

1

数据库基本概念

数据库是企业、组织或部门所涉及的存储在一起的相关数据的集合，它反映了数据本身的内容及数据之间的联系。

Visual FoxPro 是目前优秀的数据库管理系统之一。掌握数据库及数据库管理系统的基本概念，有助于在 Visual FoxPro 的可视化环境下，使用面向对象的方法开发出功能良好的数据库和应用程序。本章主要介绍数据库、数据库管理系统、关系及关系数据库的基本概念、关系数据库设计的基本知识。

1.1 数据模型

说到模型我们并不陌生，例如，一张地图、一辆汽车模型都是具体的模型。模型是现实世界特征的模拟和抽象。数据模型也是一种模型，它是现实世界数据特征的抽象。

1.1.1 现实世界的描述

数据库是某个实际问题中涉及的数据的综合，它不仅要反映数据本身的内容，而且要反映数据之间的联系。由于计算机不能直接描述现实世界中的具体事物，所以人们必须事先把具体事物转换成计算机能够处理的数据。这个过程经历了从对现实生活中事物特性的认识、概念化到计算机数据库里的具体表示的逐级抽象过程。

1. 实体的描述

现实世界中存在各种事物，事物之间存在着联系，这种联系是客观存在的，是由事物本身的性质所决定的。例如，图书馆中有图书和读者，读者借阅图书；学校的教学系统中有教师、学生和课程，教师为学生授课，学生选修课程并取得成绩。

(1) 实体。

实体是指客观存在并且相互区别的事物。例如，某个教师、某个学生、某一本图书都是

实体。实体也可以是抽象的概念或联系，如学生的一次选课。

(2) 实体的属性。

实体的属性是指描述实体的特性，即实体是通过属性来描述的。比如：学生实体的属性有学号、姓名、性别等。属性由属性名、类型和属性值组成。比如，“姓名”是属性名，类型为字符型，对于某个具体的学生而言，其属性值为“刘明”。

(3) 实体型。

属性的集合表示一种实体的类型，称为实体型。例如，图书实体的实体型表示为（书号，书名，作者，单价）；职工实体的实体型表示为（职工号，姓名，性别，出生日期，职称）。

(4) 实体集。

同一类型的实体的集合，称为实体集。例如，某单位所有职工按照职工实体型的描述得到的数据构成职工实体集。

在 Visual FoxPro 中，用“表”来存放同一类实体组成的实体集，如图 1-1 所示的学生档案表。一个“表”中包含的若干个“字段”即为实体的属性，如表中的学号、姓名、性别等均为字段；字段值的集合组成表中的一条记录，代表一个具体的实体，如表中的一行（08010402001，李刚，男，03/12/90，F，浙江杭州，01，0104）即为一条学生记录。

学号	姓名	性别	出生年月	团员否	籍贯	院系代码	专业代码
08010401001	李红丽	女	08/06/91	F	黑龙江哈尔滨	01	0102
08010401002	王晓刚	男	06/05/90	F	河北石家庄	01	0104
08010402001	李刚	男	03/12/90	F	浙江杭州	01	0104
08010201001	王心玲	女	05/06/89	T	河南南阳	01	0102
08010201002	李力	男	03/05/91	T	河南平顶山	01	0102
08080301001	王小红	男	09/02/90	T	陕西西安	08	0803
08080301002	杨晶晶	男	09/04/90	T	山东枣庄	08	0803
08080302001	王虹虹	女	05/09/91	F	湖北黄石	08	0803
08010202002	王明	男	12/09/91	F	广西柳州	01	0102
08010202003	张建明	男	11/12/91	T	河北石家庄	01	0102
08010202004	李红玲	女	09/08/90	T	山东曲阜	01	0102
08010102001	李国志	男	02/28/89	F	山西运城	01	0101
08010102002	赵明超	男	06/01/92	F	湖北武汉	01	0101
08010102003	赵艳玲	女	07/03/91	T	湖北武汉	01	0101

图 1-1 学生档案表

2. 实体间联系及联系的种类

现实世界中，事物内部以及事物之间是有联系的，这些联系在信息世界中反映为不同类型的实体之间的联系。例如，一名教师可以同时教授多个学生，每个学生也可以有多个老师。

实体间的联系共分三种类型：

(1) 一对一联系 (one-to-one relationship)。

一对一联系是双向的一对一。如果有两个实体集 A 和 B，A 中的每个实体只与 B 中的一个实体相关联，而 B 中的每个实体也只与 A 中的一个实体相关联，我们称 A 和 B 是一一对一的联系。例如，班级和班长之间就是一对一的联系。

(2) 一对多联系 (one-to-many relationship)。

如果有两个实体集 A 和 B, A 中的每个实体与 B 中的多个实体相关联, 而 B 中的每个实体至多与 A 中的一个实体相关联, 我们称 A 和 B 是一对多的联系, 而 B 和 A 则是多对一的联系。一对多联系是最普遍的联系。例如, 部门和职工这两个实体集之间存在一对多的联系, 班级和学生之间也是一对多的联系。

(3) 多对多联系 (many-to-many relationship)。

如果有两个实体集 A 和 B, A 中的每个实体与 B 中的多个实体相关联, 而 B 中的每个实体也与 A 中的多个实体相关联, 我们称 A 和 B 是多对多的联系。例如, 教师和学生之间, 供应商和商品之间都是多对多的联系。

实际上, 一对一联系为一对多联系的特例, 而一对多联系是多对多联系的特例。

1.1.2 数据模型

数据模型是数据库管理系统用来表示实体及实体间联系的方法, 是数据库设计的核心与基础。一个具体的数据模型应当正确地反映出数据之间存在的整体逻辑关系。

数据库不仅管理数据本身, 还要使用数据模型表示出数据之间的联系。任何一个数据库管理系统都是基于某种数据模型的。数据库管理系统支持三种数据模型: 层次模型、网状模型和关系模型。目前最流行的数据模型是关系模型。

1. 层次模型

用树型结构表示实体及其之间联系的模型称为层次模型。层次模型实际上是由若干个代表实体之间一对多联系的基本层次联系组成的一棵树, 树的每个结点代表一个实体类型。如图 1-2 所示为一个学校组织结构的层次模型。

2. 网状模型

用网状结构表示实体及其之间联系的模型称为网状模型。网中的每一个结点代表一个实体类型。网状模型允许一个以上的结点无双亲, 或一个结点可以有多个的双亲。如图 1-3 所示为一个学校教学实体的网状模型。

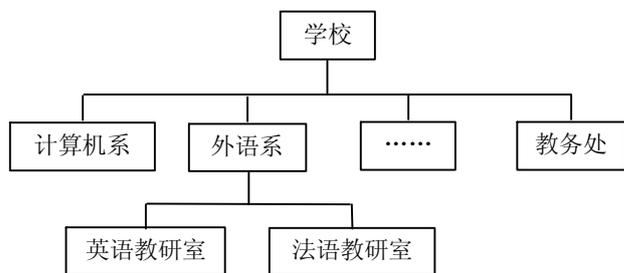


图 1-2 层次模型

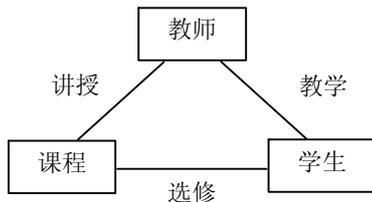


图 1-3 网状模型

3. 关系模型

用二维表结构表示实体及其之间联系的模型称为关系模型,如图 1-1 所示。在关系模型中,操作的对象和结果都是关系,每个关系都是一个二维表,无论实体本身还是实体之间的联系均用二维表来表示,使得描述实体的数据本身能够自然地反映它们之间的联系。

支持关系模型的数据库管理系统称为关系数据库管理系统,在这种系统中建立的数据库是关系数据库。关系数据库以其完备的理论基础、简单的模型、说明性的查询语言和使用方便等优点得到广泛的应用。

1.2 数据库系统

1.2.1 有关数据库的基本概念

1. 数据库 (DataBase, DB)

数据库是存储在计算机存储设备上,结构化的相关数据集合。它不仅包括描述事物的数据本身,而且还包括相关事物之间的联系。数据库中存放的数据往往面向多种应用,可以被多个用户、多个应用程序所共享。

2. 数据库管理系统 (DataBase Management System, DBMS)

数据库管理系统是负责对数据库的建立、使用和维护进行管理的大型系统软件,是数据库系统的核心组成部分。它建立在操作系统的基础上,是位于操作系统与用户之间的一层数据管理软件,负责对数据库进行统一的管理和控制。

Visual FoxPro 就是一个功能完善的 DBMS,能够实现数据库管理的各项功能,而且使用方便,适合中小型企业开发各种数据库应用系统。

1.2.2 数据库系统

1. 数据库系统 (DataBase System, DBS)

数据库系统是指引进数据库技术后的计算机系统,实现有组织地、动态地存储大量相关数据,是提供数据处理和信息资源共享的便利手段。

2. 数据库系统的特点

(1) 实现数据共享,减少数据冗余。

建立数据库时,应以面向全局的观点组织数据库中的数据,而不应像文件系统那样只考虑某一部门的局部应用,这样才能发挥数据共享的优势。

(2) 采用特定的数据模型。

数据库中的数据是有结构的,这种结构由数据库管理系统所支持的数据模型表现出来。所以根据 DBMS 的数据模型不同,数据库系统分为层次数据库系统、网状数据库系统和关系数据库系统。

(3) 具有较高的数据独立性。

在数据库系统中，数据库管理系统提供映像功能，实现了应用程序对数据的总体逻辑结构、物理存储结构之间较高的独立性。用户只需以简单的逻辑结构来操作数据，无须考虑数据存储的位置与结构。

(4) 有统一的数据控制功能。

数据库可以被多个用户或应用程序共享，数据的存取往往是并发的，即多个用户同时使用同一个数据库。数据库管理系统必须提供必要的保护措施，包括并发控制、数据的安全性控制和数据的完整性控制功能。

1.3 关系数据库

1.3.1 关系数据库

1. 关系术语

(1) 关系。

一个关系就是一张二维表。在 Visual FoxPro 中，一个关系存储为一个表文件，文件扩展名为 .dbf。一个数据库可分解成多个表。例如，学生成绩管理数据库由学生档案表、课程表及成绩表等构成。如图 1-4 所示为一个关系：“学生档案”表。

学号	姓名	性别	出生年月	团员否	籍贯	院系代码	专业代码
08010401001	李红丽	女	08/06/91	F	黑龙江哈尔滨	01	0102
08010401002	王晓刚	男	06/05/90	F	河北石家庄	01	0104
08010402001	李刚	男	03/12/90	F	浙江杭州	01	0104
08010201001	王心玲	女	05/06/89	T	河南南阳	01	0102
08010201002	李力	男	03/05/91	T	河南平顶山	01	0102
08080301001	王小虹	男	09/02/90	T	陕西西安	08	0803
08080301002	杨晶晶	男	09/04/90	T	山东枣庄	08	0803
08080302001	王虹虹	女	05/09/91	F	湖北黄石	08	0803
08010202002	王明	男	12/09/91	F	广西柳州	01	0102
08010202003	张建明	男	11/12/91	T	河北石家庄	01	0102
08010202004	李红玲	女	09/08/90	T	山东曲阜	01	0102
08010102001	李国志	男	02/28/89	F	山西运城	01	0101
08010102002	赵明超	男	05/01/92	F	湖北武汉	01	0101
08010102003	赵艳玲	女	07/03/91	T	湖北武汉	01	0101

图 1-4 “学生档案”表

对关系的描述称为关系模式。一个关系模式对应一个关系的结构，其格式为：关系名（属性名 1，属性名 2，…，属性名 n）。在 Visual FoxPro 中，关系模式表示为表的结构，其格式为：表名（字段名 1，字段名 2，…，字段名 n）。

(2) 元组（记录）。

在一个二维表中，每一行称为一个元组，对应现实中的一个实体。在 Visual FoxPro 中，

一个元组称为一条记录。图 1-4 所示的“学生档案”表中包含了 20 条记录。

(3) 属性(字段)。

二维表中的每一列称为属性，每一列有一个属性名。在 Visual FoxPro 中，属性又称为字段。如图 1-4 所示的“学生档案”表中包含了学号、姓名、性别等字段。

(4) 域。

域是指属性的取值范围。例如，成绩的取值范围是 0 到 100 之间的数值。

(5) 主关键字。

主关键字是关系中属性或属性的组合，其值能够唯一地标识一个元组。在 Visual FoxPro 中表示为字段或字段的组合，其值能够唯一地标识一条记录。例如，“学生档案”表中的“学号”字段可以作为主关键字，而“性别”字段就不行。

(6) 外部关键字。

如果表中的一个字段不是本表的主关键字或候选关键字，而是另外一个表的主关键字或候选关键字，这个字段就称为本表的外部关键字。外部关键字是用来联系两张有一定关系的表的字段，所以，外部关键字一定同时包含在两张表中。

2. 关系模型

通过以上介绍，可以将关系定义为元组的集合。关系模式是命名的属性集合，元组是属性值的集合。一个具体的关系模型是若干个有联系的关系模式的集合。在 Visual FoxPro 中，把相互之间存在联系的表放到一个数据库中统一管理，数据库文件的扩展名为 .dbc。

3. 关系模型的特点

(1) 关系中的每一项是最基本的数据项，不可再分。

(2) 在同一个关系中，不允许出现相同的属性名。

(3) 关系中不允许有完全相同的记录。

(4) 在一个关系中，记录的次序无关紧要，即任意交换两行的位置并不影响数据的实际含义。

(5) 在一个关系中，列的次序无关紧要，即任意交换两列的位置并不影响数据的实际含义。

4. 关系数据库

关系数据库是按照关系模型设计的若干关系的集合。一个关系就是一个二维表，它对应计算机中的一个数据表文件，该表文件由表文件名唯一标识，计算机通过表文件名访问该表。一个数据表由若干条记录组成，而每条记录则由若干个字段值组成。

关系数据库不仅包含若干个表，还包含表之间的关联关系。

1.3.2 表间的关联关系

在一个关系数据库中往往包含若干个表，每个表对应现实世界中的一类实体或实体之间

的联系。确定联系的目的是使表的结构合理, 在同一个数据库中, 表间的关联关系有一对一、一对多和多对多三种。

1. 一对一联系

对于一对一联系的表, 首先考虑是否可以把两个表的字段合并到一个表中。如果不能合并的话, 再观察两个表是否有同样的实体, 可以在两个表中使用同样的主关键字字段。如果两个表有不同的实体及不同的主关键字, 选择其中一个表, 把它的主关键字字段放到另一个表中作为外部关键字字段。

2. 一对多联系

一对多联系是关系数据库中最普遍的联系。要建立一对多的联系, 可以把“一方”的主关键字字段添加到“多方”的表中, 使二者具有公共字段, 然后在“多方”的表中为该字段建立普通索引, 该字段即为多方的外部关键字。

3. 多对多联系

建立两个表之间的多对多联系, 有效的方法是创建第三个表, 即纽带表。把多对多的联系分解为两个一对多的联系。创建的第三个表中包含多对多联系涉及的两个表的主关键字。纽带表不一定需要有自己的主关键字, 如果需要, 应当将它所联系的两个表的主关键字作为组合关键字并指定为主关键字。

1.3.3 关系运算

在 Visual FoxPro 中, 查询是高度非过程化的, 即用户只需明确提出“要做什么”, 而不需要说明“如何做”。然而, 要正确表示复杂的查询并非是一件简单的事。了解专门的关系运算有助于正确给出查询表达式。

(1) 选择。

选择是指从关系中找到满足条件的元组的操作。选择的条件以逻辑表达式形式给出, 满足条件的元组即被选取。例如, 从“学生档案”表中找出籍贯为“湖北”的学生, 或者是从“学生档案”表中查询所有男生的记录, 这类查询操作即为选择运算。

对于关系而言, 选择是从行的角度进行的运算, 即从水平方向抽取记录。选择运算的结果可以形成一个新的关系, 关系模式不变, 但其中的元组为原关系的一个子集。

(2) 投影。

投影是指从关系模式中指定若干个属性组成新的关系的操作。例如, 从“学生档案”表中查询学生的学号、姓名和籍贯, 这类查询操作即为投影运算。

对于关系而言, 投影是从列的角度进行的运算, 相当于对关系进行垂直方向的分解。经过投影运算可以得到一个新关系, 其关系模式所包含的属性个数往往比原关系少。

(3) 连接。

连接是指将两个关系模式拼接成一个更宽的关系模式, 生成的新关系中包含满足连接条件的元组。连接是关系的横向结合。

连接过程是通过连接条件控制的，连接条件中需给出两个关系中的公共属性，或是具有相同语义、可比的属性。连接运算的结果是满足条件的所有记录，相当于 Visual FoxPro 中的“内部连接”。

选择和投影运算的操作对象只涉及一个表，连接运算需要两个表作为操作对象。

例如，假设有职工（职工号，姓名，性别，婚否，政治面貌，工作日期，职称）和工资（职工号，姓名，基本工资，奖金，津贴）两个表，要查询基本工资高于 500 元的职工姓名、性别、职称、基本工资、奖金。

由于要查询的字段分别在两个表中，需要把这两个表连接起来，而连接的条件就是两个表的职工号对应相等，并且基本工资高于 500 元；然后再对连接的结果按照所需要的属性进行投影。

（4）等值连接和自然连接。

等值连接是指在连接运算中，按照字段值相等为条件进行的连接操作。

自然连接是指去掉重复属性的等值连接。自然连接是最常用的连接运算。

总之，在对关系数据库的查询中，利用关系的投影、选择和连接运算可以方便地分解或构造新的关系。

1.3.4 关系完整性

关系完整性规则是对关系的某种约束条件，用于保证数据的正确性、有效性和相容性。关系模型中有三种完整性约束：实体完整性、域完整性和参照完整性。

1. 实体完整性

实体完整性是指关系中的某个主关键字值不能为空，也不能具有相同值。如果主关键字值为空，则意味着存在不可识别的实体；如果主关键字的值不唯一，则失去了唯一标识记录的作用。

例如，在“学生档案”表中，“学号”字段值必须是唯一且非空的，它是区别不同学生的唯一标识，而且每个学生必须有自己的学号。

2. 域完整性

域完整性是对数据表中字段属性的约束，它包括对字段的值域、字段的类型及字段的有效性规则等的约束，它是由确定关系结构时所定义的字段属性决定的。

例如，学生成绩应大于或等于零，职工的工龄应小于年龄等。

3. 参照完整性

参照完整性是指对外部关键字的参照引用，具体说就是指关系中外部关键字必须是另一个关系的主关键字的有效值或空值。

例如，学生数据库中包含“学生档案”表和“学生成绩表”，如果“学生成绩表”中某个学生的学号在“学生档案”表中并不存在，这两个表就不满足参照完整性规则。

习题 1

一、单项选择题

1. 在关系运算中, 查找满足一定条件的元组, 相关的运算称为 ()。
A) 选择 B) 投影 C) 连接 D) 扫描
2. Visual FoxPro 是关系数据库管理系统, 所谓关系是指 ()。
A) 二维表中各记录的数据彼此有一定的关系
B) 二维表中各字段彼此有一定的关系
C) 一个表与另一个表之间有一定的关系
D) 数据模型符合并满足一定条件的二维表格
3. 数据库系统的核心是 ()。
A) 数据库 B) DBMS
C) 操作系统 D) 文件
4. DBMS 是 ()。
A) 操作系统的一部分 B) 操作系统支持下的系统软件
C) 一种编译程序 D) 一种操作系统
5. 在 Visual FoxPro 中建立数据库表时, 将年龄字段限制在 12~30 岁之间的操作属于 ()。
A) 实体完整性约束 B) 域完整性约束
C) 参照完整性约束 D) 视图完整性约束
6. 关系数据库管理系统所管理的关系是 ()。
A) 一个 .dbf 文件 B) 若干个二维表
C) 一个 .dbc 文件 D) 若干个 .dbc 文件
7. 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) 三者没有关系
8. 关系数据库的任何检索操作都是由三种基本运算组合而成的, 这三种基本运算不包括 ()。
A) 连接 B) 比较 C) 选择 D) 投影
9. 为合理组织数据, 设计数据库应遵守的原则是 ()。
A) “一事一地”的原则, 即一个表描述一个实体或实体间的一种联系
B) 表中的字段必须是原始数据和基本的数据元素, 避免在表之间出现重复字段
C) 用外部关键字保证有关联的表之间的联系
D) 以上各条原则都包括

10. () 是长期存储在计算机内有组织的、可共享的数据集合。
A) DATA B) DBS C) DB D) INFORMATION
11. 下列叙述中正确的是 ()。
A) 数据库系统是一个独立的系统, 不需要操作系统的支持
B) 数据库技术的根本目标是要解决数据的共享问题
C) 数据库管理系统就是数据库系统
D) 以上说法都不对
12. 下列叙述中正确的是 ()。
A) 为了建立一个关系, 首先要构造数据的逻辑关系
B) 表示关系的二维表中各元组的每一个分量还可以分成若干数据项
C) 一个关系的属性名表称为关系模式
D) 一个关系可以包括多个二维表

二、填空题

- 在关系数据模型中, 二维表的列称为属性, 行称为_____。
- 数据模型不仅表示反映事物本身的数据, 而且表示_____。
- 在关系数据库中, 表间关联关系的类型有_____、_____和_____。
- 自然连接是指_____。
- Visual FoxPro 不允许在主关键字字段中有重复值或_____。
- 设有学生和班级两个实体, 每个学生只能属于一个班级, 一个班级可以有 multiple 学生, 则学生和班级实体之间的联系类型是_____。
- 在数据库技术中, 实体集之间的联系可以是一对一、一对多或多对多的, 那么“学生”和“可选课程”的联系为_____。
- 人员基本信息一般包括身份证号、姓名、性别、年龄等。其中可以作为主关键字的是_____。
- 在关系数据库中, 用来表示实体之间联系的是_____。
- 在数据库管理系统提供的数据定义语言、数据操纵语言和数据控制语言中, _____负责数据的模式定义与数据的物理存取构建。

三、简答题

- 数据与信息有何区别和联系? 数据处理经历了哪几个阶段?
- 数据表之间的关联关系有几种? 试举例说明。
- 数据库、数据库管理系统和数据库系统之间是什么关系?
- 关系完整性包括哪几个方面?