

巍巍交大 百年书香
www.jiaodapress.com.cn
bookinfo@sjtu.edu.cn



丛书策划 张荣昌
责任编辑 王 清 孟海江
封面设计 唐韵设计

新时代计算机人才培养系列教材

大数据

大数据基础
数据结构与算法
大数据集群搭建维护与数据存储
大数据采集与数据处理
Hadoop应用与开发
数据可视化技术与应用
大数据分析技术与应用
数据挖掘技术与应用

云计算

云计算基础
私有云基础架构与运维
公有云服务架构与运维
云平台配置与管理
云安全技术应用
云网络技术应用
云计算运维开发
云计算应用开发

人工智能

人工智能应用基础
人工智能数学基础
人工智能数据服务
计算机视觉应用开发
深度学习应用开发
机器学习应用开发

自然语言处理及应用开发
智能语音处理及应用开发
人工智能系统部署与运维
人工智能综合项目开发

区块链

区块链应用基础
区块链核心技术
区块链部署与运维
区块链应用设计与开发
区块链项目综合实践
智能合约开发

物联网

物联网基础
物联网工程导论
物联网嵌入式技术
物联网设备装调与维护
物联网系统部署与运维
物联网工程设计与管理
物联网终端智能应用开发基础案例教程
(基于HAL/LL库)

信息安全

信息安全基础
信息安全标准与法规
信息安全工程与管理

信息安全风险评估
密码产品部署与应用
电子商务安全技术
无线网络安全技术
大数据安全技术
电子数据取证技术
终端数据安全及防泄密
数据存储与容灾
Web应用安全与防护

专业基础

数据库应用技术(基于达梦数据库)
数据库技术应用教程(SQL Server 2022版)
计算机网络基础
计算机组装与维护
Linux系统管理与服务器配置
Linux服务器安全高级运维
Windows Server操作系统基础
现代通信原理
无线传输技术
自动识别应用技术
智能传感与检测技术
路由交换技术项目实战教程
虚拟化技术基础

数据库技术应用教程 (SQL Server 2022版)

主编◎高 强 许爱军 洪锐锋



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

新时代计算机人才培养系列教材
“互联网+”新形态一体化教材

数据库技术应用教程 (SQL Server 2022版)

主编◎高 强 许爱军 洪锐锋



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

本书提供教学资源包

网址: <https://www.sjhtbook.com>



扫描二维码
关注上海交通大学出版社
官方微信



ISBN 978-7-313-31846-6

9 787313 318466 >

定价: 49.80元

新时代计算机人才培养系列教材
“互联网+”新形态一体化教材

数据库技术应用教程 (SQL Server 2022版)

主编◎高 强 许爱军 洪锐锋



扫一扫
学习资源库



上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书以 SQL Server 2022 平台为介绍对象,系统而全面地介绍了数据库基础知识、SQL Server 2022 的安装与配置、数据库的创建与管理、数据表创建与管理、数据查询、视图和索引、T-SQL 编程、存储过程和触发器、数据库安全管理、数据库备份与恢复等内容。本书以教务管理系统所使用的数据库抽象模型为应用核心,以理论够用为度,以突出实用性为原则,通过应用案例展开介绍,条目清晰,步骤明确,图文并茂,重点突出,只需按步骤学习并操作,即可轻而易举地掌握相关的知识技能。本书以计算机应用技术专业的教学需求为目标进行编写,适合计算机应用及相关专业的学生学习使用,也适合各行各业的数据库管理人员、信息系统开发人员,以及数据库初学者参考使用。

图书在版编目(CIP)数据

数据库技术应用教程 : SQL Server 2022版 / 高强,
许爱军,洪锐锋主编. -- 上海 : 上海交通大学出版社,
2025. 1 -- ISBN 978-7-313-31846-6

I. TP311.132.3

中国国家版本馆CIP数据核字第202462XV73号

数据库技术应用教程 (SQL Server 2022 版)

SHUJUKU JISHU YINGYONG JIAOCHENG (SQL Server 2022 BAN)

主 编: 高强 许爱军 洪锐锋	地 址: 上海市番禺路 951 号
出版发行: 上海交通大学出版社	电 话: 021-6407 1208
邮政编码: 200030	
印 制: 北京荣玉印刷有限公司	经 销: 全国新华书店
开 本: 889 mm × 1194 mm 1/16	印 张: 15
字 数: 421 千字	
版 次: 2025 年 1 月第 1 版	印 次: 2025 年 1 月第 1 次印刷
书 号: ISBN 978-7-313-31846-6	电子书号: ISBN 978-7-89424-943-2
定 价: 49.80 元	

版权所有 侵权必究

告读者: 如发现本书有印装质量问题请与印刷厂质量科联系

联系电话: 010-6020 6144



党的二十大报告中指出：“科技是第一生产力、人才是第一资源、创新是第一动力”。大国工匠和高技能人才作为人才强国战略的重要组成部分，在现代化国家建设中起着重要的作用。高等教育肩负着培养大国工匠和高技能人才的使命，一直以来致力于培养优质人才。

入门者如何快速达到职业岗位要求，是每位老师和学生都在思考的问题。本书的编写团队基于长期的教学经验与多年的软件开发经验，以行业最新的数据库管理系统 SQL Server 2022 为例，根据程序员和数据库管理员职业岗位要求及院校的教学特点组织内容，以教务管理数据库为主线，介绍使用 SQL Server 2022 进行数据库管理的各种操作，以及数据库应用程序开发所需要的各种知识和技能。本书是将理论、实践、应用开发三者完美结合的一体化教材。

本书在贯彻落实党的二十大精神的基础上，融入科学精神和爱国情怀，通过讲解我国计算机领域的重要事件和人物，弘扬精益求精的专业精神、职业精神和工匠精神，培养学生的创新意识，激发爱国热情，因势利导、潜移默化地将学生个人的成才梦引导到中华民族伟大复兴的中国梦的思想高度。

全书共 10 个项目，包括数据库基础知识、SQL Server 2022 的安装与配置、数据库的创建与管理、数据表创建与管理、数据查询、视图和索引、T-SQL 编程、存储过程和触发器、数据库安全管理、数据库备份与恢复。

本书的主要特点如下。

（1）本书贯彻“项目引领、案例驱动、学做合一”的理念，教学特色非常鲜明，是高等院校“数据库管理”项目化教学的理想教材。

（2）本书基于学生对院校的组织结构、专业、班级、课程开设、成绩管理等信息十分熟悉的情况，以设计教务管理系统使用的数据库抽象模型为应用核心，使学生易于理解和掌握。

（3）本书以理论够用为度，以突出实用性为原则，通过应用案例展开介绍，条目清晰，步骤明确，图文并茂，重点突出，只要按步骤学习并操作，即使是自学者也能轻而易举地掌握相关的知识技能，实现由学习者到职业人的本质提升。

本书由高强、许爱军、洪锐锋主编，潘俊、刘薇、孔令熙为副主编。在本书的编写过程中，得到了易丹、张士刚、曹伶俐、香华冠、高文丽、孟思明、李伟群、何继业、吴小禧、陈少泉、程良明等教师、工程师的大力支持，在此表示衷心的感谢！

本书中的所有例题和脚本均在 SQL Server 2022 中文版环境中运行通过，所有案例的脚本同样适用于 SQL Server 2008 及以上的各个版本。本书配有课件、电子教案、习题答案、脚本及相应素材等教学资源，对此有需要的读者可发邮件至 2393867076@qq.com 领取。

由于编者水平有限，书中若存在疏漏和不足之处，恳请广大师生和读者给予批评与指正。

编者

2024 年 8 月



目录

项目 1 数据库基础知识

任务 1.1 认识数据库	2	1.3.3 关系操作	10
1.1.1 数据库基本概念	2	任务 1.4 了解关系规范化	11
1.1.2 数据库的发展	3	1.4.1 关系规范化的必要性	11
任务 1.2 认识数据模型	3	1.4.2 函数依赖关系	12
1.2.1 数据处理的抽象描述	4	1.4.3 关系规范化处理	13
1.2.2 实体-联系模型	5	拓展阅读 中国数据库的先行者与探索者	
1.2.3 结构数据模型	7	——萨师煊	14
任务 1.3 了解关系数据库	9	项目归纳	15
1.3.1 关系数据库的基本概念及特点	9	拓展训练	15
1.3.2 数据完整性	10		

项目 2 SQL Server 2022 的安装与配置

任务 2.1 认识 SQL Server 2022	17	2.3.2 配置远程登录	27
2.1.1 SQL Server 2022 的新特性	17	2.3.3 配置“sa”管理员账户	28
2.1.2 SQL Server 2022 的管理工具	17	2.3.4 使用混合模式登录	29
任务 2.2 安装 SQL Server 2022	18	拓展阅读 华为——高斯数据库	29
2.2.1 SQL Server 2022 的安装要求	18	项目归纳	30
2.2.2 SQL Server 2022 的安装步骤	19	拓展训练	30
任务 2.3 配置 SQL Server 2022	24		
2.3.1 SQL Server 网络配置	24		

项目 3 数据库的创建与管理

任务 3.1 掌握 SQL Server 2022 数据库	32	3.1.2 数据库文件和文件组	32
3.1.1 数据库的组成对象	32	3.1.3 系统数据库和用户数据库	33

任务 3.2 创建数据库	34	3.4.2 重命名数据库	45
3.2.1 使用 SSMS 创建数据库	34	3.4.3 收缩数据库	46
3.2.2 使用 T-SQL 语句创建数据库	37	3.4.4 删除数据库	47
任务 3.3 使用 SSMS 管理数据库	40	任务 3.5 使用 SSMS 分离和附加数据库	48
3.3.1 修改数据库	40	3.5.1 分离数据库	48
3.3.2 重命名数据库	41	3.5.2 附加数据库	48
3.3.3 收缩数据库	42	拓展阅读 云计算和云数据库	49
3.3.4 删除数据库	43	项目归纳	50
任务 3.4 使用 T-SQL 语句管理数据库	43	拓展训练	50
3.4.1 修改数据库	43		

项目 4 数据表创建与管理

任务 4.1 创建数据表	53	4.3.2 重命名表名和删除数据库表	67
4.1.1 数据表的概念	53	任务 4.4 使用 SSMS 方式维护数据	68
4.1.2 数据类型	53	4.4.1 插入记录	68
4.1.3 教务管理数据库表定义	54	4.4.2 修改数据	69
4.1.4 数据完整性约束概述	56	4.4.3 删除记录	70
4.1.5 创建数据表	57	任务 4.5 使用 T-SQL 语句维护数据	70
4.1.6 字段的约束	61	4.5.1 INSERT 语句插入记录	70
任务 4.2 使用 SSMS 方式管理数据库表	64	4.5.2 UPDATE 语句修改数据	72
4.2.1 修改数据库表	64	4.5.3 DELETE 语句删除记录	73
4.2.2 重命名表名	65	拓展阅读 中国计算机的主奠基者——华罗庚	74
4.2.3 删除数据库表	65	项目归纳	75
任务 4.3 使用 T-SQL 语句管理数据库表	65	拓展训练	75
4.3.1 修改数据库表	66		

项目 5 数据查询

任务 5.1 掌握 SELECT 查询语句	79	5.2.4 限制返回行数	82
5.1.1 SELECT 查询语句	79	5.2.5 去除查询结果集的重复行	83
5.1.2 SELECT 语句的执行过程	79	5.2.6 自定义输出列	83
任务 5.2 掌握 SELECT 选择输出列子句	80	5.2.7 无数据源查询	84
5.2.1 查询所有列数据	80	5.2.8 聚合函数查询	85
5.2.2 查询指定列数据	81	任务 5.3 掌握 FROM 提供数据源子句	86
5.2.3 对输出列重命名	81	5.3.1 对数据源对象命名别名	86

5.3.2 基表数据源对象查询	86	5.6.3 随机排序查询	99
5.3.3 派生表数据源对象查询	87	5.6.4 自定义输出列排序查询	99
5.3.4 视图数据源对象查询	87	任务 5.7 使用 GROUP BY 分组查询子句	100
5.3.5 FROM 子句的连接查询	88	5.7.1 单列分组查询	100
5.3.6 自连接查询	91	5.7.2 多列分组查询	100
任务 5.4 使用 INTO 创建并插入子句	91	5.7.3 HAVING 分组选择查询	101
任务 5.5 使用 WHERE 指定选择条件子句	92	任务 5.8 掌握子查询	101
5.5.1 关系比较查询	92	5.8.1 在查询语句中使用子查询	102
5.5.2 逻辑运算查询	92	5.8.2 在插入语句中使用子查询	104
5.5.3 WHERE 子句的连接查询	93	5.8.3 在修改语句中使用子查询	105
5.5.4 范围比较查询	94	5.8.4 在删除语句中使用子查询	105
5.5.5 IN 存在查询	94	任务 5.9 使用查询设计器设计查询	105
5.5.6 模糊匹配查询	95	5.9.1 关于查询设计器	105
5.5.7 空值比较查询	98	5.9.2 使用查询设计器	106
任务 5.6 使用 ORDER BY 查询结果集排序子句	98	拓展阅读 中国国家顶级域名“CN”	109
5.6.1 单关键字排序查询	98	项目归纳	109
5.6.2 多关键字排序查询	99	拓展训练	110

项目 6 视图和索引

任务 6.1 认识视图	112	6.4.2 创建索引	118
6.1.1 视图的概念、特点	112	任务 6.5 使用 SSMS 创建与维护索引	119
6.1.2 视图的作用	112	6.5.1 创建索引	119
任务 6.2 使用 SSMS 创建与维护视图	113	6.5.2 修改索引	120
6.2.1 创建视图	113	6.5.3 禁用索引	121
6.2.2 修改视图	113	6.5.4 重新生成索引	121
6.2.3 重命名视图	114	6.5.5 重新组织索引	121
6.2.4 删除视图	115	6.5.6 删除索引	121
任务 6.3 使用 T-SQL 创建与维护视图	115	任务 6.6 使用 T-SQL 创建与维护索引	122
6.3.1 创建视图	115	6.6.1 创建索引	122
6.3.2 修改视图	116	6.6.2 修改索引	122
6.3.3 重命名与删除视图	117	6.6.3 删除索引	123
6.3.4 使用视图查询	117	拓展阅读 图灵奖与姚期智	123
任务 6.4 认识索引	118	项目归纳	124
6.4.1 索引的类型	118	拓展训练	124

项目 7 T-SQL 编程

任务 7.1 掌握 T-SQL 编程基础	126	7.4.3 使用用户自定义函数	151
7.1.1 标识符与注释	126	7.4.4 删除用户自定义函数	151
7.1.2 常量	127	任务 7.5 声明和使用游标	152
7.1.3 局部变量	127	7.5.1 声明游标	152
7.1.4 全局变量	129	7.5.2 打开游标	153
7.1.5 运算符	130	7.5.3 关闭游标	154
任务 7.2 掌握流程控制语句和批处理	133	7.5.4 释放游标	155
7.2.1 流程控制语句	133	7.5.5 提取游标数据	155
7.2.2 批处理	137	任务 7.6 应用游标	157
任务 7.3 使用系统内置函数	138	7.6.1 使用游标查询记录	157
7.3.1 聚合函数	138	7.6.2 使用游标插入记录	157
7.3.2 数学函数	139	7.6.3 使用游标删除记录	158
7.3.3 字符串函数	140	7.6.4 使用游标修改记录	159
7.3.4 日期时间函数	142	7.6.5 在函数中使用游标	160
7.3.5 排名函数	143	7.6.6 在存储过程中使用游标	160
7.3.6 数据类型转换函数	146	拓展阅读 中国计算机软件系统的开创者	
任务 7.4 编写用户自定义函数	147	——陈火旺院士	161
7.4.1 创建用户自定义函数	148	项目归纳	162
7.4.2 修改用户自定义函数	150	拓展训练	163

项目 8 存储过程和触发器

任务 8.1 创建和使用存储过程	165	8.2.2 创建触发器	175
8.1.1 存储过程概述	165	8.2.3 修改触发器	181
8.1.2 创建存储过程	166	8.2.4 删除触发器	182
8.1.3 调用存储过程	170	拓展阅读 国家最高科学技术奖	183
8.1.4 修改存储过程	171	项目归纳	183
8.1.5 删除存储过程	173	拓展训练	184
8.1.6 使用系统存储过程	173		
任务 8.2 创建和使用触发器	174		
8.2.1 触发器概述	174		

项目 9 数据库安全管理

任务 9.1 认识数据库安全管理	186	9.3.3 数据库用户	197
任务 9.2 了解 SQL Server 服务器的安全管理	186	9.3.4 特殊的数据库用户	200
9.2.1 身份验证模式	186	任务 9.4 掌握数据对象的安全管理	200
9.2.2 登录账号的创建和删除	187	9.4.1 权限的种类	200
9.2.3 服务器角色	192	9.4.2 权限的管理	201
任务 9.3 掌握数据库的安全管理	194	拓展阅读 数据库管理员的“匠人精神”	207
9.3.1 架构	194	项目归纳	207
9.3.2 数据库角色	196	拓展训练	208

项目 10 数据库备份与恢复

任务 10.1 了解备份与恢复	210	10.4.2 使用差异备份恢复数据库	217
10.1.1 备份的类型	210	任务 10.5 掌握数据库事务日志备份与恢复	218
10.1.2 备份策略	210	10.5.1 数据库的事务日志备份	218
任务 10.2 创建与删除备份设备	211	10.5.2 使用日志备份恢复数据库	219
10.2.1 创建备份设备	211	任务 10.6 导入与导出数据	220
10.2.2 删除备份设备	212	10.6.1 从 Excel 导入数据	220
任务 10.3 掌握数据库的完整备份与恢复	213	10.6.2 将数据导出到 Excel	223
10.3.1 数据库的完整备份	213	拓展阅读 数据安全是维护数据库之根本	226
10.3.2 使用完整备份恢复数据库	214	项目归纳	226
任务 10.4 掌握数据库的差异备份与恢复	216	拓展训练	226
10.4.1 数据库的差异备份	216		

参考文献	228
------------	-----

项目 1

数据库基础知识

❖ 知识目标

- (1) 了解数据库的基本概念和发展情况。
- (2) 了解数据处理的抽象描述。
- (3) 理解实体间的各种关联关系。
- (4) 掌握 E-R 图的元素、设计。
- (5) 理解数据的完整性。
- (6) 掌握关系的规范化处理。

❖ 能力目标

- (1) 能够设计、绘制 E-R 图。
- (2) 能够对关系模型进行实体完整性、域完整性和参照完整性设计。
- (3) 能够对关系模型进行规范化设计。

❖ 素质目标

- (1) 明确数据库技术的重要地位，激发科技报国的爱国情怀和使命担当。
- (2) 坚持基础知识与技能并重，培养严谨的工作态度和工作作风。

任务 1.1 认识数据库

数据库技术是计算机技术的一个重要分支，计算机应用技术发展到现在，除操作系统之外，数据库已经成为计算机应用的重要支撑系统和核心组成部分。例如，电子政务系统、电子商务系统、决策支持系统、电子邮件系统、财务金融系统等计算机应用系统均以数据库为信息存储平台，离不开数据库的支持。

1.1.1 数据库基本概念

信息、数据、数据库、数据库系统、数据库管理系统等是与数据库技术密切相关的基本概念，理解这些基本概念的含义有助于进一步深入学习和掌握数据库管理系统的应用。

1. 信息

信息（information）是经过加工处理并对人类客观行为产生影响的数据表现形式。信息具有实效性、实用性和知识性等特性。

信息与数据是有差别的，任何事物的属性都是通过数据来表示的。数据经过加工处理后，便具有了知识性，并对人类活动产生决策作用，从而形成信息。信息与数据也是有关联的，信息是数据的内涵，数据是信息的载体。同一条信息可以用不同形式的数据表示，而同一个数据也可以有不同的解释。

2. 数据

数据（data）是对客观事物属性的一种符号化的表示。从数据处理的角度看，数据是计算机处理及数据库中存储的基本对象。数据的表现形式很多，它们都可以经过数字化后存入计算机。例如，数字、字母、文字、图像、声音等在计算机中都可以以数据的形式体现。

3. 数据库

数据库（DB）是数据存储的仓库，是指长期存储在计算机内、有组织、可共享的数据的集合。数据库中的数据按一定的数据模型进行组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户所共享。

4. 数据库系统

数据库系统（DBS）是采用数据库技术构建的复杂的计算机系统，它是综合了计算机硬件、软件、数据集合和数据库管理人员、数据库规则，向用户和应用程序提供信息服务的集成系统。数据库系统由数据库、软件系统、硬件系统、数据库管理系统、数据库管理员和用户等要素组成。

5. 数据库管理系统

数据库管理系统（DBMS）是指数据库系统对数据进行管理的软件系统，是数据库系统的核心组成部分，它为用户提供了一个可以方便、有效地存取数据库信息的环境。数据库管理系统的功能由数据定义、数据库的建立和维护、数据库的运行管理、数据操纵等四个方面组成。

6. 数据处理

数据处理（data processing）是将数据转换成信息的过程，是由人、计算机等组成的能对信息进行搜索、传递、存储、加工、维护、分析、计划、控制、决策和使用的系统。经过处理，信息被加工成特定

形式的数据。

在数据处理过程中，数据计算相对简单，但是处理的数据量大，并且数据之间存在着复杂的联系，因此数据处理的关键是数据管理。

数据管理是指对数据进行搜集、整理、组织、存储和检索等操作，这部分操作是数据处理业务必不可少的基本环节。因此读者必须学习和掌握数据管理的技术，以便为数据处理提供有力的支持。有效的数据管理可以提高数据的使用率，减轻程序开发人员的负担。数据库技术就是针对数据管理的计算机软件技术。

1.1.2 数据库的发展

数据库技术是由于数据管理任务的需要而产生的，数据管理技术发展至今大致可分为人工管理、文件管理、数据库管理等三个阶段。

1. 人工管理阶段

人工管理阶段出现在计算机应用于数据管理的初期。计算机没有问世之前，主要是使用纸张等媒介对数据进行管理；计算机问世初期，由于计算机的软、硬件技术不像今天那么先进，用户的应用程序与一组数据直接对应，程序不仅要设计数据处理的方法，还要直接操作数据。在这种管理方式下，用户的应用程序与数据之间相互结合、不可分割，当数据有所变动时程序也必须随之改变，因此，程序与数据的独立性差，程序之间的数据不能互相传递，缺少共享性。

2. 文件管理阶段

文件管理阶段是把有关的数据组织成一种数据文件，数据可以脱离程序独立存在。在这种管理方式下，用户或应用程序通过文件管理系统对数据文件中的数据进行加工处理，应用程序的数据具有一定的独立性，比人工管理方式前进了一步。但是，数据文件仍然高度依赖其应用程序，不能被多个程序所共享。

3. 数据库管理阶段

随着数据量的不断增大，文件管理方式显然在操作效率和数据共享方面不能满足实际要求，为了解决在海量数据、数据共享和操作效率等方面存在的问题，数据库出现了。

数据库管理阶段对所有的数据实行统一的规划管理，所有数据形成一个数据中心，构成一个数据仓库。在这种管理方式下，应用程序不再只与孤立的数据文件相对应，而是能够从整体数据集中取出某个子集作为逻辑文件与应用程序相对应，并通过数据库管理系统实现逻辑文件与物理数据之间的映射。在数据库管理系统环境下，应用程序对数据的管理和访问灵活方便，数据与应用程序之间相互独立，使得程序的编制质量和效率大大提高，数据的冗余度极大减少，数据的共享性显著增强。

任务 1.2 认识数据模型

数据库管理系统是基于某个数据模型设计出来的，而数据模型是对现实世界数据特征的模拟和抽象。完整的数据模型主要包括数据结构、数据操作和数据完整性约束三个部分。其中，数据结构用来描述实体之间的构成和联系；数据操作是指对数据库的查询和更新等操作；数据完整性约束则是指施加在数据

上的限制和规则，目的是使数据库中的数据更具可用性和有效性。

数据模型的设计要满足三方面的要求：一是能够真实地模拟现实世界；二是容易理解；三是能够在计算机上实现。

根据应用目的的不同，数据模型分为两种：一种是信息模型，也称概念模型，它从用户的观点来对数据和信息进行建模，反映了信息从现实世界到信息世界的转换，它不涉及计算机软硬件的具体细节，只注重符号表达和用户的理解能力，典型的信息模型有 E-R 模型（实体－联系模型），信息模型主要用于数据库设计阶段；另一种是结构数据模型，结构数据模型主要用于 DBMS 的实现，它反映了信息从信息世界到机器世界的转换，描述了计算机中数据的逻辑结构、信息在存储器上的具体组织等。常见的结构数据模型有三种，即层次模型、网状模型和关系模型。

1.2.1 数据处理的抽象描述

人们在研究和处理数据的时候，通常把对数据的描述分为三个世界，即现实世界、信息世界、机器世界，这三个世界对信息描述的转换过程，就是将客观现实的信息反映到计算机数据库中的过程。

1. 现实世界

客观存在的世界就是现实世界，它独立于人们的思想之外。现实世界存在无数的对象和事物，每一个对象或事物可以看成一个个个体，每个个体有一项或多项特征信息。例如，把人看成对象时，有身高、体重、年龄、肤色等基本特征。

2. 信息世界

信息世界是现实世界在人们头脑中的反映，人的思维将现实世界中对象或事物的特征抽象化后用文字符号表示出来，就形成了信息世界。描述信息世界的常用术语如下。

（1）实体。客观存在的并且可以相互区别的事物称为实体，实体可以是具体的事物，也可以是抽象的事件。例如，一名学生，一台计算机，上课、比赛等。

实体是信息世界的基本单位，相同类型的实体的集合称为实体集。例如，一个班级由多位学生实体组成，则这个班级称为学生实体的实体集。

（2）属性。实体的特性称为属性，一个实体可以有多个属性，每一个属性都有其数据类型和数据的取值范围。例如，学生实体可由学号、姓名、成绩等若干属性来描述。

（3）键。在一个实体集中能唯一标识一个实体的属性称为键。键可以是一个属性，也可以是多个属性的组合。例如，学号、身份证号等能唯一地标识一个学生，它们都可作为学生实体的键。

（4）联系。实体之间相互作用、互相制约的关系称为实体间的联系，也称为关联或关联关系。实体之间的关联关系主要有一对一、一对多（多对一）和多对多三种。

3. 机器世界

机器世界又称为数据世界。信息世界中的信息经过抽象和组织，以数据的形式存储在计算机中，从而形成了机器世界。在机器世界中，用于描述数据的基本术语与信息世界中的术语是一一对应的，只是在文字叙述上略有差异。

（1）字段。字段也称为数据项，用于标记实体的属性。在一张关系表中，每一列称为一个字段。字段与信息世界的“属性”相对应。例如，在学生信息表中，一个学生就是一个实体，它包含了“学号”“姓名”“性别”“年龄”等字段。

(2) 记录。记录是具有逻辑关系的一个或多个字段的集合。记录与信息世界中的“实体”相对应，一个记录描述了一个实体的基本信息。

(3) 关键字。能够唯一标识一条记录的字段称为关键字。关键字与信息世界中的“键”相对应。关键字可以是一个字段，也可以由多个字段组合而成。例如，学生信息表中的学号、身份证号等能唯一地标识一名学生，它们都是关键字。

(4) 文件。文件是记录的集合。文件对应于信息世界中的“实体集”。文件的存储形式有很多种，如顺序文件、随机文件、索引文件等。

4. 三个世界信息描述的对应关系

从现实世界到信息世界再到机器世界，事物被逐层抽象，信息被逻辑化、符号化。表 1-1 展示了三个世界信息描述的对应关系。

表 1-1 三个世界信息描述的对应关系

现实世界	信息世界	机器世界
特征	属性	字段
唯一特征	键	关键字
事物	实体	记录
	实体集	文件

1.2.2 实体 – 联系模型

实体 – 联系模型（简称 E-R 模型）是目前最常用的信息模型，它不涉及计算机软、硬件的具体细节，便于分析与理解。

1. E-R 图的基本元素

E-R 模型使用 E-R 图来表达，E-R 图主要由实体、属性、联系和连线等元素符号组成。E-R 图基本元素的表达符号和用途如表 1-2 所示。

表 1-2 E-R 图基本元素的表达符号和用途

符号	用途
矩形框	表示实体，框内写实体的名称
椭圆框	表示实体的属性，框内写属性的名称，该属性是实体的非主键
粗边椭圆框	表示实体的属性，框内写属性的名称，该属性是实体的主键
线条	用于实体、属性、联系之间的连接
菱形框	表示实体间的联系，框内写联系的名称，旁边标上关联关系的类型

2. E-R 图的设计

完整的 E-R 图必须清楚地表达实体、属性和联系三者之间的关系。下面根据常见的关联关系的种类分别介绍 E-R 图的设计方法。

1) 一对一关联关系

实体集 A 中的每一个实体, 在实体集 B 中最多只能有一个实体与之有关系, 反之亦然, 则称实体集 A 与实体集 B 的关系是一对一的关联关系, 记为 $1:1$ 。

【例 1-1】使用 E-R 图描述学校和校长两个实体的关联关系。

分析: 因为一个学校只有一个校长, 而一个校长只在一个学校任校长, 所以学校和校长的关系是一对一的关联关系。设学校实体的属性主要有编号、校名、地址等; 校长实体的属性主要有编号、姓名、职称等。使用 E-R 图描述如图 1-1 所示。

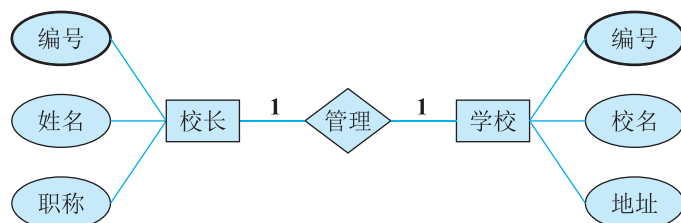


图 1-1 一对一关联关系

2) 一对多或多对一关联关系

如果实体集 A 中的每一个实体, 在实体集 B 中有 n 个实体 ($n \geq 0$) 与之有关系, 则称实体集 A 与实体集 B 是一对多的关联关系, 记为 $1:n$ 。如果实体集 A 中有 n 个实体 ($n \geq 0$) 与实体集 B 中的每一个实体有关系, 则称实体集 A 与实体集 B 是多对一的关联关系, 记为 $n:1$ 。

一对多与多对一的关联关系是同一种关系, 只是观察问题的方向不同而已。

【例 1-2】使用 E-R 图描述班级与学生两个实体的关联关系。

分析: 因为一个班有多名学生, 而每个学生只能是某个班级的成员, 所以班级与学生的关系是一对多的关联关系。设班级实体的主要属性有编号、班名、教室等; 学生实体的主要属性有编号、姓名、年龄等。使用 E-R 图描述如图 1-2 所示。

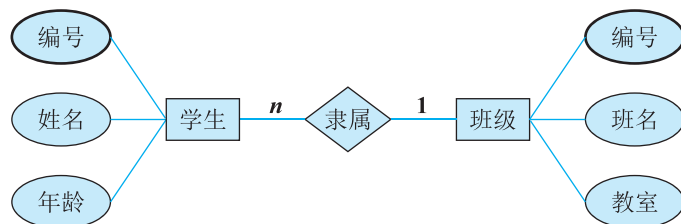


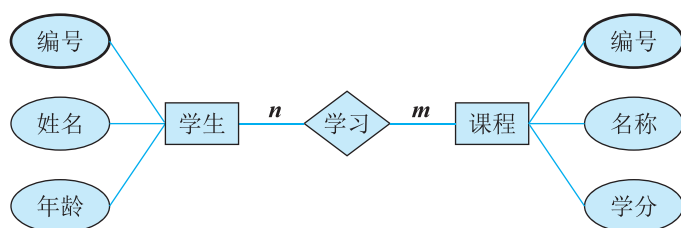
图 1-2 一对多或多对一关联关系

3) 多对多关联关系

如果实体集 A 中的每一个实体, 在实体集 B 中有 n 个实体 ($n \geq 0$) 与之有关系。反之, 对于实体集 B 中的每一个实体, 在实体集 A 中也有 m 个实体 ($m \geq 0$) 与之有关系, 则称实体集 A 与实体集 B 是多对多的关联关系, 记为 $n:m$ 。

【例 1-3】使用 E-R 图描述课程与学生两个实体的关联关系。

分析: 因为一门课程同时有若干名学生选修, 而一个学生可以同时选修多门课程, 所以课程与学生之间的关系是多对多的关联关系。设课程实体的主要属性有编号、名称、学分等; 学生实体的主要属性有编号、姓名、年龄等。使用 E-R 图描述如图 1-3 所示。

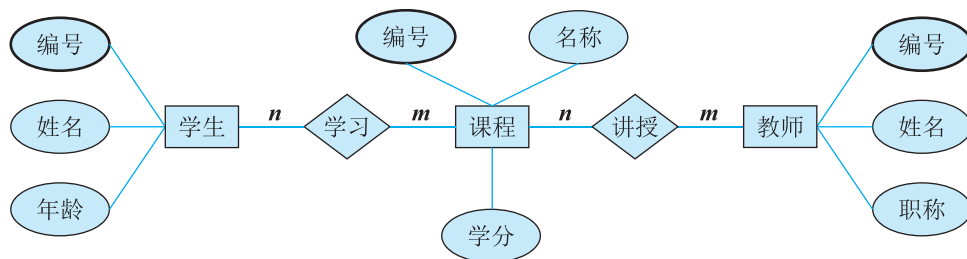


4) 多个实体之间的关联关系

如果实体集 A 与实体集 B 发生某种关系（如一对多等），而实体集 B 与实体集 C 也发生某种关系（如一对多、一对一、多对多等），则多个实体之间可能两两产生关联关系，这是实际应用中常见的综合性关联关系。

【例 1-4】使用 E-R 图描述学生、课程和教师三个实体之间的关联关系。

分析：学生实体和课程实体的关联关系由【例 1-3】描述。教师实体与课程实体的关系是一个教师可以授 n 门（ $n \geq 0$ ）课程，一门课程也可以由多个教师讲授，所以，教师实体与课程实体也是多对多的关联关系。假定教师实体的主要属性有编号、姓名、职称等。描述学生、课程和教师三个实体之间的关联关系的 E-R 图，如图 1-4 所示。



在一个应用系统中，把所有实体间产生的关系利用 E-R 图描述出来并组合成一个总的 E-R 图后，就得到了该应用系统的完整的 E-R 模型图。

设计 E-R 模型时，由于设计人员的经验或见解不同，针对同一个企业运营模式，不同人员设计的 E-R 模型是有差异的，因此，E-R 图设计不是唯一的。在实际应用中，当 E-R 图比较庞大时，为了使 E-R 图表达更抽象，可以将属性省略，只保留实体和联系两部分内容。

总之，使用 E-R 图表达的 E-R 模型直观易懂，它是系统开发人员和客户沟通的良好工具。对客户来说，它概括了企业运营的方式和各种联系；对于系统开发人员来说，它从概念上描述了一个应用系统数据库的信息组织。根据 E-R 图，结合具体的 DBMS 可以设计出相应的数据库管理系统的结构数据模型。设计完善、高质量的 E-R 图是数据库设计的一个重要步骤。

1.2.3 结构数据模型

1. 层次模型

以树形结构表示实体之间联系的模型称为层次模型。在这种模型中，数据被组织成由“根”开始的“树”，每个实体由根节点开始沿着不同的分支放在不同的层次上。如果不再向下分支，则分支序列中最

后的节点称为“叶”节点。层次模型实例如图 1-5 所示。

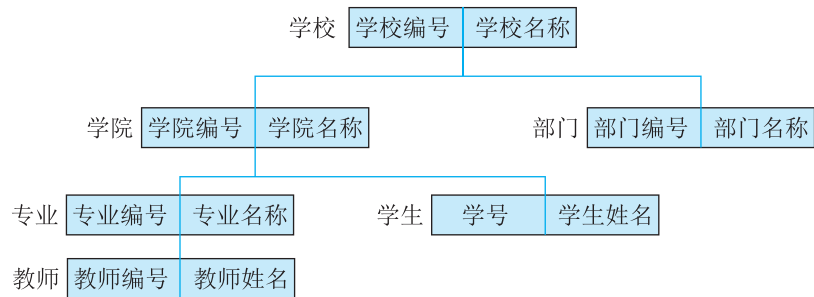


图 1-5 层次模型实例

层次模型的主要特征：有且仅有一个节点无父节点，这个节点称为根节点；除根节点外，每一个节点有且只有一个父节点；父节点与子节点之间是一对多的关联关系；层次模型无法表达多对多的关联关系。

2. 网状模型

用网状结构表示实体之间联系的模型称为网状模型。网状模型是层次模型的扩展，网中的每一个节点代表一个实体类型。网状模型实例如图 1-6 所示。

网状模型的主要特征：允许一个以上的节点没有父节点；一个节点可以有多个父节点；节点间是多对多的关系。

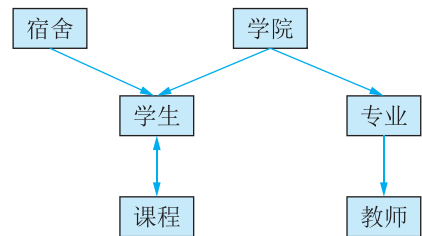


图 1-6 网状模型实例

3. 关系模型

采用二维表的结构表示实体之间的联系的模型称为关系模型。关系模型以关系数学理论为基础，操作的对象和结果都是二维表，这种二维表与日常生活、工作中所使用的行列表格的概念完全一样，所以，关系模型是目前应用最为广泛的数据模型。

在关系模型中，无论实体本身还是实体间的联系均用称为“关系”的二维表来表示。表 1-3 列出了学生入学时的概要信息，这是一张典型的二维关系表格，表中各行的数据从层次隶属的角度看是没有关系的，即表中的行数据是平行的，不代表从属关系。

表 1-3 学生入学时的概要信息

学号	姓名	性别	出生日期	入学总分	入学时间	班级编号
0101220101	李表华	女	2003-01-01	504	2022-09-01	010101
0101220102	陈小东	男	2002-09-24	486	2022-09-01	010101
0101230103	李明明	女	2003-11-26	483	2023-09-01	010103

关系模型概念清晰，结构简单，用户比较容易理解，具有较好的数据独立性和安全保密性。但是，

关系模型也有缺点，其查询效率有时不如层次模型和网状模型。

任务 1.3 了解关系数据库

关系模型是当今最流行的数据模型，较好地解决了层次模型和网状模型中存在的问题，目前大部分流行的数据库管理系统产品如 Oracle、SQL Server、MySQL 等都是基于关系模型实现的。

1.3.1 关系数据库的基本概念及特点

关系数据库是基于关系模型建立的数据库。关系数据库由一个或多个称为“关系”的二维表组成，每个关系又由一条或多条记录组成，而每条记录则由一个或多个称为“字段”的属性组成。

1. 基本概念

(1) 关系。关系是指一张有规则的、没有重复行或重复列的二维表。一个关系对应于一张二维表，每张二维表（下面常用“关系”简称）有一个表名。

(2) 记录。在关系中，水平方向的行称为记录，关系中的每一行称为一条记录。记录也称为实体或元组（下面常用“记录”简称）。

(3) 属性。在关系中，垂直方向的列称为属性。属性在关系中也叫列、栏、字段等（下面常用“字段”简称），每一列有一个属性名，也叫列名、栏名、字段名等（下面常用“字段名”简称）。关系中的字段是有值的，它等于当前被处理的记录所对应的字段值。

(4) 域。字段的取值范围称为域。例如，“性别”字段只能取“男”或“女”两者之一。

(5) 关联。关联是指关系与关系之间记录数据相互关联的方式。关联也称联系或关联关系（下面常用“关联关系”简称）。关联关系是数据库规范化处理时需要重点解决与优化处理的核心问题之一，主要有一对一、一对多（多对一）、多对多三种。

(6) 关键字。在关系中，能够唯一标识一条记录的字段或多个字段的组合称为关键字。关键字也称为“键”。例如，在表 1-3 中，“学号”能唯一标识一名学生，可以充当关键字；但学生的“姓名”不能拿来充当关键字，因为同一个班级可能出现两个或多个同名同姓的学生。

(7) 候选关键字。当关系中存在两个或多个“键”时，称这些“键”为候选关键字或候选键。

(8) 主关键字。在关系中，要从一个或多个候选关键字中选出一个关键字作为主关键字，主关键字也称主键（下面常用“主键”简称）。在一个关系中只能有一个主键。

(9) 外部关键字。在关系中，如果存在这样一个字段：一方面，它不是关系中的关键字，另一方面，它的取值来自另一个关系的主键，则此字段称为外部关键字，也叫外键。

2. 特点

(1) 关系中的字段不可再分，必须是最基本的原子数据项。

(2) 关系中每一字段的值的数据类型相同，并且来自同一个域。

(3) 字段在关系中的左、右次序不影响关系。

(4) 记录在关系中的上、下次序不影响关系。

(5) 同一关系中不允许出现两个相同的记录。

(6) 同一关系中不允许出现两个相同的字段。

1.3.2 数据完整性

数据完整性是为保证数据库中关系数据的正确性和相容性，对关系模型提出的约束条件和需要遵守的规则。数据完整性通常包括实体完整性、域完整性和参照完整性。

1. 实体完整性

实体完整性是在关系内部建立的记录与记录之间的约束规则。它通过关系的主键来实现约束，要求关系中主键的值不能为空，并保证数据具有唯一性，这样就能确保关系中不会出现重复的记录。实体完整性解决关系内部记录与记录之间的数据约束问题。

2. 域完整性

域完整性是指在关系内部建立的对字段的定义和取值范围进行限制的约束规则。域完整性解决的是字段自身的数据约束问题。

3. 参照完整性

参照完整性是指在关系与关系之间（或在同一个关系的不同记录之间）建立的约束规则。它通过一个关系的外键与另一个关系的主键建立起关联关系来进行关系之间的约束。参照完整性主要用于解决关系与关系之间的数据约束问题。

1.3.3 关系操作

关系数据库的操作主要有选择、投影和连接三种操作。数学中的集合运算是关系操作的理论基础。

1. 选择操作

选择操作也叫筛选操作，是指从指定的关系中选择出满足某种条件的记录从而形成一个新的关系的操作。

2. 投影操作

投影操作是对指定的关系进行字段的投影，它从指定的关系中抽取出某些满足需要的字段列以形成一个新的关系，因此，投影操作可以看成对某个关系进行列的选择操作。

3. 连接操作

连接操作是从两个或多个关系的记录中选择出满足某一给定条件的记录作为操作结果。在连接操作中，一般都包含选择操作和投影操作，如果包含选择操作则会改变结果关系的记录条数，如果包含投影操作则会改变结果关系的字段个数。

两个关系进行连接操作后，如果连接操作没有包含投影操作，则结果关系的字段个数是两个关系的字段的并集；如果连接操作没有包含选择操作，则结果关系的记录条数是两个关系记录的笛卡尔乘积，例如，设两个关系的记录数分别为 N 和 M ，两个关系进行不包含选择操作的连接操作后，结果关系的记录数为 N 和 M 的笛卡尔乘积，可表示为 $N \times M$ 。

任务 1.4 了解关系规范化

数据库常常包含多个关系，为了使数据库中关系的结构更加合理，消除各种操作异常，使数据冗余尽可能地小，便于查询、插入、更新和删除操作，数据库中的关系在设计时一般都要经过规范化处理。关系规范化的目的是使得关系模式中的每个属性都是不可再分的数据项，使得关系模式尽可能地满足第三范式甚至第五范式的基本要求。

1.4.1 关系规范化的必要性

下面通过分析表 1-4 的数据结构来理解关系规范化的必要性。表中的数据以学号排序，列出了每位学生的专业、各门课程的授课教师和课程成绩等信息。从学生的角度看，课程的成绩信息比较详尽，但从数据库管理的角度去分析，这个关系却存在诸多问题。

表 1-4 综合信息表

学号	姓名	性别	专业	教师编号	教师姓名	职称	课程编号	课程名称	成绩
0101	张晓青	女	软件（Java）	0001	李光策	讲师	J001	软件工程	90
0101	张晓青	女	软件（Java）	0002	孙静祥	教授	J002	数据结构	86
0102	王天明	男	软件（Java）	0001	李光策	讲师	J001	软件工程	90
0102	王天明	男	软件（Java）	0002	孙静祥	教授	J002	数据结构	76

1. 存在非原子属性

首先，关系表中的“专业”属性不是原子元素，还可以继续分解成“专业名称”和“专业方向”两个属性。

2. 存在大量的数据冗余

例如，关系中每个教师有“教师编号，教师姓名，职称”三个数据项，如果一个班有 50 个学生，则每门课程的教师信息将有 50 份相同的数据，存在大量的数据冗余。

3. 容易造成插入、修改和删除异常

从关系中插入、修改和删除数据时，只要数据满足域完整性规则，操作就不会产生异常，但容易造成数据不一致。例如，修改某位教师的“职称”数据时，可能会漏改或错改了其他行该教师的职称数据。如果深入分析，该关系还容易产生插入、修改和删除等操作异常。

提示：限于篇幅，关于操作异常这里不做深入探讨，有兴趣的读者可参考数据库设计原理方面的资料或书籍。

通过上述分析，可以确认表 1-4 是一个不好的关系。一个好的关系，不应存在非原子属性，不应发生插入异常或删除异常，应该尽可能地避免出现修改异常、减少冗余数据。如何满足这些要求正是关系规范化要解决的问题。

关系的设计主要包括函数依赖、关系范式和模式设计等三方面的内容，其中，函数依赖起着核心作

用,下面做简要介绍。

1.4.2 函数依赖关系

关系的规范化主要是处理关系中的属性对其他某个或某些属性数据的依赖程度,这些依赖条件主要是函数依赖条件。在概念上,函数依赖是指一个或一组属性的取值可以决定其他属性的取值。

常见的函数依赖有完全函数依赖、部分函数依赖、传递函数依赖三种。

在介绍函数的依赖关系之前,先介绍关系的另一种表示方法,即“关系模式”法。

格式:

关系名 (属性名[, ...n])

例如:

专业 (专业编号, 专业名称)

1. 函数依赖定义

设 R 是关系模式中的一个关系, X 、 Y 是 R 的属性子集。如果 X 的值确定后, 属性集 Y 的值也同时被确定, 则称 X 函数决定 Y , 也就是说 Y 函数依赖于 X , 记为 $X \rightarrow Y$; 如果可逆则为互相依赖, 记为 $X \leftrightarrow Y$; 如果不存在函数依赖, 则记为 $X \nrightarrow Y$ 。

例如, 设有关系“学生(学号, 姓名, 性别)”, 则属性“姓名, 性别”都函数依赖于“学号”, 即有“学号 $\rightarrow$ 姓名, 学号 $\rightarrow$ 性别”, 换句话说, “学号”可以决定“姓名”“性别”等属性的取值。

2. 完全函数依赖

如果 Y 属性集依赖于 X , 而不依赖于 X 的任何一个真子集, 则称 Y 完全函数依赖于 X , 记为 $X \twoheadrightarrow Y$ 。

例如, 设有关系“成绩(学号, 课程号, 成绩)”, 因为“学号”或“课程号”各自不能单独确定一个学生某门课程的成绩, 只有两者联合起来才能确定, 所以可记为完全函数依赖“(学号, 课程号) $\twoheadrightarrow$ 成绩”。

关系模式中, 如果仅存在一个主键, 则非主属性(不属于主键属性集中的属性)都完全函数依赖于主键。

3. 部分函数依赖

设 X' 是 X 的某一个特定的子集, 如果 $X \rightarrow Y$, $X' \nrightarrow Y$, 则称 Y 部分函数依赖于 X , 记为 $X \twoheadrightarrow Y$ 。

例如, 设有关系“学生(学号, 姓名, 性别, 班级编号, 班级名称)”, 因为有(学号, 班级编号) $\rightarrow$ 班级名称, 班级编号 $\rightarrow$ 班级名称, 所以可记为部分函数依赖“(学号, 班级编号) $\twoheadrightarrow$ 班级名称”。

关系模式中存在部分函数依赖是因为某些属性既完全函数依赖于组合键, 又完全函数依赖于组合键中的某个键。

4. 传递函数依赖

如果 $X \rightarrow Y$, 而且 $Y \rightarrow Z$, 但 $Y \nrightarrow X$, 则称 Z 传递函数依赖于 X , 记为 $X \twoheadrightarrow Z$ 。

例如, 设有关系“教学(课程号, 教师号, 教师电话)”, 因为有“课程号 $\rightarrow$ 教师号, 教师号 $\rightarrow$ 教师电话, 教师号 $\nrightarrow$ 课程号”, 所以可记为传递函数依赖“课程号 $\twoheadrightarrow$ 教师电话”。

关系模式中如果存在两类或两类以上不同性质的属性集，则关系模式必然存在传递函数依赖。

1.4.3 关系规范化处理

根据一个关系满足数据依赖程度的不同，关系可规范化为第一范式（1NF）、第二范式（2NF）、第三范式（3NF）、第四范式（4NF）、第五范式（5NF）等。实际应用时，通常只规范化到第三范式即能满足基本要求。

1. 第一范式

如果关系模式中的每个属性都是不可分解的基本数据项（简称原子属性），则称这个关系是满足第一范式的关系。第一范式是最低级别的范式。

【例 1-5】将表 1-4 中的关系规范化以满足第一范式的要求。

分析：因为“专业”属性包含“专业名称”“专业方向”两个原子属性，所以该关系不满足第一范式的要求，如果将“专业”属性分解为“专业名称”“专业方向”两个属性，就满足了第一范式的要求。为便于后面的第二、第三范式的规范化，表 1-4 改用关系模式表示如下：

综合信息表（学号，姓名，性别，专业名称，专业方向，教师编号，教师姓名，职称，课程编号，课程名称，成绩）

从关系模式可以看出，尽管关系满足第一范式的要求，但没有解决数据冗余和操作异常等问题，所以，需要引入第二、第三范式的处理。

2. 第二范式

如果关系模式中的所有非主属性都完全函数依赖于任意的候选键，则称这个关系是满足第二范式的关系。第二范式要求关系在满足第一范式的基础上，去除那些部分函数依赖于主键的非主属性，使得关系的所有非主属性都完全函数依赖于主键。

如果关系中出现非主属性对主键的部分函数依赖，则关系不满足第二范式，所以，规范化处理时，要从消除非主属性对主键属性的部分函数依赖入手。规范化的方法是，对关系进行投影处理，将关系模式分解为两个或多个新的关系模式。

【例 1-6】将“例 1-5”得到的关系模式规范化处理，以满足第二范式的要求。

分析：分析“综合信息表”关系模式，因为有“(学号，教师编号)→(教师姓名，职称)”，“教师编号→(教师姓名，职称)”，所以存在部分函数依赖；类似的，关系模式中还存在专业、课程等信息的部分函数依赖，所以关系模式不满足第二范式的要求。

根据消除部分函数依赖的操作，原有关系可拆分成四个关系：

- (1) 学生（学号，姓名，性别，专业名称，专业方向）
- (2) 教师（教师编号，教师姓名，职称）
- (3) 课程（课程编号，课程名称）
- (4) 成绩（学号，课程编号，教师编号，成绩）

关系模式分解后，消除了部分函数依赖，达到了第二范式的要求，但是，有些新的关系模式仍然因为存在“传递函数依赖”而导致操作异常，所以还要继续进行满足第三范式的规范化处理。

3. 第三范式

如果关系模式中的所有非主属性对任意候选键都不存在传递函数依赖，则称这个关系是满足第三范式的关系。第三范式要求关系在满足第二范式的基础上，除去所有传递函数依赖于主键的非主属性。

【例 1-7】将“例 1-6”得到的关系模式规范化处理，以满足第三范式的要求。

分析：因为满足第三范式的关系模式一定满足第二范式的要求，所以，第三范式的规范化处理可以从第二范式的规范化结果开始。分析“学生（学号，姓名，性别，专业名称，专业方向）”关系模式，因为“学号→专业名称，专业名称!→学号，专业名称→专业方向”，所以出现传递函数依赖“学号→专业方向”。这时，仍然需要将关系再次分解以消去传递函数依赖。

根据第三范式的要求，将上述“学生”关系模式进一步分解为如下三个关系模式：

- (1) 学生（学号，姓名，性别，方向编号）
- (2) 方向（方向编号，方向名称，专业编号）
- (3) 专业（专业编号，专业名称）

至此，“综合信息表”关系模式规范化完毕，得到六个满足第三范式要求的关系模式：

- (1) 学生（学号，姓名，性别，方向编号）
- (2) 方向（方向编号，方向名称，专业编号）
- (3) 专业（专业编号，专业名称）
- (4) 教师（教师编号，教师姓名，职称）
- (5) 课程（课程编号，课程名称）
- (6) 成绩（学号，课程编号，教师编号，成绩）

关系规范化还可以进行到第四、第五范式等，但并不是规范化程度越高越好，因为关系规范化会带来一些其他问题。例如，规范化程度越高，关系模式的表现力就越差，数据操纵就会越复杂，操纵的代价就会越高。因此，有时出于编程效率等方面的考虑，关系模式会人为地加入一些其他属性，使关系只满足第二范式的要求。



拓展阅读

中国数据库的先行者与探索者——萨师煊

“数据库是什么？”1978年，当中国人民大学经济信息管理系首任系主任萨师煊第一次将“数据库”三个字写在黑板上时，恢复高考后的第一批大学生也产生了同样的疑问。

这被视为数据库理论正式传入中国的标志性事件。

萨师煊教授是我国数据库学科的奠基人之一，作为数据库学术活动的积极倡导者和组织者，他在中国计算机学会中担任多项重要职务，如常务理事、软件专业委员会常务委员兼数据库学组组长，以及数据库专业委员会名誉主任委员。他以卓越的人格魅力和渊博的学识，团结了全国的数据库工作者，成为我国数据库界备受尊敬的组织者和领军人物。

萨师煊教授的贡献不仅在于技术和学术上的卓越成就，更在于他在人才培养和推动数据库学科发展方面所做的开创性工作。他的领导和努力使我国的数据库技术在全球范围内崭露头角，为中国的数据库技术发展奠定了坚实的基础。

项目归纳

(1) 数据模型分为两种, 一是信息模型, 二是结构数据模型。常见的结构数据模型有层次模型、网状模型和关系模型三种。一个完整的数据模型通常包括数据结构、数据操作和数据完整性约束等三个部分。当前流行的关系数据库管理系统都是关系模型的具体实现。

(2) E-R 模型是目前最常用的信息模型, 它从用户的观点来对数据和信息进行建模。由于系统分析人员的经验或见解不同, 针对同一个企业运营模式, 不同人员设计出来的 E-R 模型有所差异, 因此, E-R 模型设计不是唯一的。

(3) 关系数据库的操作主要有选择、投影和连接三种。其中, 选择操作用于从关系中选择满足条件的记录行; 投影操作用于从关系中选择满足需要的列; 连接操作可以将两个或多个表进行拼接, 它通常包含选择和投影两种操作。

(4) 关系规范化时, 第一范式的规范化使得关系模式中的每个属性都是不可分解的基本数据项; 第二范式的规范化是去除那些部分函数依赖于主键的非主属性, 使得关系模式中的所有非主属性都完全函数依赖于主键; 第三范式的规范化是除去关系模式中那些传递函数依赖于主键的非主属性。

拓展训练

(1) 使用关系规范化技术对表 1-5 中的订单信息进行规范化处理, 要求规范化后的关系模式能满足第三范式的要求。

表 1-5 订单信息

订单号	销售员	部门	订货量	订货日期	交货日期	客户	客户电话	货品名称	售价	成本价	供货商
8001	张明	销售部	20	2022-01-11	2022-03-21	王虹	136·····	电脑 (P4)	5100	4500	联想
8002	李洪	生产部	10	2023-01-01	2023-03-01	李明	159·····	打印机	4300	4000	爱普生
8003	王名利	销售部	5	2023-02-01	2022-05-01	张宇	133·····	显示器	2300	1900	三星
8004	李洪	生产部	20	2023-04-11	2023-05-20	王虹	136·····	电脑 (P5)	5100	4500	联想
8005	张明	销售部	10	2023-05-10	2023-05-25	张宇	133·····	显示器	2300	1900	三星

(2) 根据你对“行包托运”业务的了解情况, 使用 E-R 模型技术为从事行包托运的企业设计一个“行包托运管理系统”信息模型, 然后将 E-R 模型转化为关系模式表示。



拓展训练 1

项目 2

SQL Server 2022 的 安装与配置

❖ 知识目标

- (1) 了解 SQL Server 2022 的新特性。
- (2) 熟悉 SQL Server 2022 管理工具的功能。
- (3) 了解 SQL Server 2022 的安装要求。
- (4) 掌握 SQL Server 2022 的安装、配置。

❖ 能力目标

- (1) 能够阅读数据库系统有关的英文术语。
- (2) 能够在安装 SQL Server 2022 的过程中进行系统初步配置。
- (3) 能够根据具体需要进行 SQL Server 2022 组件管理与配置。
- (4) 能够使用 SSMS 的“对象资源管理器”连接服务器、查看数据库对象。

❖ 素质目标

- (1) 通过安装和配置 SQL Server 2022，培养学习和使用各种软件的能力。
- (2) 通过了解中国数据库产品的现状，树立高远的理想追求和深沉的爱国情怀，培养服务科技强国建设的责任感和使命感。

任务 2.1 认识 SQL Server 2022

SQL Server 2022 是微软公司在 2022 年年底推出的最新关系型数据库软件，它不仅包含以往各个版本的全部功能，还具有更好的云服务功能，能够对本地运营数据进行无缝分析，克服数据孤岛，实现整个数据资产的可见性，提高系统的安全性、合规性和可用性。

2.1.1 SQL Server 2022 的新特性

SQL Server 2022 基于 SQL Server 2019 的强大功能，提升了 SQL Server 2019 的性能及可信性、高效性和智能性，使得公司可以以很高的安全性、可靠性和可扩展性来运行其最关键任务的应用程序，从而降低开发和管理其数据基础设施的各成本。

SQL Server 2022 的新特性如表 2-1 所示。

表 2-1 SQL Server 2022 的新特性

新特性	说明
通过 Azure 实现业务连续性	通过 Azure SQL 托管实例中的链接功能，帮助确保云中完全托管灾难恢复的正常运行时间，持续将数据复制到云端或从云端复制数据
对本地运营数据进行无缝分析	通过 Azure Synapse Link 在云中使用 Spark 和 SQL 运行时分析所有数据，打破运营商店和分析商店之间的壁垒
整个数据资产的可见性	使用 Microsoft Purview 管理数据资产以克服数据孤岛
行业领先的性能和可用性	利用性能和可用性来加快查询速度并帮助确保业务连续性，无需更改代码即可提升查询性能，为跨多个位置的用户保持多写入环境平稳运行

2.1.2 SQL Server 2022 的管理工具

SQL Server 2022 的管理工具需要单独安装，管理工具及其说明如表 2-2 所示。

表 2-2 管理工具及其说明

管理工具	说明
SQL Server 集成管理器 (SQL Server Management Studio, SSMS)	SSMS 是一个集成环境，用于访问、配置、管理和开发 SQL Server 的所有组件。SSMS 使各种技术水平的开发人员和管理员都能使用 SQL Server。SSMS 的安装需要 Internet Explorer 8 SP1 或更高版本的浏览器
SQL Server 配置管理器 (SQL Server Configuration Manager)	SQL Server 配置管理器为 SQL Server 服务、服务器协议、客户端协议和客户端别名提供基本配置管理
SQL Server 事件探查器 (SQL Server Profiler)	SQL Server 事件探查器提供了一个图形用户界面，用于监视数据库引擎实例和分析服务实例
数据库引擎优化顾问 (Database Engine Tuning Advisor)	数据库引擎优化顾问用于协助创建索引、索引视图和分区最佳组合
商业智能开发平台 (Business Intelligence Development Studio, BIDS)	BIDS 是分析服务、报表服务和集成服务解决方案的集成开发环境 (Integrated Development Environment, IDE)，其安装需要 Internet Explorer 6 SP1 或更高版本的浏览器

续表

管理工具	说明
连接组件（Connecting Component）	安装用于客户机和服务器之间通信的组件，以及用于 DB-Library、ODBC 和 OLEDB 的网络库

任务 2.2

安装 SQL Server 2022

2.2.1 SQL Server 2022 的安装要求

SQL Server 2022 的版本、硬件要求及软件要求如下。

1. 版本

SQL Server 2022 提供了多种版本以满足不同用户的要求，各版本特点如表 2-3 所示。

表 2-3 SQL Server 2022 版本特点

版本	特点
企业版（Enterprise）	企业版是一个全面的数据管理和业务智能平台，提供大型企业进行联机事务处理（OLTP）、高度复杂的数据分析，数据仓库系统和网站所需的高性能支持
标准版（Standard）	标准版提供电子商务、数据仓库和解决方案所需的基本功能，适用于中小型企业的数据管理和分析
开发版（Developer）	开发版包括了企业版功能，它使开发人员可以建立、测试基于 SQL Server 的应用程序，是独立软件供应商、咨询人员、系统集成商、解决方案供应商、创建和测试应用程序的企业开发人员的理想选择
Web 版（Web）	Web 版是针对 Windows 服务器、面向 Web 服务的环境而设计。Web 版为实现低成本、大规模、高可用性的 Web 应用或客户托管解决方案提供了必要的工具
速成版（Express）	速成版也称精简版和简易版，可免费提供，是学习和构建桌面及小型服务器应用程序的理想选择，也是独立软件供应商、非专业开发人员和热衷于构建客户端应用程序的人员的最佳选择

2. 硬件要求

在安装 SQL Server 2022 之前，需要检查计算机系统的软、硬件配置是否符合要求，然后才能进入安装阶段。硬件配置主要从处理器、内存、硬盘等主要设备的要求去检查。

1) 处理器

要正常安装和使用 SQL Server 2022，建议使用 2.0 GHz 或更快的处理器。

2) 内存

SQL Server 2022 需要的内存数量最少为 1 GB，为了获得数据处理的最高效率，内存越大越好，推荐使用 4 GB 以上的内存，同时建议使用专用的服务器。

3) 硬盘

SQL Server 2022 对硬盘的速度要求不高，但同样越快越好，且剩余空间不要少于 6 GB。

3. 软件要求

安装 64 位版本和 32 位版本的 SQL Server 2022 的软件要求相同，都对网络协议、.NET Framework、Internet 软件等网络组件有一定的要求，如表 2-4 所示。

表 2-4 安装 SQL Server 2022 的软件要求

软件要求	说明
网络协议	SQL Serer 2022 支持的操作系统具有内置网络软件。独立安装的命名实例和默认实例支持以下网络协议：共享内存、命名管道、TCP/IP 和 VIA
.NET Framework	在选择数据库引擎（Database Engine）、报表服务（Reporting Services）、主数据服务（Master Data Services）、数据质量服务（Data Quality Services）等时，.NET 3.5 SP1 是 SQL Server 2022 必需的，但不再由 SQL Server 安装程序安装
Internet 软件	微软管理控制台（Microsoft Management Console, MMC）、SQL Server 数据工具（SQL Server Data Tool, SSDT）、报表服务的报表设计器组件和 HTML 帮助都需要 Internet Explorer 8 或更高版本的浏览器

2.2.2 SQL Server 2022 的安装步骤

下载 SQL Server 2022 开发版后，按照以下步骤完成 SQL Server 2022 的安装。

（1）关闭防火墙，保持网络通畅，并以管理员身份运行 SQL2022-SSEI-Dev.exe，在弹出的选择安装类型界面中选择“基本”，如图 2-1 所示。



图 2-1 选择安装类型界面

（2）在弹出的选择语言界面中，默认选择“中文（简体）”，单击“接受”按钮，如图 2-2 所示。



图 2-2 选择语言界面

（3）在弹出的指定安装位置界面中，默认安装位置是 C 盘，由于操作系统通常安装在 C 盘，为避免占用系统盘的空间，建议切换到其他盘，单击“浏览”按钮可以设置其他位置，本书安装的位置为“D:\Program Files\Microsoft SQL Server”，安装位置设置完成后，单击“安装”按钮，如图 2-3 所示。

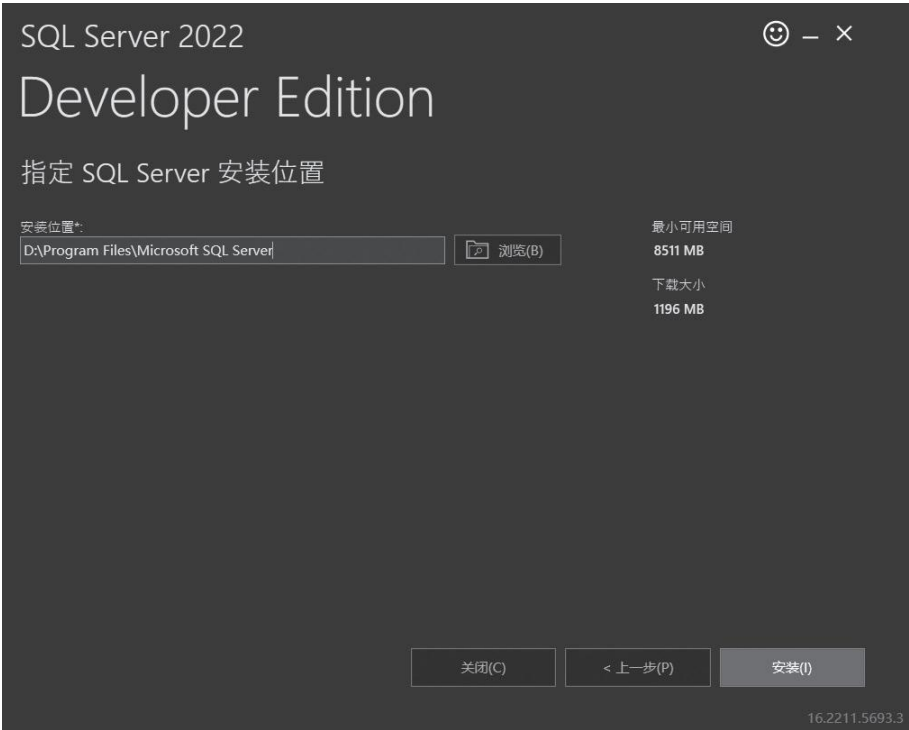


图 2-3 指定安装位置界面

(4) 系统会自动从网站下载相关文件并自动启动安装，直到安装完成，安装过程界面如图 2-4 所示。

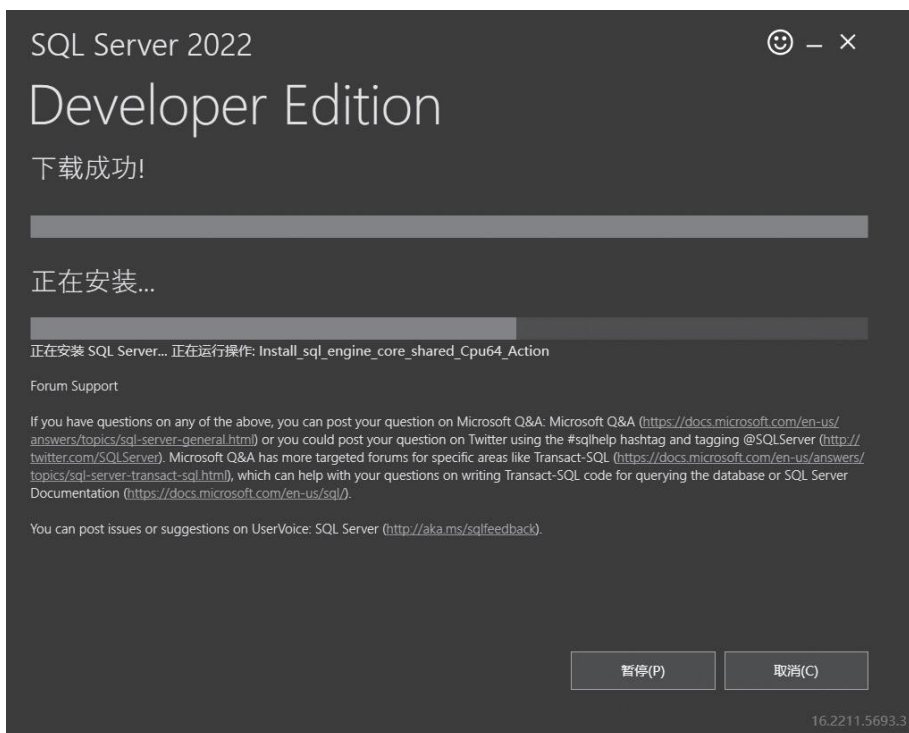


图 2-4 安装过程界面

(5) 安装完成后，界面如图 2-5 所示。

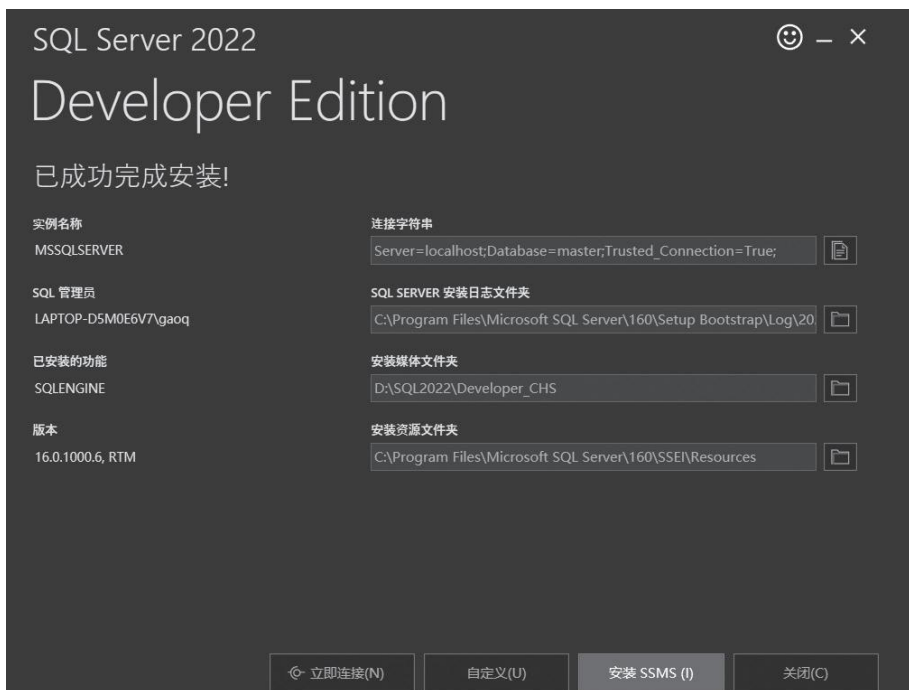


图 2-5 SQL Server 2022 安装完成界面

（6）在如图 2-5 所示的安装完成界面中单击“安装 SSMS”按钮，系统会打开下载 SSMS 页面（图 2-6），单击“下载 SQL Server Management Studio (SSMS) 20.1”链接，下载 SSMS-Setup-CHS.exe。



图 2-6 下载 SSMS 界面

（7）以管理员身份运行 SSMS-Setup-CHS.exe，在弹出的指定安装位置界面中，默认安装位置是 C 盘，与步骤（3）相似，建议将其安装到别的盘和文件夹，单击“更改”按钮可以设置其他位置，本书安装的位置为“D:\Program Files\Microsoft SQL Server Management Studio 20”，如图 2-7 所示。安装位置设置完成后，单击“安装”按钮，开始安装。



图 2-7 指定 SSMS 安装位置界面

安装完成界面如图 2-8 所示。如果出现错误导致安装无法完成，则需要根据错误提示查找解决方法，重新安装。



图 2-8 SSMS 安装完成界面

(8) 单击图 2-8 中的“关闭”按钮，结束 SSMS 的安装。返回如图 2-5 所示的界面，单击“关闭”按钮，在弹出的对话框中单击“是”按钮，如图 2-9 所示，结束 SQL Server 2022 的安装。

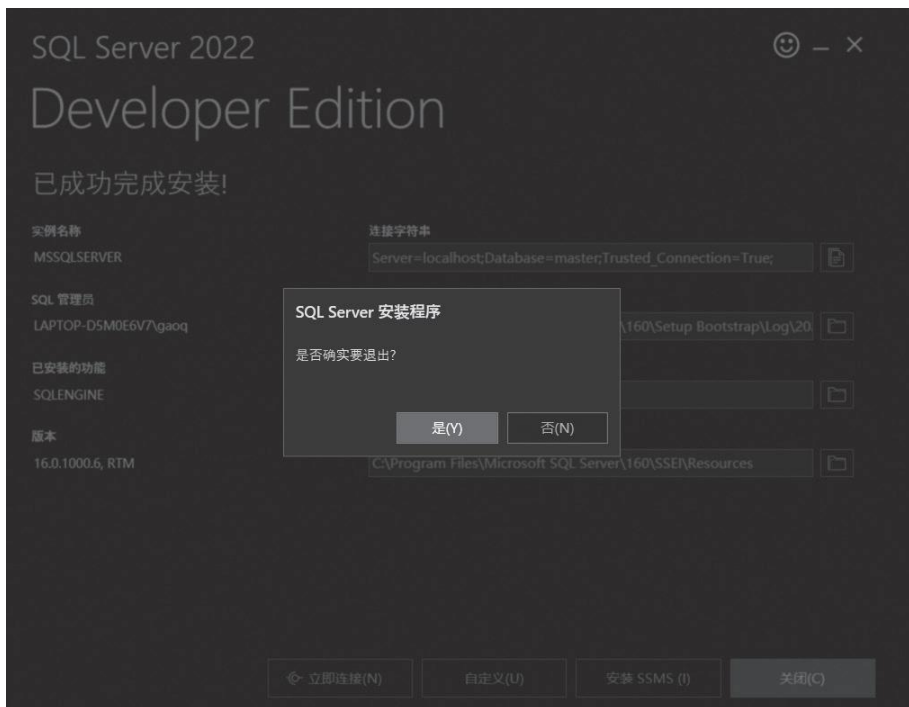


图 2-9 结束 SQL Server 2022 安装界面

任务 2.3 配置 SQL Server 2022

SQL Server 2022 安装结束后，需要进一步配置才能使用，常规的配置主要有 TCP/IP 协议配置、远程登录配置、“sa”管理员登录账户配置和用户权限配置等，其中用户权限配置等内容将在项目 9 “数据库安全管理”中详细介绍。

2.3.1 SQL Server 网络配置

SQL Server 2022 数据库管理系统对实例的操作主要使用 TCP/IP 协议进行通信，所以，系统安装成功后，还要配置 TCP/IP 协议参数。操作过程如下。

1. 进入 SQL Server 配置管理器

执行“开始”→“所有应用”→“Microsoft SQL Server 2022”→“SQL Server 2022 配置管理器”菜单项命令，弹出如图 2-10 所示的 SQL Server 配置管理器窗口。

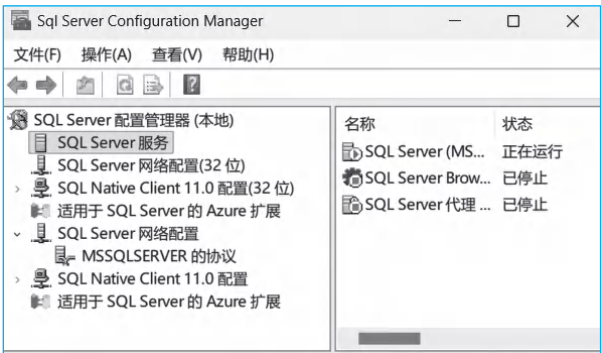


图 2-10 SQL Server 配置管理器窗口

2. 启用实例的 TCP/IP 协议

SQL Server 2022 安装后，TCP/IP 协议默认是“已禁用”的，必须把它设置为“已启用”状态。展开窗口左边的“SQL Server 网络配置”显示出所有实例的协议项，单击“MSSQLSERVER 的协议”项，窗口的右边显示出该实例的协议状态，如图 2-11 所示。

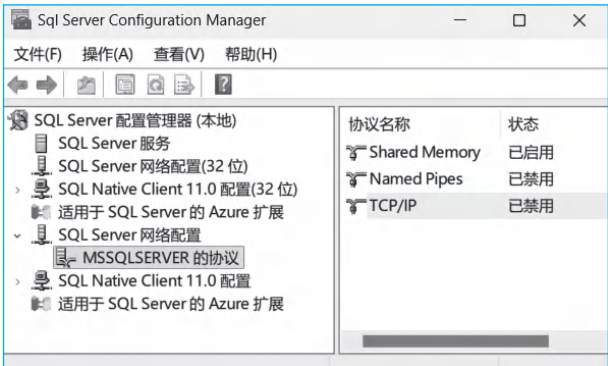


图 2-11 实例协议状态

右击“TCP/IP”项，弹出快捷菜单（图 2-12），单击“启用”菜单项，弹出“警告”对话框，提

示需要重新启动该实例的服务协议才能生效，如图 2-13 所示。

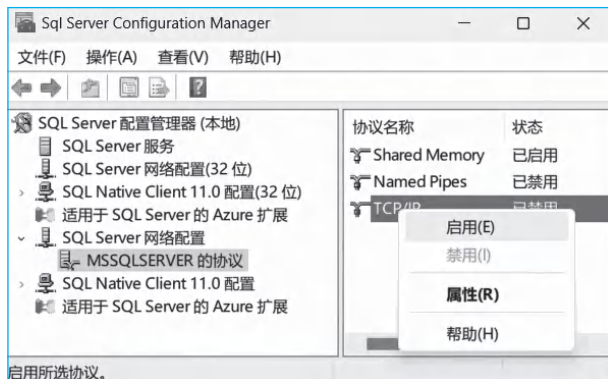


图 2-12 启用 TCP/IP 协议

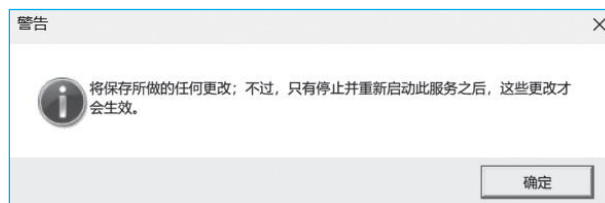


图 2-13 “警告”对话框

3. 配置 TCP/IP 协议属性

配置 TCP/IP 协议的属性主要是设置 TCP 端口号和端口的状态。SQL Server 默认使用的端口号为“1433”。在如图 2-12 所示的窗口中，单击快捷菜单的“属性”项，在属性窗口单击“IP 地址”选项卡，“IP 地址”选项卡列出了各个 IP 地址的配置情况，找到本机的 IP 地址项，将其 TCP 端口设置为“1433”，将“活动”和“已启用”项设置为“是”，如图 2-14 所示。

继续找到“IPALL”地址项，将其 TCP 端口号也设置为“1433”，如图 2-15 所示。上述内容设置完成后，需要停止该项服务然后再重新启动该项服务或重启计算机。



图 2-14 设置本机 IP 的 TCP 端口号



图 2-15 设置 IPALL 的 TCP 端口号

4. 操作 SQL Server 服务

在图 2-10 左边的区域单击“SQL Server 服务”项，右边区域显示已安装的所有服务项，如图 2-16 所示。右击“SQL Server(MSSQLSERVER)”服务项，弹出快捷菜单，如图 2-17 所示。

1) 对 SQL Server 服务进行操作

通过如图 2-17 所示的快捷菜单，可以对该项服务进行启动、停止、暂停、继续、重启等操作。这些操作的效果与在 Windows 操作系统中，执行“控制面板”→“管理工具”→“服务”命令所打开的“服务”窗口的相应操作效果一样。

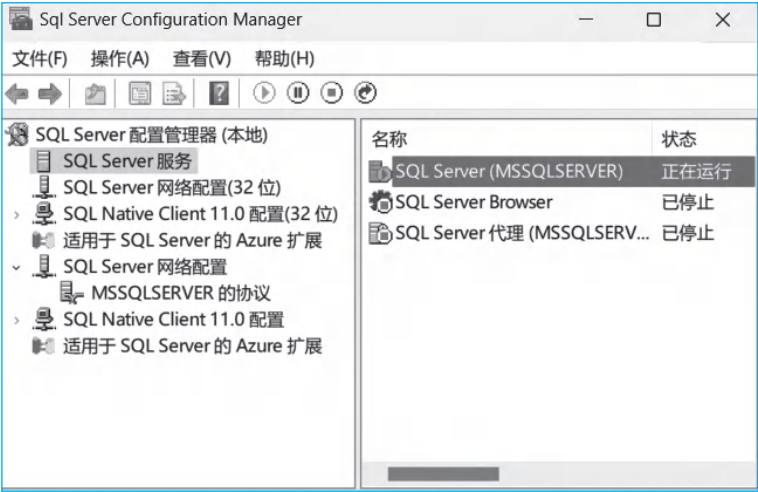


图 2-16 SQL Server 服务

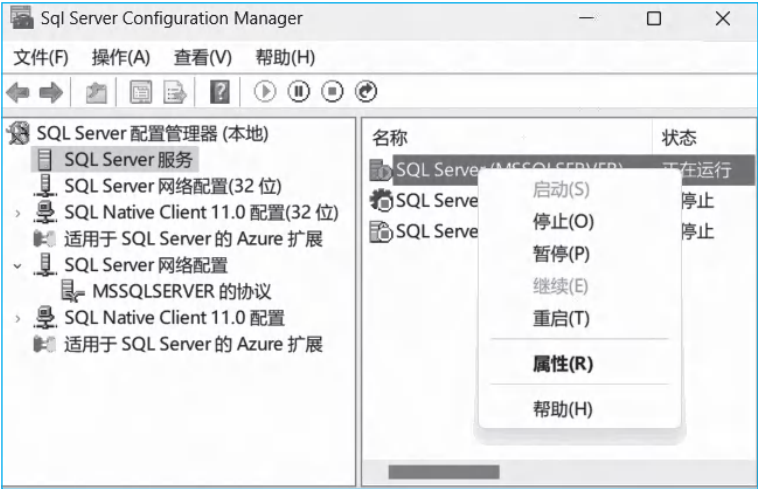


图 2-17 SQL Server 服务快捷菜单

2) 配置 SQL Server 服务属性

SQL Server 2022 安装时所设置的有关参数在“SQL Server 服务”的属性窗口可以进行修改。在如图 2-17 所示的快捷菜单中，单击“属性”项，进入服务属性窗口，窗口中有四个选项卡。“登录”选项卡如图 2-18 所示，主要用于设置登录身份、管理员账户^①及其密码等。其中，登录身份通常设置

^① 本书正文中的“账户”同软件界面中的“帐户”。

为“本地服务”；“服务”选项卡如图 2-19 所示，主要用于设置服务的启动方式，其中，“自动”的启动模式是指在计算机开机时自启动，常用的服务通常设置“自动”模式，“已禁用”模式表示该服务已被禁用，“手动”模式表示需要通过手工操作启动该项服务。



图 2-18 属性 - 登录选项卡



图 2-19 属性 - 服务选项卡

2.3.2 配置远程登录

配置 SQL Server 2022 远程登录的目的，一是将应用程序与数据库服务器分离，使得应用程序与数据库管理系统可以分别安装在不同的专用服务器上；二是使得数据库管理员可以远程登录数据库服务器，实现远程管理控制。远程登录主要使用主机名称或 IP 地址对异地服务器的数据库进行登录操作。远程登录配置的操作如下。

1. 连接 SQL Server 2022 服务器

通过网络配置，安装 SQL Server 2022 的计算机已经可以连接到数据库服务器，具体操作如下：执行“开始”→“所有应用”→“Microsoft SQL Server Tools 20”→“SQL Server Management Studio 20”菜单项命令，弹出如图 2-20 所示的“连接到服务器”窗口，操作如下。

- (1) 服务器类型：选用“数据库引擎”。
- (2) 服务器名称：采用默认名称即可。
- (3) 加密：选用“可选”。
- (4) 身份验证：首次连接数据库服务器时，主要用于配置参数，请使用管理员身份登录操作系统，此处选用“Windows 身份验证”模式。
- (5) 单击“连接”按钮，如果前面的所有安装与配置有问题，则弹出错误提示窗口。否则，则进

入 SQL Server Management Studio 管理器（简称 SSMS 管理器，是 SQL Server 2022 的重要管理工具，SQL Server 2022 的绝大部分操作将在 SSMS 管理器环境下完成）操作窗口，如图 2-21 所示。



图 2-20 “连接到服务器”窗口

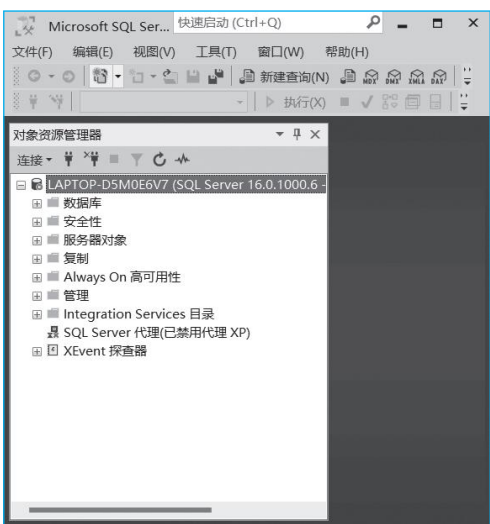


图 2-21 SSMS 管理器

2. 配置远程登录

在图 2-21 左边区域的“对象资源管理器”中右击“LAPTOP-D5M0E6V7”服务器名称，弹出快捷菜单，单击“属性”菜单，弹出“服务器属性”窗口，在窗口左边的“选择页”列表中，单击“连接”选项，然后勾选“允许远程连接到此服务器”。

3. 配置服务器身份验证

在图 2-21 左边的列表中单击“安全性”选项，然后在“服务器身份验证”一项中选择“SQL Server 和 Windows 身份验证模式”（即使用混合模式），其他保持默认值。最后单击“确定”按钮完成配置。

2.3.3 配置“sa”管理员账户

使用“Windows 身份验证”连接数据库服务器一般只限于数据库管理员的操作，其他情况一般使用“混合模式”进行连接（登录），其中“sa”账户是系统内置的权限最高的管理员账户，拥有“sa”账户的密码，意味着操作员能完全控制 SQL Server 2022 数据库管理系统的一切操作。其他账户一般需要“sa”账户授权后才能使用。

配置“sa”登录账户的主要内容有密码设置、状态设置，其他数据项视需要而定，一般使用默认值即可。

1. 配置 sa 账户密码

在如图 2-21 所示的“对象资源管理器”窗口中，展开“安全性”节点，再展开“登录名”节点，右击“sa”项，在快捷菜单中单击“属性”，弹出“登录属性-sa”配置窗口。

在“登录属性-sa”窗口中单击“常规”选项，然后在右边区域设置“sa”账户的登录密码，如图 2-22 所示。因为“sa”账户具有最高级别的操作权限，所以，密码设置必须有足够的强度。

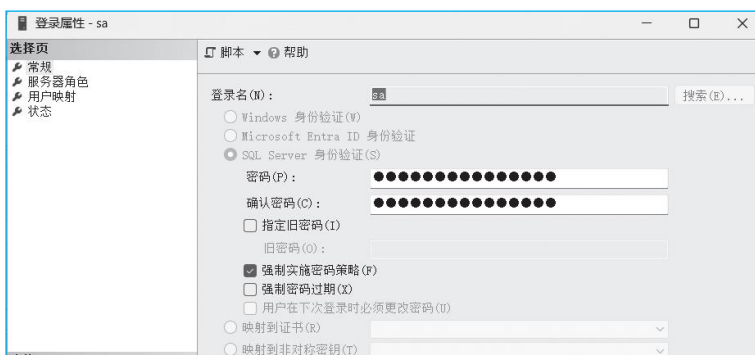


图 2-22 设置“sa”账户密码

2. 配置“sa”账户状态

在如图 2-22 所示的对话框中，单击“状态”选项，将“是否允许连接到数据库引擎”项设置为“授予”，将“登录”项设置为“启用”，其他选项保持默认值，单击“确定”按钮完成“sa”账户的常规设置。

2.3.4 使用混合模式登录

数据库管理员远程登录管理数据库或应用程序从数据库中提取数据等操作都是使用混合模式登录实例，在安装数据库管理系统的计算机上也常用这种方式登录实例进行操作。使用混合模式登录时，需要提供具有访问权限的账户名称及密码。操作过程如下。

运行 SSMS 管理器，打开“连接到服务器”窗口，如图 2-20 所示，连接参数设置如下。

- (1) 服务器类型：选用“数据库引擎”。
- (2) 服务器名称：采用默认名称。
- (3) 身份验证：使用混合模式时，只能选用“SQL Server 身份验证”。
- (4) 登录名：输入账户名称，这里使用“sa”管理员账户。
- (5) 密码：输入“sa”账户的密码。

(6) 单击“连接”按钮，如果网络配置和“sa”账户的密码不正确，则弹出错误提示窗口。否则，进入“SSMS 管理器”操作窗口，如图 2-21 所示。

拓展阅读

华为——高斯数据库

华为、阿里和中兴旗下的数据库产品是当前国产数据库领域的三大支柱，展现着我国在信息技术自主研发方面的巨大进步。通过这些科技巨头的不断研发和创新，越来越多的中国本土企业开始采用国产数据库，这不仅提高了技术自主性，还极大增强了我国信息数据的安全性。

华为自 2007 年开始研发的高斯数据库（Gauss DB），经过 17 年的不断优化和三代产品的迭代发展，如今已经成为国产数据库中的佼佼者，位居国产数据库应用数量首位。高斯数据库的技术水平已经得到了国内顶尖金融机构如招商银行、工商银行的验证，并且达到了国内三大电信运营商的使用标准，为广大用户的通信和数据安全提供了坚实的保障。这些成就不仅展示了华为在数据库技术上的领先地位，也体现了我国科技企业在信息化自主创新道路上的不懈努力和重大贡献。

项目归纳

（1）SQL Server 2022 有 5 个版本，每个版本都具有不同的特征及不同的安装要求。在安装 SQL Server 2022 的过程中有几个问题需要注意：定义实例、设置服务账号和选择身份验证模式。

（2）安装 SQL Server 2022 之前，需要检查计算机系统的硬件和软件配置是否符合要求，然后才能进入安装阶段。

（3）SQL Server 2022 提供了丰富的管理工具集，其中，SSMS 是最重要的管理工具，它集图形工具和脚本编辑器于一体，可用于访问、配置和管理 SQL Server 的产品组件。

（4）SQL Server 配置管理器用于启动、暂停、恢复和停止 SQL Server 相关服务，还可以配置网络协议和管理网络配置。

拓展训练

（1）SQL Server 2022 包含哪些版本？它们的主要特点是什么？

（2）SQL Server 2022 提供了哪些管理工具？它们的主要作用是什么？



拓展训练 2