用于数据库的设计阶段，它与具体的数据库管理系统无关。

2. 逻辑模型

组织层数据模型称为逻辑模型，它从数据的组织层来描述数据并按计算机系统的观点对数据建模，它指明我们采用什么样的数据逻辑结构来组织数据。数据的逻辑模型有层次模型、网状模型及关系模型。具体的数据管理系统都是基于以上三种逻辑模型中的一种。因此，逻辑模型与所使用的数据库管理系统相关，同时逻辑模型也易于在计算机上实现。

为了使在某一具体数据管理系统管理下的数据模型比较真实地模拟现实世界，应首先将现实世界抽象为信息世界，也就是说，将现实世界中的事物及其联系抽象为某一种信息结构，这种结构反映了人们对现实世界的认识和理解，它不依赖于具体的计算机系统和数据库管理系统，而是概念层次的数据模型，也就是上述的概念模型。然后再把概念模型转换成计算机上数据库管理系统支持的逻辑模型。从现实世界到信息世界是通过人的“抽象”思维来完成的，其中体现了人对现实世界的认识和理解，从概念模型到逻辑模型使用的是“转换”，也就是说，仅作形式上的变化，而其中的信息结构并不变化。因此，得到数据模型的过程是

先有概念模型，后有逻辑模型，概念模型是逻辑模型的基础。如果在获得概念模型的抽象过程中，人的认知符合现实世界，那么，最终得到的逻辑数据模型就能较好地满足数据模型的三个要求。

1.3.1 概念模型与 E-R 图

从概念模型与逻辑模型的关系可知，概念模型是数据库设计过程的一个中间层次，也是一个关键环节，它决定了最终得到的数据模型是否能真实地反映现实世界。它既是数据库设计人员的设计工具，也是数据库设计人员与用户交流的工具。现在采用的概念模型主要是实体 - 联系（Entity Relationship，ER）模型。表示概念模型的工具是由 P.P.Chen 于 1976 年提出的 E-R 图，使用 E-R 图描述的结果称为 E-R 模型。E-R 模型是建立于对现实世界的这样一种认识：现实世界由一组称为实体的基本对象以及这些对象间的联系构成。

1. 实体

客观存在并且可以相互区别的事物称为实体。如一名学生，一件商品，一份订单；也可以是抽象的事件，如一次比赛，一笔交易。

一个数据管理软件中往往存储许多类似的实体数据。例如，某一超市中有多位供货商，需要在数据库中存储多位供货商的数据，而所有供货商的数据结构都是相同的，如都由名称、地址、电话联系人这些数据组成。因此，这些供货商实体都具有相同的属性，但对于不同的供货商，这些属性的值不同。我们把具有相同属性的一类实体抽象为一个实体型（Entity Type）。实体型是用实体型的名字和一组属性定义。例如，“供货商（供货商号，名称，地址，电话，联系人）”就是一个实体型，实体型所表示的实体集合中的任一实体称为该实体型的实例，简称实体。同型实体的集合称为实体集（Entity Set）。例如，全体供货商就是一个实体集。

在 E-R 图中，用矩形表示实体型，矩形框内写明实体名。如图 1-1 所示。

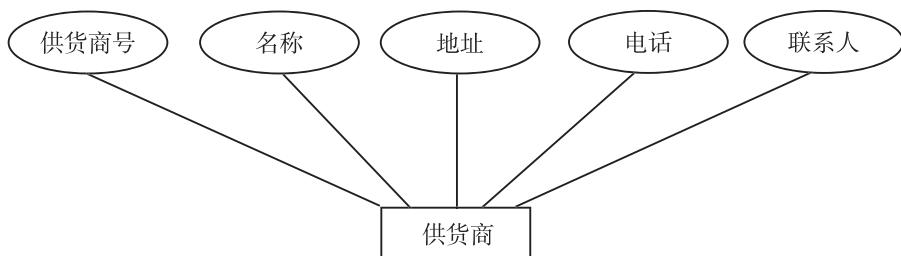


图 1-1 供货商实体及其属性联系

2. 属性

实体的某一特征称为属性（Attributes）。如供货商有供货商号、名称、地址、电话、联系人等方面的属性。

属性有“型”和“值”的区别，属性名是属性的型，如“（供货商号，名称，地址，电话，联系人）”就是属性的型，它规定了数据的语义。值是属性的具体内容，如“（000010，三环乳制品公司，向阳街，87623891，姜文波）”是对应上述属性的值。

在 E-R 图中，实体的属性用椭圆表示，椭圆内写明属性名，并用无向边将实体与其属性连接起来。

3. 联系 (Relationship)

现实世界中，事物内部及事物之间是相互联系的。这种联系必然要反映到信息世界中，事物内部的联系和事物之间的联系对应地在信息世界中实体联系被抽象为实体型内部的联系和实体型之间的联系。实体型内部的联系通常是指反映为组成实体型的各属性之间的联系。实体型之间的联系通常是指不同实体集中实体的对应关系。

两个不同实体型之间的联系有以下三种情况。

(1) 一对一联系 (1:1)。实体集 A 中的一个实体最多与实体集 B 中的一个实体相对应 (相联系)，反之亦然，则称实体集 A 与实体集 B 的联系为一对一的联系，记作 1:1。例如，一个部门中只有一名员工做主管，如果企业规定主管不能兼任，部门与员工之间的管理联系则为一对一联系，如图 1-2 (a) 所示。

(2) 一对多联系 (1:n)。实体集 A 中的一个实体与实体集 B 中的多个实体相对应，反过来，实体集 B 中的一个实体最多与实体集 A 中的一个实体相对应，记作 1:n。例如，一个部门中有多名员工，一名员工只属于一个部门，部门与员工之间的从属联系则为一对多联系，如图 1-2 (b) 所示。

(3) 多对多联系 (m:n)。实体集 A 中的一个实体与实体集 B 中的多个实体相对应，而实体集 B 中的一个实体又与实体集 A 的多个实体相对应，记作 m:n。例如，一门课程同时可以有多个学生选修，而每个学生又可以同时选修多门课程，学生与课程之间的选课联系则为多对多联系，如图 1-2 (c) 所示。

在以上三种联系中，一对一联系是一对多联系的特例，一对多联系是多对多联系的特例。实际上，两个以上不同实体型也可以存在这三种联系，例如，采购联系将采购员、商品与供货商三个实体联系起来，如图 1-3 所示；同一实体集内的各实体之间也可以存在三种联系，例如，在员工实体中，一名作为部门主管的员工管理多名同部门的其他员工，如图 1-4 所示。

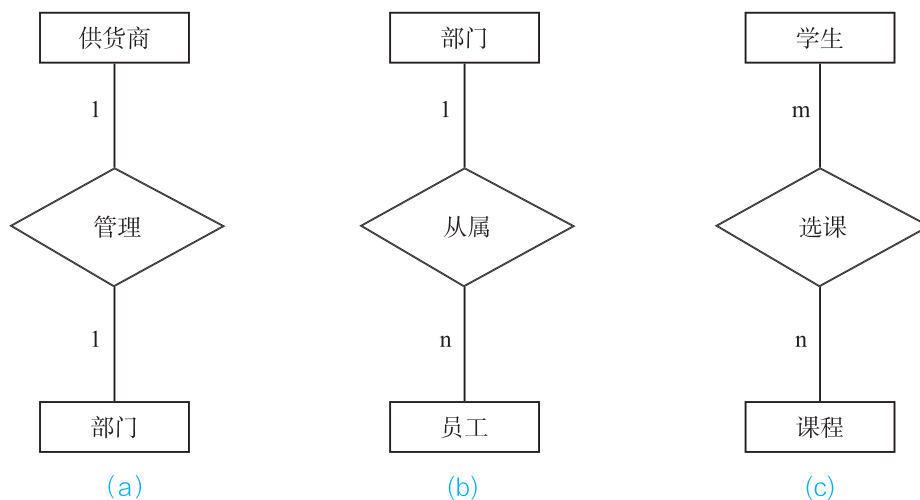


图 1-2 两个实体之间的联系

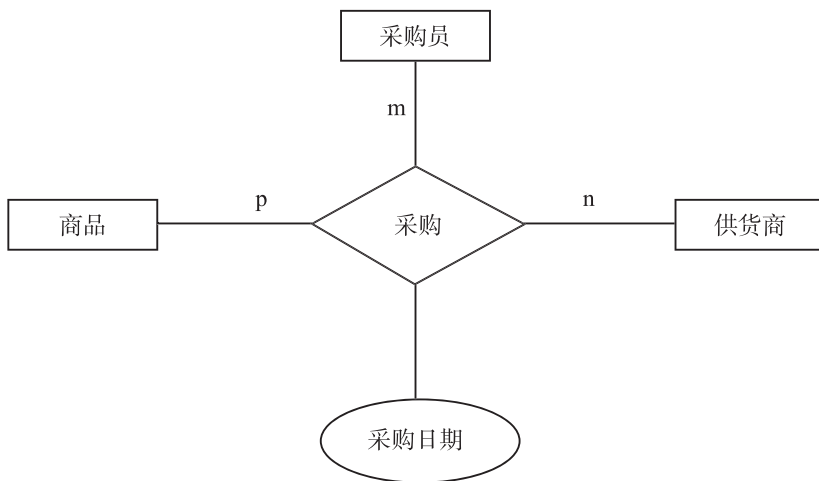


图 1-3 多个实体联系

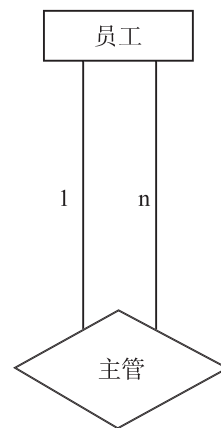


图 1-4 同一实体联系

在 E-R 图中，联系用菱形表示，菱形框内写明联系名，并用无向边与有关实体型连接，同时，在无向边旁标明联系的类型。如果联系具有属性，也要把属性和联系用无向边连接上。

1.3.2 逻辑数据模型

概念模型接近现实世界，比较抽象，不便于计算机的实现，因此，还必须将其转换成计算机能够处理的逻辑数据模型。逻辑数据模型有层次模型、网状模型和关系模型。这些数据模型按其存储数据的逻辑结构来命名。层次模型、网状模型又称为非关系模型。三种模型的根本区别在于数据结构不同，即数据之间联系的表示方式不同。由于层次模型与网状模型已很少使用，在此不再赘述。

1. 关系模型

关系模型是发展较晚的一种数据模型。一个关系模型的逻辑结构是一张规范的二维表。如表 1-1 所示。

关系模型是建立在严格的数学概念基础上的。关系模型的概念单一，无论实体还是实体之间的联系都用关系来表示，对数据进行检索的结果也是关系（即表）。关系模型具有结构简单、清晰，用户易懂易用的优点，因此，关系模型是当今数据管理中主要的数据模型。

表 1-1 员工关系

员工号	姓名	性别	岗位	基本工资	部门号
000010	张振国	男	经理	3500.00	000010
000020	张丽	女	收银员	2000.00	000030
000030	刘强	男	会计	3500.00	000010
000040	向秀丽	女	主管	6000.00	000030
000050	李文婷	女	收银员	3000.00	000030
000060	王卫东	男	采购员	4500.00	000020
000070	郑小娟	女	主管	5000.00	000020
000080	赵治军	男	主管	5000.00	000040

2. Excel 数据清单与关系模型

Excel 在数据组织、数据管理、数据计算、数据分析等方面都具有强大的功能。使用 Excel 可以用简洁的方式对数据进行类似于数据库的管理。

Excel 中用数据清单规范数据工作表以满足关系数据模型的要求，数据清单是指工作表中这样的一个区域：

- (1) 数据清单是工作表中的一个矩形表格区域；
- (2) 数据清单的第一行一定是标题行，标题行内是各列数据的字段名；
- (3) 标题行下面的每一行是一条对应各字段的数据记录。

为了使表格中的数据符合关系模型的要求，有利于 Excel 检测和选定数据清单，用户在建立数据清单时，必须按照以下原则：

- (1) 避免在一个工作表中建立多个数据清单；
- (2) 数据清单中不能有空白的行和空白的列；
- (3) 在同一工作表中的数据清单应该与其他数据间至少留出一个空行或空列。

3. 关系模型的基本概念

(1) 关系：一个关系就是一张规范的二维表，每个关系都有一个关系名。在 Excel 中，又称其为数据清单，在这个数据清单中包含数据或数据间的联系。

(2) 关系模式：对关系结构的描述称为关系模式。一个关系模式对应一个关系的结构。关系模式可记作：

关系名 (属性名 1, 属性名 2, …, 属性名 n)

例如，表 1-1 所示的员工关系的关系模式可记作：

员工 (员工号, 姓名, 性别, 岗位, 基本工资, 部门号)

在 Excel 中，为了方便对数据清单的管理，通常用一个工作表存储一个数据清单的数据，这样 Excel 中的数据清单可以使用“工作表名 (字段名 1, 字段名 2, …, 字段名 n)”表示。数据库设计人员设计得出 E-R 模型后，为方便将 E-R 模型表示成关系，设计人员通过关系模式来描述转换得到的关系，最后用工作表中的数据清单体现关系。

(3) 元组：关系中的一行称为一个元组。它实际上是对一个具体实体的描述。例如“(000010, 张振国, 男, 经理, 3500.00, 000010)”就是一个元组，它是张振国这名员工的描述信息。表 1-1 所示的员工关系共有 11 个元组。元组在 Excel 数据清单中被称为记录。

(4) 属性：关系中的列称为属性，实体常常有多个属性，体现在关系中就有多个列。例如，表 1-1 所示的员工关系的属性有员工号、姓名、性别、岗位、基本工资、部门号。属性在 Excel 数据清单中称为字段。

(5) 域：属性的取值范围。例如性别属性的域为“男”和“女”。

(6) 分量：元组中的每个属性值称为元组的一个分量。例如，元组 (000010, 张振国, 男, 经理, 3500.00, 000010) 有 6 个分量，对应“姓名”属性的分量是“张振国”。分量在 Excel 数据清单中对应某条记录中一个字段或单元格的值。

(7) 候选键与主键：能唯一标识元组且不包括多余属性的最小属性组合称为关系的候选键。一个候选键可以由一个属性组成，也可以由一组属性组成。例如，在员工关系中，员工号是唯一的，因此，员工号是候选键。再例如，在表 1-2 的订单明细关系中仅用“订单号”

或仅用“明细项号”无法区分各元组，而“订单号”与“明细项号”组合在一起才能唯一标识一条订单明细项，因此，“(订单号，明细项号)”是订单明细关系候选键。一个关系中可以有多个候选键，其中，有一个被频繁使用的候选键被指定为主键。

表 1-2 订单明细关系

订 单 号	明 细 项 号	商 品 号	数 量	订 货 价
100602001	1	000080	100	10
100602002	1	000060	10	52
100807001	1	000070	200	3.5
100807001	2	000030	50	1.1
100807002	1	000040	20	7

(8) 主属性：候选键中的各属性称为主属性。例如订单明细关系中的“订单号”和“明细项号”都是主属性。

4. 关系的特点

关系是一张规范的二维表，也就是说，一张二维表要成为关系则必须满足一定的要求。最基本的要求是关系中的每个属性必须是不可再分的基本数据单元，即表中不能有表。除此之外还有以下四个特点。

- (1) 在同一个关系中不能出现相同的属性名，即不允许同一个数据表中有重名字段。
- (2) 关系中不允许有完全相同的元组，即一个数据表中没有重复行。
- (3) 在一个关系中元组的次序无关紧要。
- (4) 在一个关系中列的次序无关紧要。

1.4

商品进销存系统的数据库设计

我们将通过案例的形式学习和使用 Excel VBA 开发一个实用的数据库应用系统。本书将以某商业企业的商品进销存系统为案例，讲解一个商品进销存数据处理系统的设计与实现过程。本节主要讲述该系统的数据模型设计过程。

1.4.1 商品进销存系统背景介绍

某仓储式零售业企业提供商品零售服务，该企业设置了管理部、营业部、采购部等部门以保证日常经营活动。每个部门按部门职责设置必要的岗位来完成具体的业务活动。该企业的主要业务环节为商品的采购与销售。在采购业务中，由采购员负责与供货商签订订单，供货商按订单上要求的供货日期提供货物，并由采购员安排入库保管；在销售业务中，收银员按顾客选购的货物开具发票，并收取货款。为保证供货商按时供货履约，该企业规定：只有订单商品入库后，才能由会计记账并由出纳支付订单货款。为了及时了解企业的经营情况，以便作出正确的订货和销售策略，由企业经理对经营情况进行核算，计算出企业的利润、商品库存、应付货款等数据。这些数据的计算方法如下：

商品库存量 = 商品进货量 - 商品销售量

商品进货平均单价 = 商品进货额 ÷ 商品进货量

商品库存成本 = 商品进货平均单价 × 商品库存量

商品利润 = 商品销售额 - 商品进货额 + 商品库存成本

其中的商品进货额是将所有订单中涉及该商品的明细项的进货量与订货价相乘后再累加得出的，其中的该商品进货量、销售量，销售额也是采用累加的方法得到的。而这时仅仅是得到了某一商品的统计数据。而要得到整个企业的相关统计数据则还需要将各商品的相关数据进行汇总。对一个正常经营的企业来说，这些数据量是相当大的。如果依靠有限人力来完成这些工作，是无法及时完成的，同时也是不经济的，因此必须使用数据处理系统才能完成任务。

1.4.2 商品进销存系统功能介绍

按照上述的系统背景的要求，商品进销存系统应具有以下功能：为方便对企业员工的管理，系统应支持对企业中的部门、员工的管理；为方便对进销业务的管理，系统应支持对商品、供应货、签订订单、订单入库、订单付款、销售发票的业务活动管理；为方便了解和掌握商品库存量，系统应设置对商品库存的查询功能；为方便经理对相关业务经营情况的掌握，系统还应设置业务查询、统计报表等功能。该系统的所有功能如表 1-3 所示。

表 1-3 商品进销存系统的主要功能

系 统 功 能		系 统 功 能	
基础数据维护	商品目录	查 询	订单查询
	部门设置		发票查询
	员工管理		经营成果查询
	供应商管理	统计报表	商品销售报表
日 常 交 易	销售发票		员工销售报表
	采购订单		应付款报表
	采购入库		库存及销售利润报表
	订单付款	系统功能	系统登录
			退出

1.4.3 商品进销存系统数据模型设计

数据处理系统的数据模型设计至关重要，数据模型设计合理与否关系到 Excel 数据表中的数据是否真实地反映企业的业务运营情况。因此，我们必须按前面所讲的过程设计数据库。首先设计 E-R 模型，然后将 E-R 模型转换为关系数据模型。

1.E-R 模型

根据 1.4.2 中所述的系统功能要求，在数据库中必须有部门、员工、订单、发票、商品、供货商，发票明细、订单明细这些实体。在仔细分析各实体之间的关系后，我们得到的 E-R 模型如图 1-5 所示。

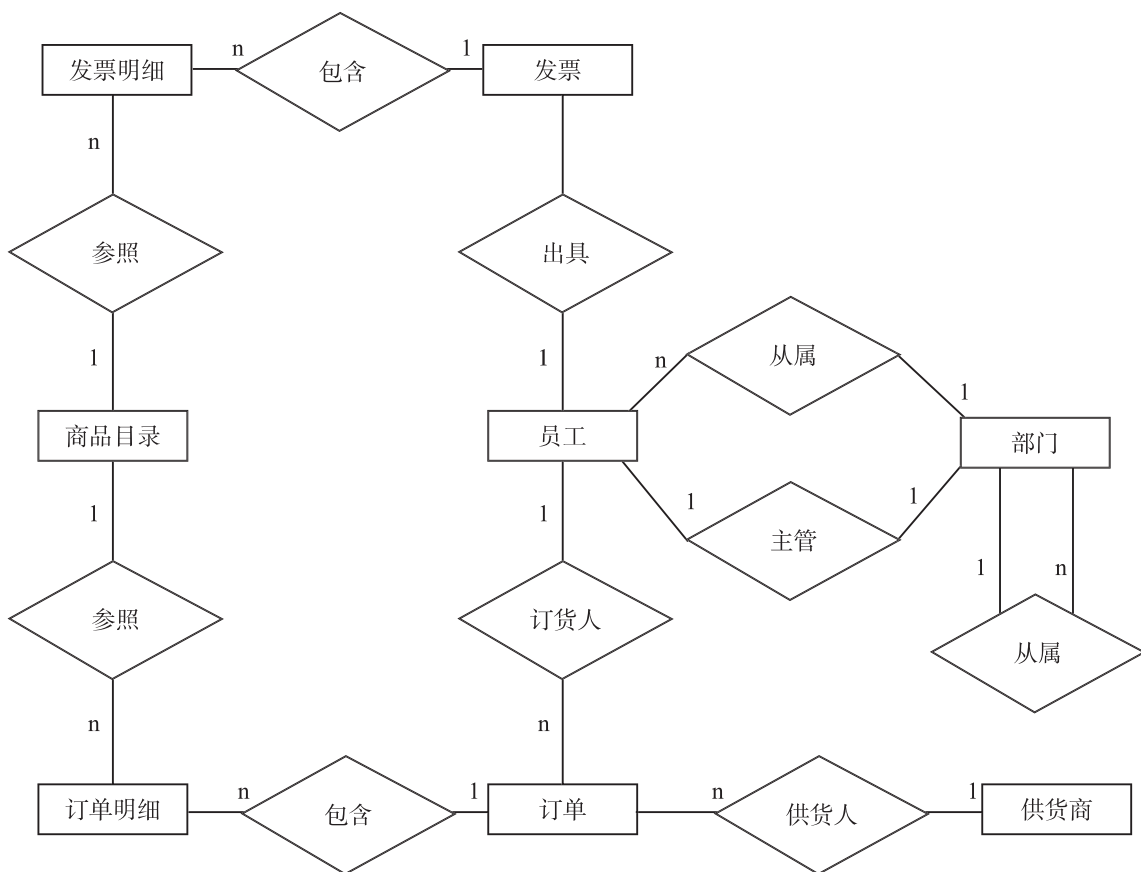


图 1-5 商品进销存系统数据库 E-R 模型

2. 关系模型

根据 E-R 模型和业务数据处理的要求，我们确定每个实体的属性后，再将其中的实体转换成关系，得到如下的基本关系模式。

部门（部门号，名称，经理号，上级部门，办公地点）
 员工（员工号，姓名，性别，岗位，基本工资，部门号，出生日期，简历，婚否，照片，登录密码）
 采购订单（订单号，订单日期，供货人，订货人，已入库，入库日期已付款，付款日期供货日期）
 订单明细（订单号，明细项号，商品号，数量，订货价）
 发票（发票号，日期，收款人）
 发票明细（发票号，明细项号，商品号，数量，售价）
 供货商（供货商号，名称，地址，电话，联系人）
 商品目录（商品号，品名，类别，售价，单位）

这些基本关系模式是构成商品进销存数据处理系统的数据模型，它将是我们的 Excel 工作簿及其数据表的依据。

本章小结

数据处理系统是当代计算机系统的重要组成部分。本章通过对数据处理技术发展的介绍，阐述了数据处理技术和数据管理技术产生和发展的过程。

由于计算机不能直接处理现实世界的具体事物，所以必须事先对现实世界的事物抽象，转换成计算机能处理的数据模型，从概念模型到逻辑数据模型，是数据库设计的重要方法。在概念模型中，介绍了基于 E-R 图的概念模型的表示方法。在数据模型中，详细介绍了关系

模型。最后本章以商品进销存系统为例，介绍了该系统中关系数据模型的建模过程。



一、单选题

1. 数据处理技术发展的三个阶段中，没有管理决策功能的是（ ）。
A. 管理信息系统 B. 决策支持系统
C. 电子数据处理系统 D. 数据库系统
2. 在数据管理技术发展的阶段中，没有专门的软件对数据进行管理的是（ ）。
A. 人工管理阶段 B. 文件系统阶段
C. 数据库阶段 D. 电子数据处理阶段
3. 数据处理的核心是（ ）。
A. 编译系统 B. 数据库
C. 操作系统 D. 数据管理
4. 概念模型是从（ ）观点对数据和信息建模。
A. 计算机系统 B. 文件系统 C. 程序设计人员 D. 用户
5. 数据库的概念设计中，最常用的数据模型是（ ）。
A. 形象模型 B. 实体联系模型 C. 物理模型 D. 逻辑模型
6. 以下内容，哪一项不是概念模型中的实体（ ）。
A. 一件商品 B. 一场考试 C. 一个员工号 D. 以上都是
7. 在 E-R 图中，用来表示属性的图形是（ ）。
A. 矩形 B. 椭圆形 C. 菱形 D. 平行四边形
8. 实体之间联系的方式有（ ）。
A. 1 种 B. 2 种 C. 3 种 D. 4 种
9. 常见的数据模型有三种，它们是（ ）。
A. 网状、关系和语义 B. 层次、关系和网状
C. 环状、层次和关系 D. 字段名、字段类型和记录
10. Excel 中，数据清单的（ ）对应数据库中的字段。
A. 行 B. 行标志 C. 列 D. 列标志

二、填空题

1. 数据处理技术经历了_____、_____和_____三个阶段。
2. 数据处理的核心是_____。
3. 根据数据模型的应用目的不同，数据模型分为_____和_____。
4. E-R 图是表示_____的方法。
5. 逻辑数据模型可分为_____、_____和_____。
6. Excel 支持的数据模型是_____。
7. 能唯一标识元组且不包括多余属性的最小属性组合称为关系的_____。

三、简答题

1. 简述信息与数据的区别和联系。
2. 试说明关系、元组、属性、域之间的关系。

3. 简述候选键、主键的联系与区别。

4. 为某图书馆设计一个图书借阅管理系统数据库的关系模型，为简单处理，系统中只涉及对读者、图书的管理。要求该系统能够记录读者的借阅情况以及读者和图书的基本情况。具体要求为：

（1）每本书有书号、书名、作者、附光盘、出版社、出版时间、单价、册数、简介、封面、借出数这些属性。其中书号是区分不同书籍的唯一标识。

（2）每位读者有编号、姓名、单位、类型四个属性。其中每位读者的编号是唯一的。读者分为教师、教工、学生三种类型，借书期限分别为 90 天、45 天、30 天。

（3）除上述要求外，当读者借书时需要记录其借书的日期以方便还书时计算是否超期及超期天数。

请先画出上述系统的 E-R 图，然后将该 E-R 图转换为关系模式。