

21世纪高等学校规划教材 | 计算机科学与技术



数据库原理及应用

(SQL Server) (第3版)

李俊山 叶霞 罗蓉 刘东 编著



清华大学出版社

数据库原理及应用

(SQL Server)



李俊山 叶霞 罗蓉 刘东 编著

清华大学出版社
北京

内 容 简 介

本书基于“数据库原理 + SQL Server 数据库 + ADO.NET 数据库访问技术 + VB.NET 主语言”架构及其内容体系,全面、系统地介绍了数据库系统的基本概念、基本原理、基本技术和基本设计方法。全书共分 11 章,内容包括数据库系统概述、关系运算、数据库应用系统设计方法、SQL Server 2012 数据库与数据表的创建、关系数据库语言 SQL、关系数据库模式的规范化设计、Transact-SQL 与存储过程、数据库应用系统体系结构与访问技术、数据库应用系统设计与实现、数据库保护技术、数据库技术新发展等。内容不仅覆盖了关系数据库原理、数据库应用系统设计方法和技术,而且较全面地介绍了利用 ADO.NET 访问数据库的过程和方法,介绍了 VB.NET 程序设计基础和 VB.NET 绑定数据方法;给出了一个完整的教学管理数据库应用系统的设计案例及程序源码,具有鲜明的理论与应用紧密结合特色。

本书内容上兼顾了数据库课程教学的应用型和工程型要求,内容选材新颖精练,理论与实践并重,针对性与系统性强。语言精练,表述通俗,图文并茂;案例驱动,配有大量习题,便于复习、练习和综合设计实践。

本书可作为高等院校本科计算机科学与技术、软件工程、网络工程、信息安全、物联网工程、数字媒体技术、信息工程、信息与计算科学、信息管理与信息系统、地理信息系统、电子商务等专业的数据库课程教材,也可作为电气类专业研究生的数据库课程教材,还可供相关领域专业的高等院校师生、从事计算机软硬件研究和信息系统设计的科技人员和工程技术人员参考。

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。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据库原理及应用:SQL Server/李俊山等编著.—3 版.—北京:清华大学出版社,2017
(21 世纪高等学校规划教材·计算机科学与技术)
ISBN 978-7-302-47321-3

I. ①数… II. ①李… III. ①关系数据库系统 IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2017)第 124471 号

责任编辑:魏江江 薛 阳

封面设计:傅瑞学

责任校对:梁 毅

责任印制:李红英

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:010-62770175 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

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:清华大学印刷厂

经 销:全国新华书店

开 本:185mm×260mm 印 张:26.25 字 数:641 千字

版 次:2009 年 6 月第 1 版 2017 年 9 月第 3 版 印 次:2017 年 9 月第 1 次印刷

印 数:23001~25000

定 价:59.00 元

产品编号:074840-01

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与应用。

(7) 21 世纪高等学校规划教材·电子商务。

(8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail:weiijj@tup.tsinghua.edu.cn

第3版前言

随着信息技术的迅猛发展和信息化社会水平的进一步提升,数据库技术已经成为国家信息基础设施和信息化社会中的最重要的支撑技术之一;基于数据库技术和数据库管理系统(DBMS)的应用软件和装备制造系统中的控制软件的研发,已经成为计算机及其相关专业领域技术和管理人员的必备知识和基本技能。数据库技术已在国民经济的各个领域得到了十分广泛的应用,并在推动科技发展和社会进步方面起着越来越重要的作用。

本书第1版和第2版自2009年和2012年出版以来,在几十多所院校的计算机类专业、电子信息类专业、管理类专业和其他相关专业的教学中得到了采用,许多学生、教师和读者对本书的进一步改版都给予了特别的关心,并提出了许多宝贵的建议。出版第3版的原因一是将第2版中基于SQL Server 2005的概念、平台软件环境介绍、SQL Server 语句练习、数据库应用系统设计应用等方面的内容,都改写成基于SQL Server 2012,以适应数据库软件的版本更新。二是完善和优化第2版中的部分内容和实例,特别是进一步完善和优化了基于SQL Server 2012的数据库应用系统设计案例的设计过程及程序代码,进一步突出了面向综合设计能力培养的主题,提高了对学生学习和设计过程的指导性。三是删除了一些不重要的内容。

本教材在内容选取上兼顾了应用型人才和工程型人才培养的要求,构建了以“数据库原理+SQL Server数据库+ADO.NET数据库访问技术+VB.NET主语言”为架构的数据库课程内容体系,具有理论与应用结合的鲜明特色,教学参考及应用价值高。

数据库课程理论抽象,实践性强,对数据库整体架构理解的要求高。本教材在编写手法上旨在努力贯彻“深、活、透、新”四字教学要求,即把数据库设计理论讲深,把数据库的SQL语言讲活,把数据库设计方法讲透,把数据库应用技术讲新,以便使课程教学效果力争达到:在讲解数据库系统的整体概念时能让学生自己搭起数据库的总体架子,在讲解数据库设计理论时能让学生穿起定义、定理之间的链子,在讲解数据库的保护技术时能让学生厘清它们内部的口子,在讲解数据库应用系统设计时能让学生掌握与实际结合的路子,在讲解数据库新技术时能给学生指出新技术未来的影子。

本教材第3版共分11章,第1章是数据库系统概述,第2章介绍关系运算,第3章是数据库应用系统设计方法,第4章介绍SQL Server 2012数据库与数据表的创建,第5章是关系数据库语言SQL,第6章介绍关系数据库模式的规范化设计,第7章是Transact-SQL与存储过程,第8章介绍数据库应用系统体系结构与访问技术,第9章是数据库应用系统设计与实现,第10章介绍数据库保护技术,第11章是数据库技术新发展。内容覆盖了关系数据库原理、数据库应用系统设计方法和设计技术。

本书可作为高等院校本科计算机科学与技术、软件工程、网络工程、信息安全、物联网工程、数字媒体技术、信息工程、信息与计算科学、信息管理与信息系统、地理信息系统、电子商务等专业的数据库课程教材,也可作为相关专业研究生的数据库课程教材,还可供相关领域

专业的高等院校师生、从事计算机软件研究和信息系统设计的科技人员和工程技术人员参考。

本书的第1章、第2章、第3章、第6章、第10章由李俊山编写,第9章、附录A、附录B、附录C由叶霞编写,第5章、第8章由李俊山和叶霞编写,第4章由罗蓉编写,第7章由李俊山和罗蓉编写,第11章由李俊山和刘东编写。另外,李建华、赵方舟、杨威、张娇、杨亚威等参与了部分资料整理和早期版本的数据库应用系统案例程序的编写。

由于作者水平有限,书中难免有不当之处,敬请广大读者和专家批评指正。

李俊山

2017年3月

E-mail: lijunshan403@163.com

目 录

第 1 章 数据库系统概述	1
1.1 数据与数据管理	1
1.1.1 信息与数据	1
1.1.2 数据管理技术的发展	3
1.2 数据库系统组成	4
1.2.1 数据库	4
1.2.2 数据库管理系统	5
1.2.3 数据库应用系统	6
1.3 数据描述与数据模型	7
1.3.1 现实世界的的数据描述	7
1.3.2 数据模型	8
1.3.3 层次模型	8
1.3.4 网状模型	9
1.4 关系模型	10
1.4.1 关系模型的基本概念	10
1.4.2 关系的键与关系的属性	14
1.5 数据库的内部体系结构	15
1.5.1 数据库内部体系结构中的三级模式结构	16
1.5.2 数据库内部体系结构中的两级映像与数据独立性	19
1.5.3 数据库内部体系结构的概念	21
习题 1	21
第 2 章 关系运算	22
2.1 关系的数学定义	22
2.1.1 笛卡儿积的数学定义	22
2.1.2 关系的数学定义	23
2.2 关系代数	24
2.2.1 基于传统集合理论的关系运算	24
2.2.2 关系代数特有的关系运算	27
2.2.3 关系运算综合示例	30
2.2.4 关系代数运算在关系数据库查询操作中的应用	33
2.3 关系演算	36

2.3.1	元组关系演算	36
2.3.2	域关系演算	39
* 2.4	3种关系运算表达能力的等价性	40
2.4.1	关系演算表达式的安全性约束	40
2.4.2	安全的元组关系演算表达式满足的条件	41
2.4.3	安全的域关系演算表达式满足的条件	42
2.4.4	3种关系运算表达能力的等价性定理	42
习题 2		42
第 3 章	数据库应用系统设计方法	45
3.1	数据库应用系统设计概述	45
3.1.1	数据库应用系统的生命周期	45
3.1.2	数据库应用系统设计方法	46
3.1.3	数据库应用系统研发、管理和使用人员视图级别	47
3.2	用户需求分析	48
3.2.1	用户需求分析过程	48
3.2.2	数据流图及用户业务处理逻辑描述	50
3.2.3	数据字典	52
3.2.4	数据库应用系统的功能需求	53
3.2.5	数据库应用系统环境配置与安全性需求	54
3.3	数据库概念结构设计	55
3.3.1	属性表概念结构设计方法	55
3.3.2	实体与实体集	56
3.3.3	实体集之间的联系及联系集	57
3.3.4	E-R 图设计方法	58
3.3.5	实体-联系模型设计中的一些特殊情况	60
3.3.6	基于 E-R 图的概念结构设计步骤和方法	62
3.4	数据库逻辑结构设计	64
3.4.1	属性表表示的概念结构向关系模型的转换	65
3.4.2	E-R 图表示的概念结构向关系模型的转换	65
3.4.3	关系数据库模式的规范化设计及优化	68
3.5	数据库物理结构设计	69
3.5.1	数据库的物理文件与数据表	69
3.5.2	数据库物理文件的存储位置	70
3.5.3	索引技术(数据表的存储结构与组织方式 1)	71
3.5.4	数据聚簇(数据表的存储结构与组织方式 2)	76
3.5.5	数据库物理结构评价	77
3.6	数据库实现技术简介	78
3.7	数据库应用系统运行与系统维护	79

3.7.1	软件维护	79
3.7.2	运行与维护时期的主要工作	80
习题 3		80
第 4 章	SQL Server 2012 数据库与数据表的创建	82
4.1	SQL Server 2012 关系数据库系统	82
4.1.1	SQL Server 2012 的基础知识	82
4.1.2	SQL Server Management Studio	84
4.1.3	SQL Server 2012 的服务器管理	86
4.1.4	SQL Server 2012 的联机丛书和教程	92
4.2	SQL Server 2012 数据库的存储结构	93
4.3	SQL Server 2012 数据库创建	94
4.3.1	SQL Server 2012 数据库	94
4.3.2	使用 SQL Server Management Studio 创建数据库	95
4.4	创建数据表	98
4.4.1	使用表设计器创建表	98
4.4.2	表结构的修改	102
4.4.3	表的删除	105
4.5	表中数据的插入和更新	105
4.6	创建索引	108
4.6.1	SQL Server 中的索引	108
4.6.2	建立索引的基本原则	109
4.6.3	使用 SQL Server Management Studio 操作索引	110
习题 4		115
第 5 章	关系数据库语言 SQL	116
5.1	SQL 的功能与特点	116
5.1.1	SQL 的功能	116
5.1.2	SQL 的特点	117
5.2	表的基本操作	118
5.2.1	表的创建、修改与撤销	118
5.2.2	表中数据的插入、修改和删除	124
5.3	SQL 的数据查询	126
5.3.1	投影查询	126
5.3.2	选择查询	129
5.3.3	分组查询	131
5.3.4	结果的排序查询	132
5.3.5	模糊查询(字符串匹配)	133
5.3.6	SQL 语言中的常用函数	134

5.3.7	表的连接查询	137
5.3.8	嵌套查询	138
5.3.9	谓词演算查询	139
5.3.10	并、交、差运算查询	143
5.4	SQL 的视图操作	145
5.4.1	视图的创建	146
5.4.2	在视图上进行查询操作	147
5.4.3	用户视图对数据库应用系统设计及系统性能带来的影响	147
5.5	SQL 中带有子查询的数据更新操作	148
5.5.1	带有子查询的数据插入操作	148
5.5.2	带有子查询条件的数据更新操作	149
5.6	嵌入式 SQL 与游标应用	150
5.6.1	嵌入式 SQL	150
5.6.2	SQL Server 2012 的游标及其使用	151
习题 5		158
第 6 章	关系数据库模式的规范化设计	160
6.1	关系约束与关系模式的表示	160
6.2	对关系模式进行规范化设计的必要性	161
6.3	函数依赖	163
6.3.1	函数依赖的定义	163
6.3.2	具有函数依赖约束的关系模式	164
6.3.3	函数依赖的逻辑蕴涵	165
6.4	函数依赖的公理体系	165
6.4.1	阿姆斯特朗公理	165
6.4.2	阿姆斯特朗公理的推论	166
6.4.3	X 关于 F 的闭包及其计算	167
6.4.4	最小函数依赖集	168
6.5	关系模式的分解	172
6.5.1	关系模式分解的概念	172
6.5.2	保持无损的分解	174
6.5.3	保持依赖的分解	178
6.6	关系模式的规范化	179
6.6.1	候选键的求解方法	179
6.6.2	第一范式	182
6.6.3	第二范式	182
6.6.4	第三范式	184
6.6.5	鲍依斯-柯德范式	186
6.6.6	范式之间的关系和关系模式的规范化	187

6.6.7 向 3NF 的模式分解算法	187
6.7 关系模式的规范化方法小结	189
习题 6	190
第 7 章 Transact-SQL 与存储过程	193
7.1 Transact-SQL 基础	193
7.2 Transact-SQL 的语言要素	195
7.2.1 常量	195
7.2.2 变量	196
7.2.3 表达式与运算符	198
7.2.4 Transact-SQL 函数	199
7.3 Transact-SQL 流程控制语句	203
7.3.1 BEGIN...END 语句	203
7.3.2 IF...ELSE 语句	203
7.3.3 CASE 语句	204
7.3.4 WHILE 语句	204
7.3.5 WAITFOR 语句	205
7.3.6 其他语句	205
7.4 基于 Transact-SQL 的数据库创建与管理应用	206
7.4.1 利用 Transact-SQL 语句创建数据库	206
7.4.2 利用 Transact-SQL 语句管理数据库	208
7.5 存储过程	211
7.5.1 存储过程基础	211
7.5.2 创建存储过程	213
7.5.3 执行存储过程	216
7.5.4 管理存储过程	217
习题 7	221
第 8 章 数据库应用系统体系结构与访问技术	223
8.1 数据库应用系统的体系结构	223
8.1.1 集中式计算模式的数据库应用系统	223
8.1.2 C-S 结构的数据库应用系统	224
8.1.3 B-S 结构的数据库应用系统	227
8.2 ADO.NET 数据库访问接口	228
8.2.1 认识 ADO.NET	229
8.2.2 ADO.NET 常用对象	231
8.2.3 ADO.NET 访问数据库的方法	233
8.3 VB.NET 程序设计基础	235
8.3.1 窗体	236

8.3.2	控件	238
8.3.3	属性	240
8.3.4	事件	241
8.3.5	方法	242
8.4	用 VB.NET 绑定数据	242
8.4.1	VB.NET 数据绑定概念	242
8.4.2	DataGridView 控件绑定数据源	243
8.4.3	用代码方式绑定数据源	251
	习题 8	254
第 9 章	数据库应用系统设计与实现	255
9.1	系统功能分析	255
9.2	数据库结构创建	256
9.3	新建应用系统项目	258
9.4	应用系统设计与实现	260
9.4.1	登录模块	260
9.4.2	学生信息添加功能模块	265
9.4.3	学生信息查询功能模块	271
9.4.4	成绩信息查询功能模块	274
9.4.5	成绩信息维护功能模块	279
9.4.6	主界面模块	286
9.5	系统调试运行	290
	习题 9	295
第 10 章	数据库保护技术	296
10.1	事务机制	296
10.1.1	事务的概念	296
10.1.2	事务的特性	297
10.1.3	事务的类型	298
10.2	数据库的安全性	300
10.2.1	数据库安全的威胁	301
10.2.2	数据库安全控制	302
10.2.3	视图机制	304
10.2.4	审计	305
10.2.5	SQL Server 2012 的安全机制	305
10.3	数据库的完整性	316
10.3.1	域完整性约束	316
10.3.2	实体完整性约束	317
10.3.3	参照完整性约束	318

10.3.4	用户定义完整性约束	320
10.3.5	SQL Server 2012 的完整性约束	321
10.4	数据库恢复	330
10.4.1	数据库的故障分类	330
10.4.2	数据库故障的基本恢复方式	331
10.4.3	恢复策略	333
10.4.4	具有检查点的恢复技术	335
10.4.5	数据库镜像	336
10.4.6	SQL Server 数据库的备份和还原	337
10.5	并发控制	343
10.5.1	数据库并发操作带来的数据不一致性问题	344
10.5.2	锁	345
10.5.3	锁协议	346
10.5.4	封锁带来的问题——活锁与死锁	347
10.5.5	并发调度的可串行性	349
10.5.6	两段锁协议	350
10.5.7	锁的粒度	351
习题 10		353
第 11 章	数据库技术新发展	354
11.1	数据库新技术发展的动因	354
11.1.1	新应用领域对数据库存储、检索和管理技术的需求	354
11.1.2	关系数据库系统的局限性	356
11.1.3	数据库技术新发展	357
11.2	面向对象数据库系统	357
11.2.1	对象的逻辑结构	357
11.2.2	对象间的限制和联系	358
11.2.3	面向对象数据库管理系统的基本要求	359
11.2.4	面向对象数据库管理系统的组成	360
11.2.5	面向对象数据库系统的发展	361
11.3	主动数据库系统	361
11.3.1	主动数据库的知识模型	362
11.3.2	主动数据库系统的组成	363
11.3.3	主动数据库管理系统的功能	364
11.3.4	主动数据库的实现	364
11.3.5	事件知识库的实现	365
11.3.6	事件监视器的实现	365
11.3.7	主动数据库系统的发展	365
11.4	嵌入式数据库	366

11.4.1	嵌入式数据库的基本架构	366
11.4.2	嵌入式数据库的主要特点	366
11.4.3	嵌入式数据库的应用	367
11.5	非关系型数据库 NoSQL	368
11.5.1	Web 2.0 动态网站对数据库性能的需求	368
11.5.2	NoSQL 数据库的概念	369
11.5.3	NoSQL 数据库的特征	369
11.5.4	NoSQL 数据库的分类	370
11.5.5	NoSQL 数据库发展展望	372
11.6	面向应用领域的数据库新技术	373
11.6.1	工程数据库	373
11.6.2	统计数据库	373
11.6.3	空间数据库	374
习题 11		375
附录 A	安装 SQL Server 2012	376
附录 B	安装 Visual Studio 2010	391
附录 C	大学教学管理信息系统案例程序代码	397
参考文献		405

第1章

章

数据库系统概述

数据库技术是计算机科学与技术学科中的一个十分活跃而重要的分支,已形成了一整套较为完整的理论与技术体系,其应用已经遍及国民经济和国防技术的所有领域,并成为国家和军队信息化建设中的关键支撑技术。

本章将从数据及数据管理的概念着手,系统地介绍数据库系统及其设计技术所涉及的基本概念和方法,主要包括数据库系统的基本概念、数据描述与数据模型、关系模型、数据库系统的内部体系结构等,以便对数据库有一个概括的了解。

1.1 数据与数据管理

建立数据库的目的是为数据管理和数据处理提供环境支持,而在讲到数据处理时又必须提及信息处理及其与数据处理的关系。所以下面先从信息和数据的概念出发,进一步引出数据处理与信息处理的关系、数据管理技术的发展,进而对数据库系统的相关基本概念进行详细的介绍。

1.1.1 信息与数据

数据是数据库系统研究和处理的基本对象,但在谈到数据时,不可避免地会涉及信息的概念。信息和数据是两个不同的概念,但它们互相联系,密不可分。

1. 信息

信息(Information)在不同的应用领域,其含义有所不同。从信息的基本含义和其具有的内涵和外延来说,可将信息定义为:信息是事物属性的标识。这里“事物”泛指存在于人类社会、思维活动和自然界中的一切可能的对象;“属性”可以是事物的基本特征,可以是对事物存在方式的描述,也可以是对事物运动状态的表现形式的刻画。也就是说,事物以其存在的方式和运动状态的表现形式的不同而具有不同的属性标识。“存在方式”反映了事物的内部结构和外部联系。“运动状态”反映了事物在时间和空间上变化所展示的特征、态势和规律。

信息具有如下基本特征。

(1) 可度量性。信息可采用某种度量单位,例如香农给出的信息熵,进行度量,并进一步进行信息编码。计算机中的二进制就是一种最典型的信息编码方式。

(2) 可识别性。根据其信息源的不同,信息可以通过感官直接识别,或通过各种测试手段间接识别。

(3) 可表述性。信息可通过数字、字符、符号、曲线、图表、图形和图像等进行表述。

(4) 可存储性。信息可以以不同的方式存储在不同的介质上。

(5) 可压缩性。人们通过对信息的加工、整理、概括、归纳,可使其精练而浓缩,即用尽可能少的信息量描述一事物。

(6) 可传递性。信息可以通过语言、表情、动作、报刊、书籍、广播、电视、电话、卫星等手段进行传递。信息的可传递性是信息的本质特征。

(7) 可转换性。信息可以由一种形态转换成另一种形态。

(8) 可处理性。人脑的思维功能实现了智能性的信息处理过程。计算机为信息化社会提供了强大的信息处理功能。

(9) 特定范围有效性。信息在特定的范围内是有效的和可利用的,否则就是无效的。

(10) 可共享性。信息具有的扩散性,使得信息具有了可共享性。

2. 数据

所谓数据(Data),就是记录在某种物理载体上的可以被鉴别的符号;是用符号表示的用于反映客观世界中客体属性的记录。数据的基本形式是数字、字符和字符串,广义的数据还可以包括文字、报表、图形、图像、语音等。

数据具有以下基本特征。

1) 数据具有“型”和“值”之分

数据的型是指数据的结构,也即数据的内容构成及其对外的联系。例如,学号、姓名、性别、出生年月、专业代码、班级这6个数据项,可描述学生的基本信息;课程号、课程名、学时这3个数据项,可描述开设课程的基本信息。而下面的结构体:

```
struct stu_c
{
    int snum[9];
    char sname[20];
    char cnum[7];
    char cname[30];
    float grade;
}
```

更直观地给出了由学号、姓名、课程号、课程名、分数共5个数据项组成的学生学习课程数据的“数据具有‘型’”的概念的描述。

按数据的型赋予数据的具体值称为数据的值。例如,表示一个具体学生的数据值为“201401002,李建平,男,1996-8-20,上海,s0401,201401”。

2) 数据具有数据类型和取值范围之约束条件

数据因其描述的对象或属性的不同而具有不同的数据类型。由于数据类型的不同,其表示和存储方式及能进行的运算也不同。同理,数据因其描述的对象或属性的不同,具有不同的取值范围。例如,性别的取值范围为{男,女}。

3) 数据可以通过观察、测量和考核等手段获得

通常情况下,通过观察可以获得定性数据。例如,直接观察不同人员时得到的“老年”、