

1

数据库概述

任务 1 需求分析

数据库技术是数据管理的核心技术，是计算机科学的重要分支。数据库技术已应用于社会的各个领域。需求分析是整个数据库设计的基础，是系统开发中的一个关键过程。

现给出一个项目实例，本书的所有章节均围绕该项目实例展开。某高校要开发一个“学生信息管理系统”，业务需求为：能够满足学校对院系信息、班级信息、课程信息、学生学籍信息以及学生成绩信息的管理需求；能够满足各用户对上述信息的查询需求；能够满足各用户对学籍信息和学生成绩信息的分类统计需求。

1.1.1 任务分析

了解数据、数据库、数据库管理系统和数据库设计的基本概念；掌握进行需求分析的过程和方法。

1.1.2 相关知识

数据是数据库中存储的基本对象，是描述客观事物及其活动的抽象符号。数据的种类很多，它不仅包括数值，还包括文本、图像、音频和视频等，它们都可以经过数字化后存入计算机。

随着互联网技术的迅速发展，人类已经开始进入了大数据时代。这些数据有些是有用的，有些是无用的。对于软件开发人员，他们首先要做的就是从大量的数据中提取出对所研发系统有用的数据，并将这些数据进行整理、分类、组织，然后存储在仓库内，这个仓库就是数据库。所谓的数据库就是指长期存储在计算机内的、有组织的、可共享的、可统一管理的数据集合。

数据库中的数据按一定的数据模型组织、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，并可为各种用户共享。理解数据库最简单的方法，就是把它想象成一个容器，它是存放数据的物理位置。

数据库管理系统（Database Management System）是一种操纵和管理数据库的系统软件，用于建立、使用和维护数据库，简称DBMS。它对数据库进行统一的管理和控制，以保证数据库的安全性和完整性。用户通过DBMS访问数据库中的数据，数据库管理员也通过DBMS进行数据库的维护工作。DBMS 可使多个应用程序和用户用不同的方法在同一时刻或不同时刻去建立、修改和访问数据库。数据库的逻辑数据模型分为层次模型、网状模型、关系模型及面向对象模型。现在主流的数据库管理系统都是关系型数据库，如MS SQL Server、Oracle、Sybase、DB2、MySQL和Access等。关系型数据库是建立在关系模型的基础上，由一系列的二维表格组成，将数据分类存储在多个二维表中，操作起来相对容易。

数据库设计（Database Design）是指根据用户的需求，在某一具体的数据库管理系统上，设计数据库的结构和建立数据库的过程。具体来说就是针对要研发的应用系统，首先进行需求分析，并将分析结果以业务流程图、系统功能结构图等形式展现出来；然后在需求分析的基础上进行概念模型设计，再将设计好的概念模型转化为关系模型，并对关系模型进行规范化处理；最后根据规范后的关系模型来创建数据库及其应用系统，以满足用户的需求。

需求分析的任务就是明确用户的各种需求，并在此基础上确定所研发系统的功能。系统功能必须充分考虑今后可能的扩充和改变，不能仅仅按当前应用需求来设计数据库。在这个过程中，系统分析员和软件工程师确定用户的需求，只有在确定了这些需求后，他们才能够分析和寻求新系统的解决方法。需求分析是否做得充分与准确，决定了在此基础上构建数据库大厦的速度与质量。需求分析做得不好，甚至会导致整个数据库设计返工重做。因此需求分析在软件开发的过程中具有举足轻重的地位，在一个大型软件系统的开发中，它的作用甚至要远远大于程序设计。

1.1.3 任务实施

绘制业务流程图和系统功能结构图，可以使用多种绘图软件，也可以使用一些专门的数据库设计工具。其中的 Microsoft Office Visio 是一款多功能绘图工具，下面将介绍如何使用该软件的一般绘图工具，来绘制“学生信息管理系统”的业务流程图和系统功能结构图。

1. 画出“学生信息管理系统”的业务流程图

步骤 1：启动 Microsoft Office Visio，单击主菜单“文件”|“新建”|“流程图”|“基本流程图”，如图 1-1 所示。

步骤 2：在左边的“基本流程图形状”窗格上单击“进程”（矩形）标签，并将其拖动到绘图页上表示业务流程图的“活动”；还可以使用复制、粘贴的功能绘制相同的形状；选中要进行布局的形状，单击主菜单“形状”|“对齐形状”，可进行形状的布局，如图 1-2 所示。

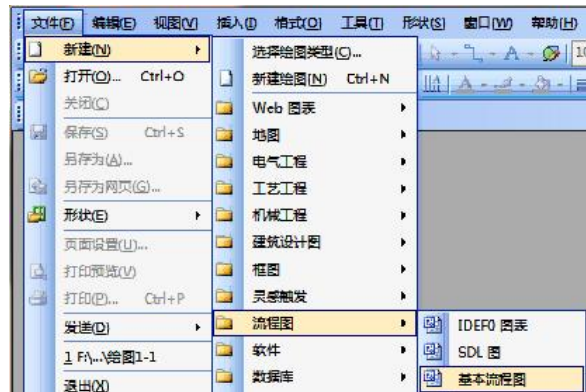


图 1-1 新建业务流程图文件

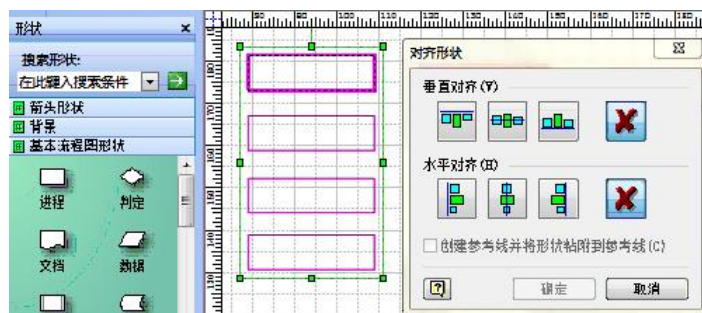


图 1-2 绘制业务流程图的“活动”

步骤 3: 单击工具栏中的“连接线工具”, 可以在绘图页上绘制业务流程图的“逻辑关系线”。单击工具栏中的“线条粗细”“线型”和“线端”, 可以对逻辑关系线进行相关设置。也可以单击工具栏中的“绘图工具”, 手动绘制矩形、圆形、椭圆形、直线、曲线等形状, 如图 1-3 所示。

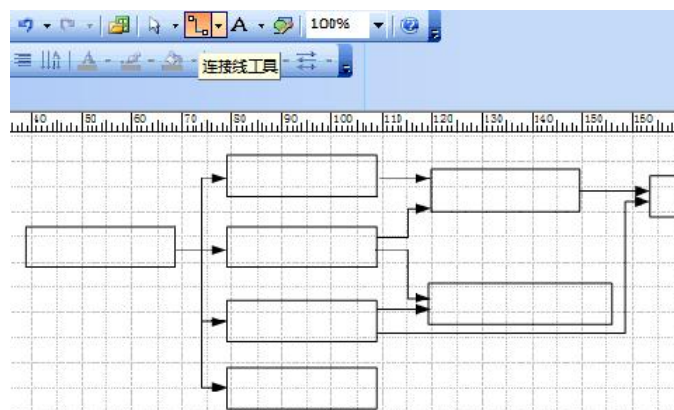


图 1-3 绘制业务流程图的“逻辑关系线”

步骤 4: 单击绘图页中的形状, 即可在其中输入文字, 单击工具栏中的“**A 文本工具**”, 可在绘图页上输入文字, 选择主菜单或快捷菜单的“**格式**”|“**文本**”/“**线条**”/“**填充**”等可进行形状的格式化, 如图 1-4 所示。

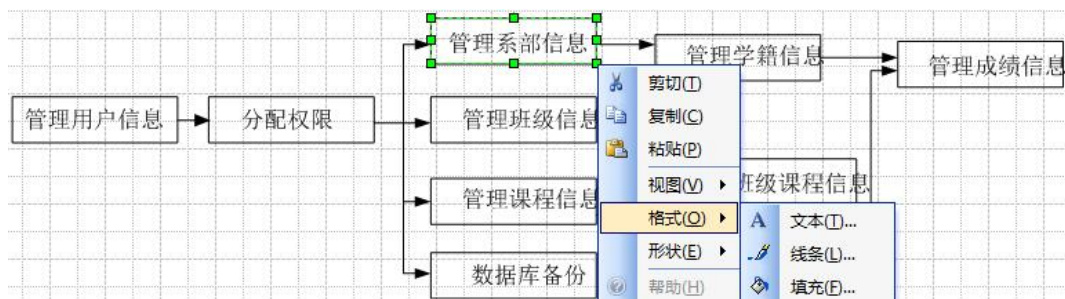


图 1-4 业务流程图的文字输入及格式化

步骤 5: 最终绘制好的“学生信息管理系统”业务流程图如图 1-5 所示。

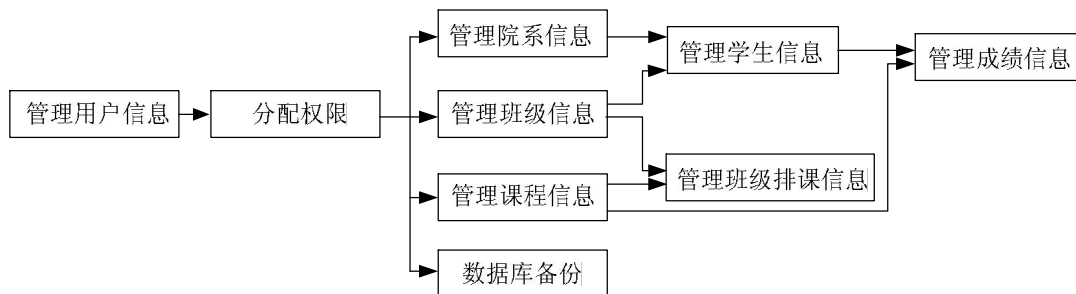


图 1-5 “学生信息管理系统”业务流程图

2. 画出“学生信息管理系统”的系统功能结构图

步骤 1: 启动 Microsoft Office Visio, 单击主菜单“**文件**”|“**新建**”|“**流程图**”|“**基本流程图**”, 新建一个系统功能图文件。

步骤 2: 在左边的“**基本流程图形状**”窗格上单击“**进程**”(矩形)标签, 并将其拖动到绘图页上表示系统功能结构图的“**功能**”; 还可以使用复制、粘贴的功能绘制相同的形状; 选中要进行布局的形状, 单击主菜单“**形状**”|“**对齐形状**”, 可进行形状的布局。

步骤 3: 单击工具栏中的“**连接线工具**”, 可以在绘图页上绘制箭头, 表示系统功能结构图中各功能之间的从属关系。单击工具栏中的“**线条粗细**”“**线型**”和“**线端**”, 对箭头线进行相关设置。也可以单击工具栏中的“**绘图工具**”, 手动绘制矩形、圆形、椭圆形、直线和曲线等形状。

步骤 4: 单击绘图页中的形状, 即可在其中输入文字, 单击工具栏中的“**A 文本工具**”, 可在绘图页上输入文字, 选择主菜单或快捷菜单的“**格式**”|“**文本**”/“**线条**”/“**填充**”等可进行形状的格式化。

步骤 5：最终绘制好的“学生信息管理系统”的系统功能结构图如图 1-6 所示。

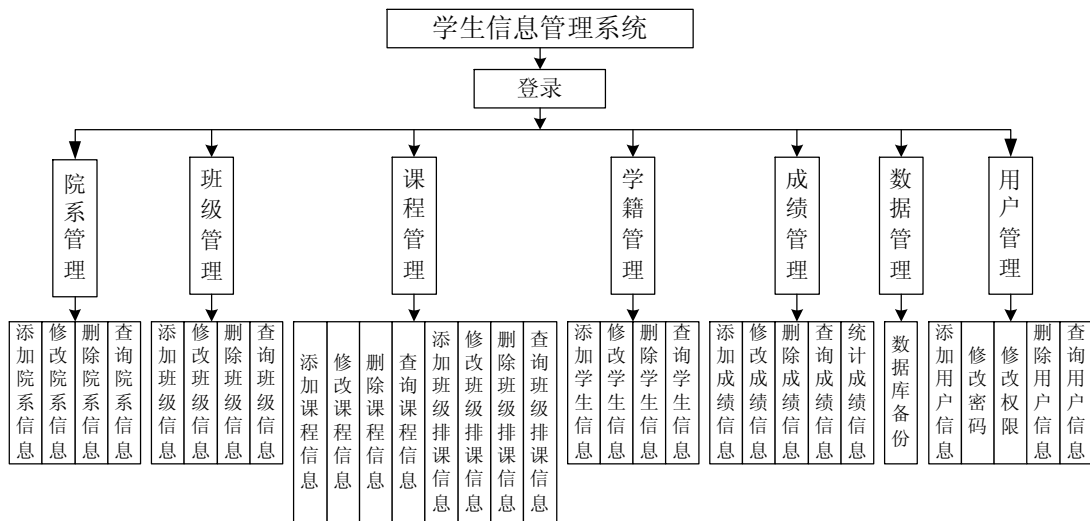


图 1-6 “学生信息管理系统”的系统功能结构图

3. “学生信息管理系统”的系统功能分析

通过图 1-6，可以得出“学生信息管理系统”的功能需求如下：

院系管理功能：用于添加、修改、删除和查询院系信息。

班级管理功能：用于添加、修改、删除和查询班级信息。

课程管理功能：用于添加、修改、删除和查询课程信息；用于添加、修改、删除和查询班级排课信息。

学籍管理功能：用于添加、修改、删除、查询和分类查询学生的学籍信息。

成绩管理功能：用于添加、修改、删除、查询和分类统计学生的课程考试成绩信息。

数据管理功能：用于对数据库进行备份操作。

用户管理功能：用于添加、删除和查询用户信息，还可以进行修改密码和修改权限的操作。

权限说明：该系统的用户类别有三个，分别是 **admin**（管理员用户）、**teacher**（教师用户）和 **student**（学生用户），不同的用户具有不同的权限。其中管理员用户具有所有功能的操作权限及修改密码功能，可以对所有用户进行修改，还可以修改不同用户的用户类别。教师用户除了具有对所有信息的查询权限外，还具有对学生成绩的添加、修改、删除和分类统计功能的操作权限，而对于修改密码功能，只能修改自己的密码。学生用户具有对所有信息的查询权限，对于修改密码功能，也只能修改自己的密码。学生用户可由系统自动生成，当添加学生信息时，系统会根据该条信息自动生成用户名和密码均为学生学号的学生用户；当删除学生信息时，系统也会自动删除由该学生生成的用户信息。

进入该系统时，首先要进行登录。然后根据登录的用户类别，决定该用户可以看到主界面

的哪些菜单选项。通过设置权限，可以提高系统的安全性。

任务 2 概念模型设计

将需求分析得到的用户需求抽象为信息结构（即概念模型）的过程就是概念结构设计，它是整个数据库设计的关键。

根据任务 1 的需求分析结果，设计“学生信息管理系统”的概念模型。

1.2.1 任务分析

掌握实体、属性和联系等概念模型的基本要素，会使用实体联系图（E-R 图）来描述实体、属性和实体间的联系。

1.2.2 相关知识

实体：指现实世界中客观存在并可以相互区别的事物，可以是人或物等实际对象，也可以是一个具体的概念。例如：学生、课程都是现实世界中客观存在并可相互区分的事物，所以学生、课程都是实体。

属性：实体或联系所具有的某些特征称之为属性，一个实体或一个联系可以使用许多属性来描述。例如：学生实体可以用学号、姓名、性别、出生年月、政治面貌、班级编号、班级名称、家庭住址、邮编和备注等属性来描述。

属性值：属性的具体取值，称之为该属性的属性值。例如：“秦伊诺”就是学生实体姓名属性的具体取值，所以“秦伊诺”就是学生实体姓名属性的属性值。

实例：实体表示的是一类事物，其中的一个具体事物称之为该实体的一个实例。例如：“01201401001”“秦伊诺”“女”“1995-10-08”“团员”“01201401”“14 级计算机应用技术”，“河南省信阳市平桥区文喜小区 14 栋 2 单元 302”“464000”“品学兼优”，则具体表示了一个学生，它就是学生实体的一个实例。

关键字：能够唯一标识出实体或联系的属性或属性的组合，称之为该实体或该联系的关键字，关键字也称为码。例如：“学号”这个属性可以唯一标识“学生”实体，因此“学号”可以作为“学生”实体的关键字；由于可能出现的学生重名现象，导致“姓名”的属性值不唯一，所以“姓名”这个属性不能唯一标识“学生”实体，不是“学生”实体的关键字。

联系：实体之间的关系称之为联系。实体之间的联系有 3 种，即一对一（1:1）、一对多（1:n）和多对多（m:n）。例如：一个学生只能属于一个班级，而一个班级中可以有多个学生，所以班级实体和学生实体之间的联系是一对多的联系，记为 1:n。

E-R 图：通常使用 E-R 图（或称 E-R 模型）来描述现实世界的概念模型，E-R 图提供了表示实体、属性和联系的方法：

- (1) 实体：用矩形表示，矩形框内写明实体名称。

(2) 属性：用椭圆形表示，椭圆框内写明属性的名称，并用无向边将其与相应的实体连接起来。

(3) 联系：用菱形表示，菱形框内写明联系的名称，并用无向边分别与有关实体连接起来，同时在无向边旁标上联系的类型（1:1, 1:n, m:n）。需要注意的是，如果一个联系具有属性，则这些属性也要用无向边与该联系连接起来。

在绘制 E-R 图时，一张图往往无法表现出整个系统的 E-R 模型，这时可以将其分解成多张图。绘制 E-R 图的工具有很多，我们仍选用 Microsoft Office Visio 来进行绘制。

1.2.3 任务实施

步骤 1：启动 Microsoft Office Visio，单击主菜单“文件”|“新建”|“流程图”|“基本流程图”，新建一个系统 E-R 图文件。

步骤 2：在左边的“基本流程图形状”窗格上单击“进程”（矩形）、“判定”（菱形）标签，并将其拖动到绘图页上表示系统 E-R 图的“实体”和“联系”；单击工具栏中的“绘图工具”，手动绘制椭圆形，表示 E-R 图的“属性”。还可以使用复制、粘贴的功能绘制相同的形状；选中要进行布局的形状，单击主菜单“形状”|“对齐形状”，可进行形状的布局。

步骤 3：单击工具栏中的“连接线工具”，可以在绘图页上绘制系统 E-R 图的无向边。单击工具栏中的“线条粗细”和“线型”，可对无向边进行相关设置。也可以单击工具栏中的“绘图工具”，手动绘制矩形、椭圆形和无向边等形状。

步骤 4：单击绘图页中的形状，即可在其中输入文字，单击工具栏中的“A 文本工具”，可在绘图页上输入文字，选择主菜单或快捷菜单的“格式”|“文本”/“线条”/“填充”等可进行形状的格式化。

步骤 5：“学生信息管理系统”的“班级”和“课程”两个实体的 II 级 E-R 图如图 1-7 所示，该系统的 I 级 E-R 图如图 1-8 所示。

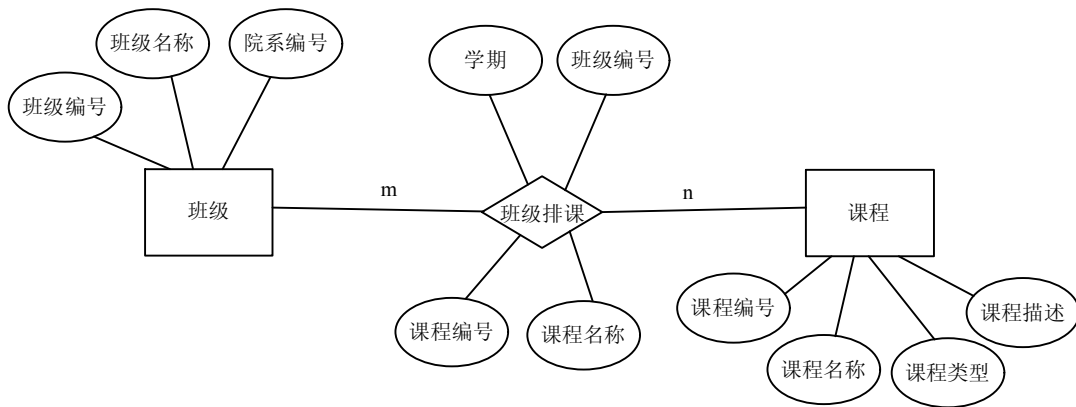


图 1-7 “班级”和“课程”两个实体的 II 级 E-R 图

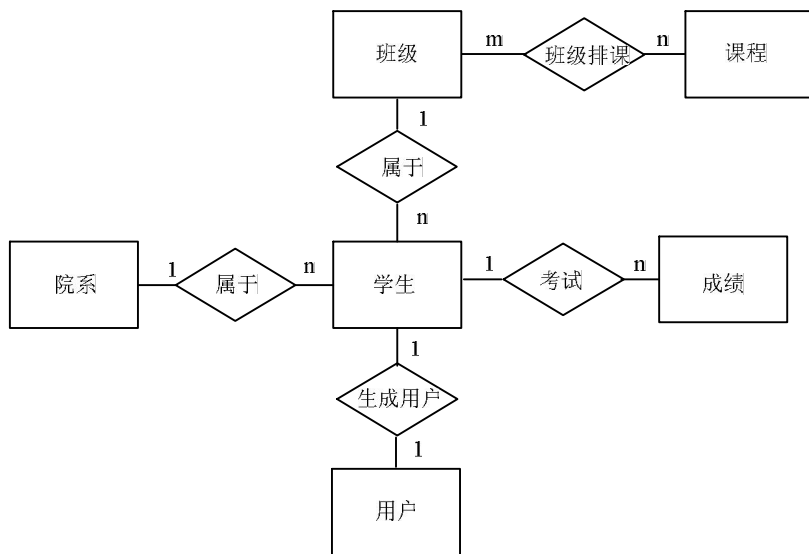


图 1-8 “学生信息管理系统”的I级 E-R 图

任务 3 关系模型设计

关系模型设计的任务就是把概念模型设计阶段设计好的 E-R 图转换为与选用数据库管理系统 (DBMS) 产品所支持的数据模型相符合的逻辑结构。关系模型是通过二维表来表示实体以及实体之间的联系的。关系模型是目前数据库系统普遍采用的数据模型,也是应用最广泛的数据模型。

将任务 2 设计的“学生信息管理系统”的概念模型转换为关系模型。

1.3.1 任务分析

掌握元组、属性、主关键字、外关键字等关系的相关术语;会将 E-R 图转换为关系模型;会对关系模型进行规范化操作;对数据完整性有清晰的认识,为以后学习实现数据完整性技术打下基础。

1.3.2 相关知识

1. 关系

关系模型是通过二维表格来表示实体以及实体之间的联系的。但并不是所有的二维表格都能称之为关系,关系必须满足以下特征:

- (1) 列是不可再分的最小数据项。
- (2) 每一列表示一个属性,且每列的名字不能相同,列的每个值的类型必须与对应属性

的类型相同。

(3) 关系中不存在完全相同的两行。

(4) 行的顺序无关紧要，列的顺序也无关紧要。

判断表 1-1、表 1-2 和表 1-3 是否是关系。

表 1-1 学生表

学号	姓名	性别	政治面貌	班级编号	班级名称	联系方式	
						家庭住址	邮编
01201401001	秦伊诺	女	团员	01201401	14 级计算机应用技术	河南省信阳市	464000
01201401002	张航	男	团员	01201401	14 级计算机应用技术	吉林省长春市	130000
01201401003	李小飞	男	党员	01201401	14 级计算机应用技术	江西省赣州市	341000
01201401004	赵文婷	男	团员	01201401	14 级计算机应用技术	福建省福州市	350000

表 1-2 用户表

用户名	密码	用户类别
admin	admin	admin
admin	admin	admin
teacher	teacher	teacher
student	student	student

表 1-3 班级排课表

学期	班级编号	课程编号	课程名称
3	01201401	001	网络数据库
3	01201401	002	网页设计与制作
2	01201401	003	Java 程序设计
1	01201401	004	高等数学

分析：表 1-1 的“联系方式”列可以再分为“家庭住址”和“邮编”两列，不满足关系的第一条特征，故表 1-1 不是关系。表 1-2 中的第一行和第二行完全相同，不满足关系的第三条特征，故表 1-2 不是关系。表 1-3 完全符合关系的所有特征，故表 1-3 是关系。

2. 关系的术语

关系中的有些术语与概念模型的术语相同或类似，但请注意它们之间的异同。

(1) 关系：满足上述特征的二维表格称之为关系，每个二维表格称为一个关系，每个关系都有一个关系名。例如：表 1-3 是一个关系，它的关系名为班级排课表。

(2) 元组：表中的一行，也叫做记录。例如：表 1-3 中的 (3, 01201401, 001, 网络数据库) 就叫做表 1-3 的一个元组，或一条记录。

(3) 属性：表中的各列，给每一个属性起一个名字，叫属性名，属性也叫做字段。例如表 1-3 中的学期，班级编号，课程编号，课程名称都是属性。

(4) 属性值：表中的各列对应的数据，也叫做字段值。例如表 1-3 中的网络数据库，就是课程名称这个属性的属性值。

(5) 候选关键字：也叫做候选键。若表中的某一属性或属性的组合能唯一地标识一个元组，则称该属性或属性的组合为候选键，候选键可以有多个。例如：若把表 1-1 的学生表规范为关系模型，则对于该表而言，学号属性能够唯一地标识学生表的一行，若在学生表中增加一个身份证号码的属性，则该属性也可以唯一地标识学生表的一行，故学号和身份证号都是学生表的候选关键字。

(6) 主关键字：在候选关键字中选择一个作为主关键字，主关键字又叫做主码或主键。在一个表中只能有一个主关键字，主关键字的值必须唯一，且不允许为空值。通常情况下，选择属性值较短的那个属性或属性的组合作为主关键字。例如：对于规范为关系的学生表，可以选用学号作为该表的主关键字；对于表 1-3 的班级排课表，可选择 (班级编号，课程编号) 这两个属性的组合作为主关键字。

(7) 外关键字：也叫做外键或外码，它是另一个表的主关键字。使用主关键字和外关键字可建立起表和表之间的联系。例如：表 1-3 的班级排课表里存放的是班级所选的课程信息，其中班级编号属性是班级信息表的主关键字，所以班级编号属性是班级排课表的外关键字；同理课程编号属性是课程表的主关键字，所以课程编号属性是班级排课表的外关键字。

(8) 主属性：能作为候选关键字的属性。例如规范为关系的学生表的学号就是主属性。

(9) 非主属性：除了主属性的其他属性。例如规范为关系的学生表的姓名、性别都是非主属性。

3. 将 E-R 图转换为关系模型

将一个实体或联系转换为表，将实体或联系的属性转换为表的列。实体或联系的主关键字转换为表的主关键字，它能唯一地标识表中的每一行。

4. 关系模型的规范化

将 E-R 图转换为关系模型后经常会出现如下问题：

数据冗余：例如学生姓名，在学生表中出现，在成绩表中也出现；课程名称在课程表中出现，在班级排课表中也出现，造成了数据的冗余。

数据可能会出现不一致的现象：如学生姓名、课程名称的重复出现，容易出现数据的不一致。再如输入课程名称不规范，有的时候输入全称，有的时候输入简称。另外在修改的时候，可能会出现遗漏的情况，造成数据的不一致。

数据维护困难：例如若某学生更改了姓名，那么就要在学生表和成绩表中都要进行修改，否则就会出现数据不一致。因此就会增加维护的工作量，造成数据的维护困难。

那么如何评价数据库关系模型设计的好坏呢？一般情况下，为了克服上述问题，都要对数据库的关系模型进行规范化，通常要求设计符合第三范式（3NF）的要求。所谓范式就是组织数据的规则，根据一个关系满足数据依赖的程度，可将范式分为 1NF、2NF、3NF、4NF、5NF，后一种范式是建立在前一种范式的基础上的，也就是说若某关系满足 3NF，那它一定满足 2NF；若某关系满足 2NF，那它一定满足 1NF。由于 4NF 和 5NF 在实际应用中意义不大，因此不予阐述。

（1）第一范式（1NF）

如果一个关系中的所有属性都不可再分，则称这个关系满足 1NF。

其实若一个二维表格能称之为关系，那么它就符合 1NF 的要求。这是因为能称之为关系的二维表格必须具备一些特征，其中有一条就是列是不可再分的最小数据项。例如表 1-1 学生表，联系方式属性可以再分成家庭住址和邮编两个属性，不符合 1NF 的要求，如何将该表规范成 1NF 呢？可以有两种方法。一种方法是将联系方式属性展开，如表 1-4 所示。

表 1-4 学生表

学号	姓名	性别	政治面貌	班级编号	班级名称	家庭住址	邮编
01201401001	秦伊诺	女	团员	01201401	14 级计算机应用技术	河南省信阳市	464000
01201401002	张航	男	团员	01201401	14 级计算机应用技术	吉林省长春市	130000
01201401003	李小飞	男	党员	01201401	14 级计算机应用技术	江西省赣州市	341000
01201401004	赵文婷	男	团员	01201401	14 级计算机应用技术	福建省福州市	350000

另一种方法是将该关系分解为两个关系，如表 1-5 和表 1-6 所示。

表 1-5 学生表

学号	姓名	性别	政治面貌	班级编号	班级名称
01201401001	秦伊诺	女	团员	01201401	14 级计算机应用技术
01201401002	张航	男	团员	01201401	14 级计算机应用技术
01201401003	李小飞	男	党员	01201401	14 级计算机应用技术
01201401004	赵文婷	男	团员	01201401	14 级计算机应用技术

（2）第二范式（2NF）

如果一个关系满足 1NF 的要求，且所有非主属性都完全函数依赖于主关键字，则称这个关系满足 2NF。

表 1-6 学生通讯录表

学号	家庭住址	邮编
01201401001	河南省信阳市	464000
01201401002	吉林省长春市	130000
01201401003	江西省赣州市	341000
01201401004	福建省福州市	350000

例如表 1-4 学生表，学号能唯一标识出该表中的每一行，所以学号是该表的主关键字。学号为“01201401001”的姓名是“秦伊诺”，学生的姓名完全能由学生的学号来决定，也就是说有一个学号就会有一个并且只能有一个姓名与它对应，则称姓名完全函数依赖于学号，也可以说学号决定了姓名。同理，表 1-4 中的性别、政治面貌、班级编号、班级名称、家庭住址和邮编等属性也完全函数依赖于学号，故表 1-4 符合 2NF 的要求。

表 1-3 班级排课表，（班级编号，课程编号）能唯一标识出该表中的每一行，所以（班级编号，课程编号）是该表的主关键字。而当给出课程编号时就可以唯一确定出该课程编号所对应的课程名称，所以课程名称属性只依赖于主关键字的课程编号，而与主关键字的班级编号无关。所以课程名称只依赖于主关键字的一部分，则说课程名称部分函数依赖于主关键字（班级编号，课程编号）。故表 1-3 不满足 2NF，如何将该表规范为 2NF 呢？可将关系进行拆分，将部分函数依赖关系中的主属性和非主属性从关系中提取出来，单独构成一个关系；将关系中余下的其他属性加上主关键字，构成一个关系。故可将表 1-3 中存在部分函数依赖关系的课程编号和课程名称提取出来单独构成课程表，而班级排课表中包含余下的其他属性和主关键字（班级编号，课程编号），如表 1-7 和表 1-8 所示。

表 1-7 课程表

课程编号	课程名称
001	网络数据库
002	网页设计与制作
003	Java 程序设计
004	高等数学

表 1-8 班级排课表

学期	班级编号	课程编号
3	01201401	001
3	01201401	002
2	01201401	003
1	01201401	004

(3) 第三范式 (3NF)

如果一个关系满足 2NF 的要求, 且关系中的任何一个非主属性都不传递函数依赖于主关键字, 则称这个关系满足 3NF。

例如表 1-4, 学号可以唯一确定一个与它对应的班级编号, 而班级编号又可以唯一确定一个与它对应的班级名称。也就是说班级名称是通过班级编号的传递而完全函数依赖于学号的, 则称学号和班级名称之间存在着传递函数依赖关系, 故表 1-4 不满足 3NF。如何将该表规范为 3NF 呢? 首先要把存在传递函数依赖关系的班级编号和班级名称提取出来单独构成一个班级表, 然后在学生表中将班级名称属性去掉, 如表 1-9 和表 1-10 所示。

表 1-9 班级表

班级编号	班级名称
01201401	14 级计算机应用技术
01201402	14 级计算机网络技术
01201403	14 级图形图像制作
01201404	14 级电子信息技术

表 1-10 学生表

学号	姓名	性别	政治面貌	班级编号	家庭住址	邮编
01201401001	秦伊诺	女	团员	01201401	河南省信阳市	464000
01201401002	张航	男	团员	01201401	吉林省长春市	130000
01201401003	李小飞	男	党员	01201401	江西省赣州市	341000

5. 数据完整性

数据完整性使得不正确、无意义和不符合规范的数据不能进入数据库, 以确保数据库中存储数据的正确性和一致性。关系模型的数据完整性包括域完整性、实体完整性和参照完整性。

(1) 域完整性, 也叫做列完整性或用户定义完整性。关系的列的值域必须满足某种特定数据类型或某种约束, 如数据类型、格式、值域范围、默认值、是否允许为空等。

例如表 1-10 学生表, 学号属性定义为长度 12 的字符型, 只能输入 12 个字符或 6 个汉字, 若输入“1111111111111” (13 个字符), 则该数据不符合对学号属性的定义, 违反了学号属性的域完整性。表 1-10 的性别属性定义为长度 2 的字符型, 只能输入 2 个字符或 1 个汉字, 并规定该列的取值只能是“男”或“女”。若输入“woman”, 则该数据不符合对性别属性的定义, 违反了性别属性的域完整性。

(2) 实体完整性, 也叫做表的完整性。指关系中必须有一个主关键字, 且主关键字不能取空值, 或者说关系中每个元组的主关键字都不能为空。

例如表 1-8 班级排课表, (班级编号, 课程编号) 是主关键字, 则它的取值不能为空, 意

味着班级编号和课程编号都不能为空，且主关键字的值要唯一，以保证该表的实体完整性。表 1-10 学生表，学号为主关键字，则它的取值必须唯一，且不能取空值，以保证该表的实体完整性。

(3) 参照完整性，也叫做引用完整性。在关系模型中，一个关系的外关键字关联着另一个关系的主关键字。则外关键字和主关键字不但数据类型及长度要定义的相同，而且对外关键字值进行插入或修改时一定要参照主关键字的值是否存在，对主关键字进行修改或删除时，也必须参照外关键字的值是否存在。

例如：表 1-7 课程表和表 1-8 班级排课表，班级排课表中的课程编号是课程表的主关键字，是班级排课表的外关键字，则对班级排课表进行插入记录或修改记录时，插入的课程编号值或修改后的课程编号值，一定要在课程表里存在。若将表 1-8 班级排课表的第一条记录的课程编号，由“001”改为“005”，则违反了这两张表的参照完整性，因为在表 1-7 中没有课程编号为“005”的记录。

如何保证数据的参照完整性，要遵循以下规则：

①对外关键字所在的表进行插入操作时，要保证插入的外关键字值在主关键字所在的表中存在。

例如若要向表 1-8 班级排课表插入数据时，插入的课程编号值一定要在表 1-7 课程表中存在。

②对外关键字所在的表进行修改操作时，要保证修改后的外关键字值在主关键字所在的表中存在。

例如若要修改表 1-8 班级排课表的数据，修改后的课程编号值一定要在表 1-7 课程表中存在。

③对主关键字所在的表进行修改操作时，首先检查要修改的主关键字的值在外关键字所在的表中是否存在，如果存在则不允许修改，或者级联修改所有与之相对应的外关键字的值。

例如要将表 1-7 课程表中的“001”课程编号修改为“005”，而表 1-8 班级排课表中存在课程编号为“001”的数据，那么此时或者不允许修改，或者允许修改，但同时要将表 1-8 中课程编号“001”修改为“005”（级联修改）。

④对主关键字所在的表进行删除操作时，首先检查要删除记录的主关键字的值在外关键字所在的表中是否存在，如果存在则不允许删除，或者级联删除所有与之相对应的外关键字的值所在的那些行。

例如要将表 1-7 课程表中课程编号为“001”的记录删除，而表 1-8 班级排课表中存在课程编号为“001”的数据，那么此时或者不允许删除，或者允许删除，但同时要将表 1-8 中课程编号为“001”的记录也删除（级联删除）。

1.3.3 任务实施

1. 将“学生信息管理系统”的 E-R 图转换为关系模型

步骤 1：根据“学生信息管理系统”的 E-R 图，将实体“院系”转换为一个关系，关系名为“院系表”，将该实体的属性“院系编号”“院系名称”“院系主任”转换为关系的列，将该

实体的主关键字“院系编号”转换为关系的主关键字。转换后的关系如表 1-11 所示。

表 1-11 院系表

院系编号	院系名称	院系主任
01	计算机科学系	张敏
02	规划设计系	程明亮
03	外语系	苏晴
04	管理系	李力

步骤 2：将实体“班级”转换为一个关系，关系名为“班级表”，将该实体的属性“班级编号”“班级名称”“院系编号”转换为关系的列，将该实体的主关键字“班级编号”转换为关系的主关键字。转换后的关系如表 1-12 所示。

表 1-12 班级表

班级编号	班级名称	院系编号
01201401	14 级计算机应用技术	01
01201402	14 级计算机网络技术	01
01201403	14 级图形图像制作	01
01201404	14 级电子信息技术	01

步骤 3：将实体“课程”转换为一个关系，关系名为“课程表”，将该实体的属性“课程编号”“课程名称”“课程类型”“课程描述”转换为关系的列，将该实体的主关键字“课程编号”转换为关系的主关键字。转换后的关系如表 1-13 所示。

表 1-13 课程表

课程编号	课程名称	课程类型	课程描述
001	网络数据库	工科	讲述 SQL Server 数据库
002	网页设计与制作	工科	基于 ASP.NET 的网页设计与制作
003	Java 程序设计	工科	基本 JAVA 的应用开发
004	高等数学	理科	高等数学是学习计算机技术的基础

步骤 4：将联系“班级排课”转换为一个关系，关系名为“班级排课表”，将该联系的属性“学期”“班级编号”“课程编号”“课程名称”转换为关系的列，将该实体的主关键字（班级编号，课程编号）转换为关系的主关键字。转换后的关系如表 1-14 所示。

表 1-14 班级排课表

学期	班级编号	课程编号	课程名称
3	01201401	001	网络数据库
3	01201401	002	网页设计与制作
2	01201401	003	Java 程序设计
1	01201401	004	高等数学

步骤 5: 将实体“学生”转换为一个关系, 关系名为“学生表”, 将该实体的属性“学号”“姓名”“性别”“出生年月”“政治面貌”“班级编号”“班级名称”“家庭住址”“邮编”“备注”转换为关系的列, 将该实体的主关键字“学号”转换为关系的主关键字。转换后的关系如表 1-15 所示。

表 1-15 学生表

学号	姓名	性别	出生年月	政治面貌	班级编号	班级名称	家庭住址	邮编	备注
01201401001	秦伊诺	女	1995-10-08	团员	01201401	14 级计算机应用技术	河南省信阳市	464000	品学兼优
01201401002	张航	男	1994-05-02	团员	01201401	14 级计算机应用技术	吉林省长春市	130000	吃苦耐劳
01201401003	李小飞	男	1994-10-01	党员	01201401	14 级计算机应用技术	江西省赣州市	341000	品学兼优
01201401004	赵文婷	男	1995-08-23	团员	01201401	14 级计算机应用技术	福建省福州市	350000	吃苦耐劳

步骤 6: 将实体“成绩”转换为一个关系, 关系名为“成绩表”, 将该实体的属性“学号”“姓名”“课程名称”“成绩”转换为关系的列, 将该实体的主关键字(学号, 课程名称)转换为关系的主关键字。转换后的关系如表 1-16 所示。

表 1-16 成绩表

学号	姓名	课程名称	成绩
01201401001	秦伊诺	网络数据库	90
01201401001	秦伊诺	JAVA 程序设计	89
01201401002	张航	网络数据库	85
01201401002	张航	JAVA 程序设计	80

步骤 7: 将实体“用户”转换为一个关系, 关系名为“用户表”, 将该实体的属性“用户名”“密码”“权限”转换为关系的列, 将该实体的主关键字“用户名”转换为关系的主关键字。

转换后的关系如表 1-17 所示。

表 1-17 用户表

用户名	密码	权限
admin	admin	admin
teacher	teacher	teacher
student	student	student
01201401001	01201401001	student

2. 对转换后的关系模型进行规范化处理

步骤 1: 将所有关系规范到 1NF。表 1-11 院系表, 表 1-12 班级表, 表 1-13 课程表, 表 1-14 班级排课表, 表 1-15 学生表, 表 1-16 成绩表, 表 1-17 用户表, 这些表的所有属性都不可再分, 故它们均符合 1NF 的要求。

步骤 2: 将所有关系规范到 2NF。表 1-11 院系表, 表 1-12 班级表, 表 1-13 课程表, 表 1-15 学生表, 表 1-17 用户表, 这些表的非主属性都完全函数依赖于主关键字, 故它们均符合 2NF 的要求。

表 1-14 的主关键字是 (班级编号, 课程编号), 而课程名称属性只依赖于课程编号, 和主关键字中的班级编号无关, 也就是说课程名称只部分函数依赖于主关键字 (班级编号, 课程编号)。故表 1-14 不符合 2NF 的要求, 需对其进行规范化处理。首先将有部分函数依赖关系的课程编号和课程名称分离出来组成一个关系, 由于分离出来的课程编号和课程名称在课程表中已存在, 所以可删除这个由刚分离属性组成的关系。在原有的表中去除课程名称属性, 规范后的关系如表 1-18 所示。

表 1-18 班级排课表

学期	班级编号	课程编号
3	01201401	001
3	01201401	002
2	01201401	003
1	01201401	004

表 1-16 成绩表的主关键字是 (学号, 课程名称), 而姓名属性只依赖于学号, 和主关键字中的课程名称无关, 也就是说姓名属性只部分函数依赖于主关键字 (学号, 课程名称)。故表 1-16 不符合 2NF 的要求, 需对其进行规范化处理。首先将有部分函数依赖关系的学号、姓名分离出来组成一个关系, 由于分离出来的学号和姓名在学生信息表中已存在, 所以可删除这个由刚分离属性组成的关系。剩余的其他属性, 即成绩, 再加上主关键字 (学号, 课程名称) 构成关系。规范后的关系如表 1-19 所示。

表 1-19 成绩表

学号	课程名称	成绩
01201401001	网络数据库	90
01201401001	Java 程序设计	89
01201401002	网络数据库	85
01201401002	Java 程序设计	80

步骤 3: 将所有关系规范到 3NF。表 1-11 院系表, 表 1-12 班级表, 表 1-13 课程表, 表 1-18 班级排课表, 表 1-19 成绩表, 表 1-17 用户表, 这些表的任何一个非主属性都不传递函数依赖于主关键字, 故它们均符合 3NF 的要求。

表 1-15 学生表的主关键字是学号, 学号可以唯一确定一个与它对应的班级编号, 而班级编号又可以唯一确定一个与它对应的班级名称。也就是说班级名称是通过班级编号的传递而完全函数依赖于学号的, 所以学号和班级名称之间存在着传递函数依赖关系。故表 1-15 不符合 3NF, 需对其进行规范化处理。首先将具有传递函数依赖关系的班级编号和班级名称分离出来组成一个关系, 由于分离出来的班级编号和班级名称在班级表中已存在, 所以可删除这个由刚分离属性组成的关系。去除表 1-15 中的班级名称属性, 规范化后的关系如表 1-20 所示。

表 1-20 学生表

学号	姓名	性别	出生年月	政治面貌	班级编号	家庭住址	邮编	备注
01201401001	秦伊诺	女	1995-10-08	团员	01201401	河南省信阳市	464000	品学兼优
01201401002	张 航	男	1994-05-02	团员	01201401	吉林省长春市	130000	吃苦耐劳
01201401003	李小飞	男	1994-10-01	党员	01201401	江西省赣州市	341000	品学兼优
01201401004	赵文婷	男	1995-08-23	团员	01201401	福建省福州市	350000	吃苦耐劳

步骤 4: 进一步规范。因为有些人习惯用课程的全称, 有些人习惯用课程的简称, 所以对表 1-19 学生成绩表中的课程名称属性进行插入、修改、删除操作时, 容易出现数据不一致的情况。而对于每一个课程编号都可以在表 1-13 课程表中找到与之对应的课程名称, 所以可将表 1-19 中的课程名称改为课程编号, 故表 1-19 的主关键字为 (学号, 课程编号), 规范化后的关系如表 1-21 所示。

表 1-21 成绩表

学号	课程编号	成绩
01201401001	001	90
01201401001	003	89
01201401002	001	85
01201401002	003	80

步骤 5: 指出规范化后的各个关系的主关键字和外关键字。

表 1-11 院系表的主关键字是院系编号, 该表无外关键字。

表 1-12 班级表的主关键字是班级编号, 该表无外关键字。

表 1-13 课程表的主关键字是课程编号, 该表无外关键字。

表 1-17 用户表的主关键字是用户名, 该表无外关键字。

表 1-18 班级排课表的主关键字是(班级编号, 课程编号), 由于班级编号是表 1-12 班级表的主关键字, 而课程编号是表 1-13 课程表的主关键字, 故班级编号和课程编号为该表的外关键字。

表 1-20 学生表的主关键字是学号, 由于班级编号是表 1-12 班级表的主关键字, 故班级编号为该表的外关键字。

表 1-21 成绩表的主关键字是(学号, 课程编号), 由于学号是表 1-20 学生表的主关键字, 课程编号是表 1-13 课程表的主关键字, 故学号和课程编号为该表的外关键字。

项目小结

本章简要介绍了与数据库相关的一些基本概念, 并围绕“学生信息管理系统”详细介绍了进行数据库设计的一般步骤。讲述了如何使用业务流程图、系统功能结构图和系统功能分析, 将需求分析的结果呈现出来; 如何根据需求分析的结果绘制 E-R 图来进行概念模型设计; 以及如何将概念模型转换为关系模型, 并对关系模型进行规范化处理。通过本章的学习, 可以使读者对数据库的设计过程有一个清晰的认识。

实训项目

综合实训: 对简化的“图书管理系统”进行数据库设计

实训目标: 掌握业务流程图和系统功能结构图的画法, 会进行系统功能分析; 掌握 E-R 图的画法, 会进行概要模型设计; 掌握将概要模型转换为关系模型的方法, 会对转换后的关系进行规范化处理。

实训要求: 对简化的“图书管理系统”进行数据库设计。

实训操作步骤:

步骤 1: 绘制“图书管理系统”的业务流程图。

启动 Microsoft Office Visio, 单击主菜单“文件”|“新建”|“流程图”|“基本流程图”, 新建一个业务流程图文件。在左边的“基本流程图形状”窗格上单击“进程”(矩形)标签, 并将其拖动到绘图页上表示业务流程图的“活动”; 单击工具栏中的“连接线工具”, 可以在绘图页上绘制业务流程图的“逻辑关系线”, 单击工具栏中的“线条粗细”“线型”和“线端”, 对逻辑关系线进行相关设置; 单击绘图页中的形状, 即可在其中输入文字, 单击工具栏中的“A

文本工具”，可在形状上输入文字。可以使用复制、粘贴的功能绘制相同的形状，也可以单击工具栏中的“绘图工具”，手动绘制矩形、圆形、椭圆形、直线、曲线等形状。选中要进行布局的形状，单击主菜单“形状”|“对齐形状”，可进行形状的布局；选择主菜单或快捷菜单的“格式”|“文本”/“线条”/“填充”等可进行形状的格式化。最终绘制好的“图书管理系统”的业务流程图如图 1-9 所示。

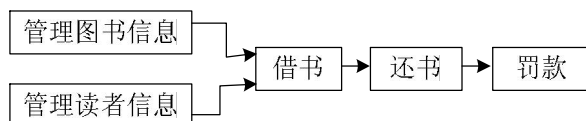


图 1-9 “图书管理系统”业务流程图

步骤 2：绘制“图书管理系统”的系统功能结构图。

启动 Microsoft Office Visio，单击主菜单“文件”|“新建”|“流程图”|“基本流程图”，新建一个系统功能图文件。在左边的“基本流程图形状”窗格上单击“进程”（矩形）标签，并将其拖动到绘图页上表示系统功能结构图的“功能”；单击工具栏中的“连接线工具”，可以在绘图页上绘制箭头，表示系统功能结构图中各功能之间的从属关系。单击工具栏中的“线条粗细”“线型”和“线端”，对箭头线进行相关设置；单击绘图页中的形状，即可在其中输入文字，单击工具栏中的“A 文本工具”，可在形状上输入文字。还可以使用复制、粘贴的功能绘制相同的形状，也可以单击工具栏中的“绘图工具”，手动绘制矩形、圆形、椭圆形、直线、曲线等形状。选中要进行布局的形状，单击主菜单“形状”|“对齐形状”，可进行形状的布局；选择主菜单或快捷菜单的“格式”|“文本”/“线条”/“填充”等可进行形状的格式化。最终绘制好的“图书管理系统”的系统功能结构图如图 1-10 所示。

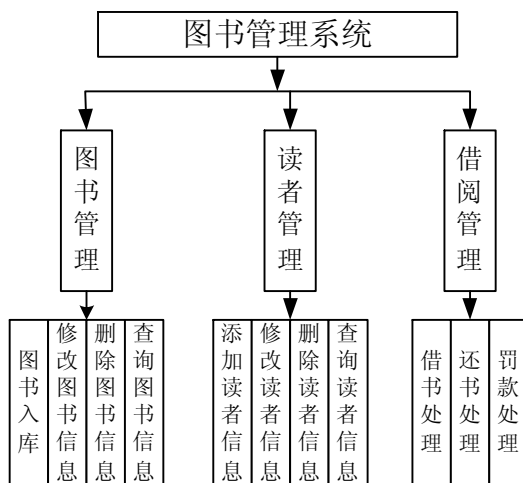


图 1-10 “图书管理系统”的系统功能结构图

步骤 3: 进行“图书管理系统”的系统功能分析。从图 1-10 可以得出系统的功能需求如下:

图书管理功能: 用于添加、修改、删除和查询图书信息。

读者管理功能: 用于添加、修改、删除和查询读者信息。

借阅管理功能: 用于对借书、还书和罚款进行处理。

本系统是简化的“图书管理系统”, 故并未考虑该系统的操作员管理和权限管理。读者可以在该简化系统的基础上, 进行系统扩展。

步骤 4: 绘制 E-R 图, 进行“图书管理系统”的概要模型设计。

启动 Microsoft Office Visio, 单击主菜单“文件”|“新建”|“流程图”|“基本流程图”, 新建一个系统 E-R 图文件。在左边的“基本流程图形状”窗格上单击“进程”(矩形)、“判定”(菱形)标签, 并将其拖动到绘图页上表示系统 E-R 图的“实体”和“联系”; 单击工具栏中的“绘图工具”, 手动绘制椭圆形, 表示 E-R 图的“属性”; 单击工具栏中的“连接线工具”, 可以在绘图页上绘制系统 E-R 图的无向边, 单击工具栏中的“线条粗细”和“线型”可对无向边进行相关设置; 单击绘图页中的形状, 即可在其中输入文字, 单击工具栏中的“A 文本工具”, 可在绘图页上输入文字。还可以使用复制、粘贴的功能绘制相同的形状, 也可以单击工具栏中的“绘图工具”, 手动绘制矩形、椭圆形和无向边等形状。选中要进行布局的形状, 单击主菜单“形状”|“对齐形状”, 可进行形状的布局; 选择主菜单或快捷菜单的“格式”|“文本”/“线条”/“填充”等可进行形状的格式化。“图书管理系统”的完整 E-R 图如图 1-11 所示。

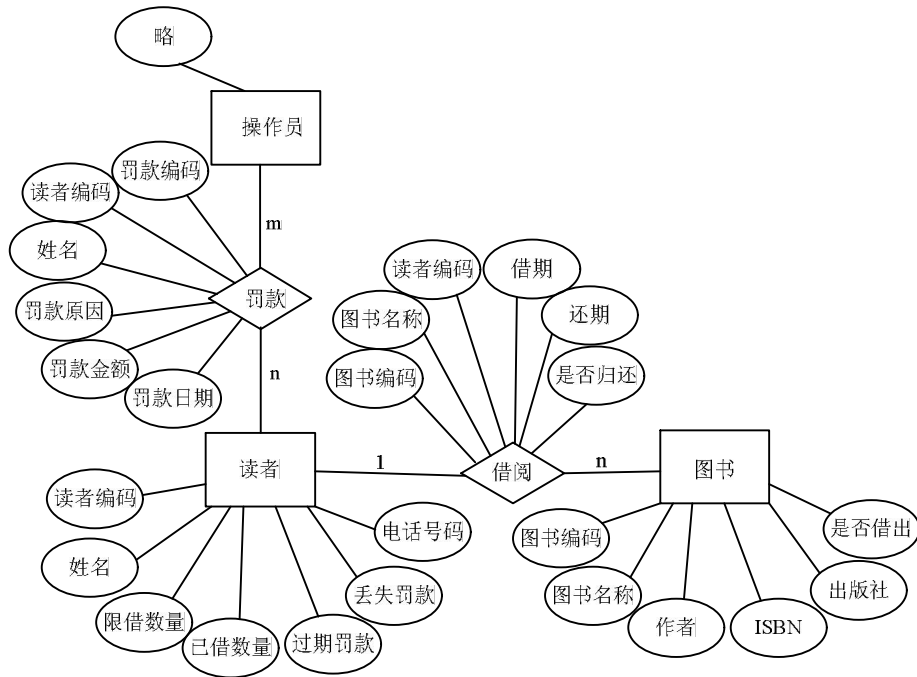


图 1-11 “图书管理系统”的完整 E-R 图

步骤 5: 将绘制好的 E-R 图转换成“图书管理系统”的关系模型。

根据“图书管理系统”的 E-R 图, 将实体“图书”转换为一个关系, 关系名为“图书信息表”。将该实体的属性“图书编码”“图书名称”“作者”“ISBN”“出版社”“是否借出”转换为关系的列, 将该实体的主关键字“图书编码”转换为关系的主关键字。转换后的关系如表 1-22 所示。

表 1-22 图书信息表

图书编码	图书名称	作者	ISBN	出版社	是否借出
00000001	Oracle 数据库应用教程	朱亚兴	9787560619781	西安电子科技大学出版社	是
00000002	Web 应用程序开发教程	唐笑林	9787040269710	高等教育出版社	是
00000003	SQL Server 2005 数据库及应用	徐人凤	9787040372052	高等教育出版社	是
00000004	Visual C#.NET 面向对象程序设计教程	赵卫伟	978111187646	机械工业出版社	是

将实体“读者”转换成一个关系, 关系名为“读者信息表”。将该实体的属性“读者编码”“姓名”“限借数量”“已借数量”“限借天数”“逾期罚款”“丢失罚款”“电话号码”转换成关系的列, 将该实体的主关键字“读者编码”转换为关系的主关键字。转换后的关系如表 1-23 所示。

表 1-23 读者信息表

读者编码	姓名	限借数量	已借数量	限借天数	逾期罚款	丢失罚款	电话号码
00001	秦伊诺	10	2	7	每天 2 角	2 倍书价	15565589908
00002	张航	10	0	7	每天 2 角	2 倍书价	15837653470
00003	李小飞	10	1	7	每天 2 角	2 倍书价	15637616784
00004	赵文婷	10	1	7	每天 2 角	2 倍书价	13403768907

将联系“借阅”转换成一个关系, 关系名为“借阅信息表”。将该联系的属性“图书编码”“图书名称”“读者编码”“借期”“还期”“是否归还”转换成关系的列, 将该联系的主关键字(图书编码, 读者编码)转换为关系的主关键字。转换后的关系如表 1-24 所示。

将联系“罚款”转换成一个关系, 关系名为“罚款信息表”。将该联系的属性“罚款编码”“读者编码”“姓名”“罚款原因”“罚款金额”“罚款日期”转换成关系的列, 将该联系的主关键字“罚款编码”转换为关系的主关键字。转换后的关系如表 1-25 所示, 其中的罚款金额的单位是元。

表 1-24 借阅信息表

图书编码	图书名称	读者编码	借期	还期	是否归还
00000001	Oracle 数据库应用教程	00001	2014-04-15	2014-04-22	是
00000002	Web 应用程序开发教程	00003	2014-04-16	2014-04-23	是
00000003	SQL Server2005 数据库及应用	00001	2014-04-16	2014-04-23	是
00000004	Visual C#.NET 面向对象程序设计教程	00004	2014-04-17	2014-04-23	否

表 1-25 罚款信息表

罚款编码	读者编码	姓名	罚款原因	罚款金额	罚款日期
20140424001	00001	秦伊诺	逾期 2 天还书	0.4	2014-04-24
20140428001	00004	赵文婷	丢失图书	40	2014-04-28

步骤 6: 对转换后的关系模型进行规范化处理。

首先将所有关系规范到 1NF。表 1-22 图书信息表, 表 1-23 读者信息表, 表 1-24 借阅信息表, 表 1-25 罚款信息表, 这些表的所有属性都不可再分, 故它们均符合 1NF 的要求。

其次将所有关系规范到 2NF。表 1-22 图书信息表, 表 1-23 读者信息表, 表 1-25 罚款信息表, 这些表的非主属性都完全函数依赖于主关键字, 故它们均符合 2NF 的要求。

表 1-24 借阅信息表的主关键字是 (图书编码, 读者编码), 而图书名称属性只依赖于图书编码, 和主关键字中的读者编码无关, 也就是说图书名称属性只部分函数依赖于主关键字 (图书编码, 读者编码)。故表 1-24 不符合 2NF 的要求, 需对其进行规范化处理。首先将有部分函数依赖关系的图书编码和图书名称分离出来组成一个关系, 由于分离出来的图书编码和图书名称在图书信息表中已存在, 所以可删除这个由刚分离属性组成的关系。在原有的表中去除图书名称属性, 规范后的关系如表 1-26 所示。

最后将所有关系规范到 3NF。表 1-22 图书信息表, 表 1-23 读者信息表, 表 1-26 借阅信息表, 这些表的任何一个非主属性都不传递函数依赖于主关键字, 故它们均符合 3NF 的要求。

表 1-26 借阅信息表

图书编码	读者编码	借期	还期	是否归还
00000001	00001	2014-04-15	2014-04-22	是
00000002	00003	2014-04-16	2014-04-23	是
00000003	00001	2014-04-16	2014-04-23	是
00000004	00004	2014-04-17	2014-04-23	否

表 1-25 罚款信息表的主关键字是罚款编码, 罚款编码可以唯一确定一个与它对应的读者编码, 而读者编码又可以唯一确定一个与它对应的姓名。也就是说姓名是通过读者编码的传递

而完全函数依赖于罚款编码的，所以罚款编码和姓名之间存在着传递函数依赖关系。故表 1-25 不符合 3NF，需对其进行规范化处理。首先将具有传递函数依赖关系的读者编码和姓名分离出来组成一个关系，由于分离出来的读者编码和姓名在读者信息表中已存在，所以可删除这个由刚分离属性组成的关系。去除表 1-25 中的姓名属性，规范化后的关系如表 1-27 所示。

表 1-27 罚款信息表

罚款编码	读者编码	罚款原因	罚款金额	罚款日期
20140424001	00001	逾期 2 天还书	0.4	2014-04-24
20140428001	00004	丢失图书	40	2014-04-28

步骤 7：指出规范化后的各个关系的主关键字和外关键字。

表 1-22 图书信息表的主关键字是图书编码，该表无外关键字。

表 1-23 读者信息表的主关键字是读者编码，该表无外关键字。

表 1-26 借阅信息表的主关键字是（图书编码，读者编码），由于图书编码是表 1-22 图书信息表的主关键字，读者编码是表 1-23 读者信息表的主关键字，所以图书编码和读者编码是该表的外关键字。

表 1-27 罚款信息表的主关键字是罚款编码，由于读者编码是表 1-23 读者信息表的主关键字，所以读者编码是该表的外关键字。