

第一部分

dBASE III /FoxBASE+题解

第一章 数据库基本概念

[1.1] 试区别下列名词:

1. 数据与数据模型
2. 数据模型与数据模式
3. 概念数据模型与概念数据模式

[解]

1. 数据是描述事物的符号记录。例如,对一个大学生,可以采用如下的记录来描述:(李丽,女,1970,北京,计算机系,1988)。可以从中得知,李丽是女大学生,1970年出生,北京人,1988年考入计算机系。这条记录就是描述一个大学生的数据,人们通过解释以至推论、综合,从数据所获得的有意义的内容即为信息。因此,数据是信息存在的一种形式。而数据模型是数据特征的抽象,它不是描述个别的数据,而是描述数据的共性。它一般包括两个方面:一是数据的静态特性,包括数据的结构和限制;二是数据的动态特性,即在数据上所定义的运算或操作。例如,关系数据模型定义了关系的结构以及存在完整性限制、引用完整性限制等静态特性,同时又定义了选择、投影、连接和关系集合运算以及更新、定义数据等操作。这些定义与具体的数据和实现的方式无关,它们是关系型数据的抽象。

2. 数据模型是数据特征的抽象,它描述数据的基本结构及其相互间的关系以及定义在数据上的操作。在数据模型中有“型”与“值”之分,型是某一类数据的结构和特性的说明,而值是型的一个实例。以一定的数据模型对一个单位的数据的类型、结构及其相互间的关系所进行的描述称为数据模式。所以,数据模式仅仅是型的描述,而不涉及具体的值。实际上,数据模式是用数据模型对现实世界某一部分的模拟,反映一个单位的数据结构及其关系,具有相对稳定性。

3. 概念数据模型是一种面向用户、面向现实世界的数据库模型,与DBMS无关,它主要用来描述一个单位的概念化结构,而概念数据模式是整个数据库的逻辑结构,是一个单位的模拟。

[1.2] 请指出DB, DBS, DBMS的区别和联系。

[解] DB(Data Base)表示数据库。数据库是以一定组织方式存储在:一起的、能为多个用户所共享的、与应用程序彼此独立的、相互关联的数据集合。在物理上按一定方法存储于若干媒质,逻辑上按一定模型构成一个整体,它改进了数据操作,减少了数据冗余,可供一个机构的各类人员使用及被多种应用程序访问。

DBS(Data Base System)表示数据库系统,是指引进数据库技术的整个计算机系统。它包含数据、硬件、软件 and 用户四个部分。其中的软件包括系统软件和应用软件,也可以认为DBS包含DB和DBMS。系统软件中最主要的是DBMS和操作系统。用户分为处理用户,联机用户和系统用户。DBS的基本结构如图1-1所示。

DBMS(Data Base Management System)是数据库管理系统。是指物理文件与用户接口

之间的一层软件,是为数据库的建立、使用和维护而配置的软件。它建立在操作系统的基础之上,对 DB 实行统一的管理和控制,用户使用的各种数据库命令及应用程序的执行,都要通过 DBMS。



图 1-1 DBS 的基本结构

DBS 包括了 DB 和 DBMS,而 DBMS 又是为管理 DB 不可缺少的软件。一般将 DBMS 及其建立的 DB 总称为数据库系统即 DBS,有时也将 DB 称为 DBS,而 DBMS 是其中必不可少的系统软件。

[1.3] 数据库系统的体系结构分为内层、概念层及外层。请用下例说明三种不同的数据观。设有一个职工管理程序,其中包括职工信息,如表 1-1 所示。

表 1-1 职工情况信息表

编号	姓名	性别	年龄	工资	电话号码	住址
T1	李立明	男	40	200	4654115	东区 301
T2	肖湘	女	35	180	4213344	东区 315
T3	刘华山	男	40	220	7378412	北区 3-3
T4	周华	男	30	120	4688112	北区 2-1
T5	曹英	女	50	210	4224466	西区 412

有三个用户要使用此数据库,用户甲要求用作职工通信;用户乙要求用作人事档案管理;用户丙要求查询并输出工资在 195 元以下的职工姓名及工资。

[解] 按用户要求,分别给出三个表,如表 1-2、表 1-3、表 1-4 所示。

表 1-2 职工通信录

编号	姓名	电话号码	住址
T1	李立明	4654115	东区 301
T2	肖湘	4213344	东区 315
T3	刘华山	7378412	北区 3-3
T4	周华	4688112	北区 2-1
T5	曹英	4224466	西区 412

表 1-3 职工档案

编号	姓名	性别	年龄
T1	李立明	男	40
T2	肖湘	女	35
T3	刘华山	男	40
T4	周华	男	30
T5	曹英	女	50

表 1-4 工资低于 195 元的职工表

编号	姓名	性别	工资
T2	肖湘	女	180
T4	周华	男	120

由此看出,用户甲、乙、丙使用的是相同的数据集合,即表 1-1 的内容。该表中给出的职工详细信息称为数据库系统体系结构中的概念层,它表示数据库系统中的所有用户对数据

的总需求,是所有用户的共享数据。每个用户有特定的数据要求及结构,用户甲、乙、丙使用的数据称为数据库系统结构的外层。用户甲、乙、丙从三个不同角度使用了数据库的部分数据,这表示外层与概念层有一定的对应关系。将整体数据,即职工详细信息存储到物理介质中的存储结构与存放方式统称为数据库的内层,如图 1-2 所示。内层如何存储,由 DBMS 自动完成,不需用户干预。

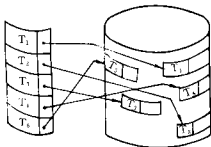


图 1-2 数据库的内层

从上述分析可知,外层是对应于用户的。一般来说,每个用户都有适当的外层与之对应,所以外层可以有多个。而概念层是为所有用户共享的,它综合所有用户的要求形成一个统一的数据存储,且只有一个。内层是实际存储的数据库,表示数据的实际存在,也只有一个。

可以从不同的角度出发,观察 DB 中的数据,这就是三种不同的数据观,即外部数据观、概念数据观和内部数据观。但是,外部数据观及概念数据观是两种观察数据的方法,并不表示

数据的实际存在,只有内部数据才表示实际存储的数据。

[1.4] 关系数据模型与网状和层次数据模型的主要区别是什么?它有何特点?

[解] 在网状和层次模型中,数据模型可以用图表示,图的结点表示记录型的集合,图中连接两结点的弧表示记录型间两两联系的集合。关系数据模型与网状和层次数据模型的最大区别是不再将数据模型表示为图的方式,而是用二维表表示实体集合属性间关系和实体集合之间的联系,数据的逻辑结构归结为满足一定条件的表格形式。

关系数据模型的特点:(1)由于表格是人们习惯使用的一种记录数据方式,因此关系数据模型具有直观性和使用的方便性。(2)关系数据模型是用数学理论来处理数据的组织和数据操作,关系方法的理论基础是关系代数。

[1.5] 给出关系的直观解释与严格定义。关系有哪些性质?

[解]

1. 关系的直观解释

关系数据模型把数据看成一张表,整个表叫做一个关系,赋予一个名字,代表实体集合。表的每一列起一个名字代表一个属性,表的每一行相当于记录,叫元组,代表一个实体。表 1-5 所示的学生成绩记载着一个班的学生成绩数据,表的每一行代表一个学生的成绩记录,每一栏代表学生成绩的一个属性。二维表由表头和表值组成,表头包含所有属性名,可以记为:

表 1-5 学生成绩

表头	姓名	性别	高等数学	程序设计	会计原理	哲学	平均成绩
表值	学文	男	90	85	92	95	90.5
	南英	女	95	90	85	80	87.5
	⋮	⋮	⋮	⋮	⋮	⋮	⋮

属性值的集合

学生成绩(姓名,性别,高等数学,程序设计,会计原理,哲学,平均成绩)

叫做关系模式,通常认为关系是一张二维表,它是元组的集合,由表头和表值组成,具有值的概念。但是,在不引起混乱的情况下,也将关系模式称为关系。关系模式中圆括号前面的叫关系名,圆括号里的是各个属性名所组成的属性名表,例如:

$R(D_1, D_2, \dots, D_n) \quad i=1, 2, \dots, n$

表示名字为 R 的关系, D_i 是属性名。

2. 关系的严格定义

给定一组域, $D_1, D_2, \dots, D_n$, 定义 $D_1, D_2, \dots, D_n$ 的笛卡尔积

$D_1 \times D_2 \times \dots \times D_n = \{ \langle d_1, d_2, \dots, d_n \rangle \mid d_i \in D_i, i=1, 2, \dots, n \}$

的子集为在域 $D_1, D_2, \dots, D_n$ 上的关系, 用 $R(D_1, D_2, \dots, D_n)$ 表示, 其中 R 是关系的名字, n 是关系的目或度。

笛卡尔积 $D_1 \times D_2 \times \dots \times D_n$ 的子集中每个元素是关系 R 的元组。例如, 若 $D_1(\text{姓名}) = \{\text{南英}, \text{学文}\}$, $D_2(\text{性别}) = \{\text{男}, \text{女}\}$, $D_3(\text{年龄}) = \{18, 20\}$, 则

$D_1 \times D_2 \times D_3 = \{ \langle \text{南英}, \text{男}, 18 \rangle, \langle \text{南英}, \text{男}, 20 \rangle, \langle \text{南英}, \text{女}, 18 \rangle, \langle \text{南英}, \text{女}, 20 \rangle, \langle \text{学文}, \text{男}, 18 \rangle, \langle \text{学文}, \text{男}, 20 \rangle, \langle \text{学文}, \text{女}, 18 \rangle, \langle \text{学文}, \text{女}, 20 \rangle \}$ 。它是所有这样的序偶的集合, 可以表示为一张二维表, 如表 1-6 所示。显然, 对一个人来说不可能既是男又是女。既是 18 岁又是 20 岁。因此, 取其子集即得 $D_1 \times D_2 \times D_3$ 上的一个关系, 代表一个实体集。如表 1-7 所示。

表 1-6

姓名	性别	年龄
南英	男	18
南英	男	20
南英	女	18
南英	女	20
学文	男	18
学文	男	20
学文	女	18
学文	女	20

表 1-7

姓名	性别	年龄
南英	女	18
学文	男	20

3. 在关系模型中, 要求关系(表格)具有以下三个性质:

(1) 表格中每个登记项必须是不可分离的数据项。

(2) 列是同质的, 即每一列里所有项目是同一类型, 各列有不同的属性名, 列的次序任意。

(3) 各行互异, 不允许有重复的行, 行的次序任意。

[1.6] 试述关系数据库模式、关系数据库、决定因素、候选关键字、关键字、主属性、非主属性和外部关键字的定义。

[解] 一个关系的属性名表 $R(A_1, A_2, \dots, A_n)$ 叫做关系模式。用于表示数据或信息的一组关系模式, 叫做关系数据库模式, 对应于这组关系模式的当前值称为关系数据库。

在关系中每个元组可以通过一个或一组属性来标识。凡能标识元组的一组属性叫做选关键字,可选定其中一个候选关键字作为关系的:关键字,简称关键字),被实际用来标识元组,具有唯一标识和无冗余性。

若关系 R 的某一属性被包含在某些关键字中,则称为主属性,而不包含在任何关键字中的属性称为非主属性。

设 X 和 Y 均为关系 R 的属性或属性组,如果 $X \rightarrow Y$ (即 Y 函数依赖于 X),则称 X 为决定因素。

若关系 R1 中的属性 X 并非 R1 的关键字,但 X 在另一个关系 R2 中是关键字,则称 X 是 R1 的外部关键字。

[1.7] 已知两个同类的学生关系 R1(学号,姓名,性别,年龄,政治面貌),R2(学号,姓名,性别,年龄,政治面貌),如表 1-8、表 1-9 所示。求 $R1 \cup R2$, $R1 - R2$, $R1 \cap R2$ 。

表 1-8 R1

学号	姓名	性别	年龄	政治面貌
93110	赵平	男	20	党员
93111	钱丽	女	21	团员
93112	赵平	女	22	群众
93113	李冬	男	19	团员
93114	周英	女	18	群众

表 1-9 R2

学号	姓名	性别	年龄	政治面貌
92111	李华	男	20	团员
92112	上官云	女	20	团员
92113	武林	男	18	党员
93110	赵平	男	20	党员
93114	周英	女	18	群众

[解]

1. $R1 \cup R2$ 为并运算,并运算的结果由属于 R1 或属于 R2 的元组组成,如表 1-10 所示。

表 1-10 $R1 \cup R2$

学号	姓名	性别	年龄	政治面貌
93110	赵平	男	20	党员
93111	钱丽	女	21	团员
93112	赵平	女	22	群众
93113	李冬	男	19	团员
93114	周英	女	18	群众
92111	李华	男	20	团员
92112	上官云	女	20	团员
92113	武林	男	18	党员

表 1-11 $R1 - R2$

学号	姓名	性别	年龄	政治面貌
93111	钱丽	女	21	团员
93112	赵平	女	22	群众
93113	李冬	男	19	团员

表 1-12 $R1 \cap R2$

学号	姓名	性别	年龄	政治面貌
93110	赵平	男	20	党员
93114	周英	女	18	群众

2. $R1 - R2$ 为差运算,差运算的结果由属于 R1 而不属于 R2 的元组组成,如表 1-11 所示。

3. 交运算 $R_1 \cap R_2$ 的结果由既属于 R_1 又属于 R_2 的元组组成,如表 1-12 所示。

[1.8] 设关系 R_3 (学号,姓名,高等数学,程序设计,会计原理,哲学)和关系 R_4 (专业,籍贯)分别如表 1-13、表 1-14 所示,关系 R_1 如表 1-8 所示。求 $R_1 \times R_4, R_1 \bowtie_{2=2} R_3, R_1 \bowtie_{2=2} R_3, \delta_{[3] < 40 \cup [4] < 40 \cup [5] < 40 \cup [6] < 60}(R_3), \prod_{2,3}(R_3)$ 。

表 1-13 R_3

学号	姓名	高等数学	程序设计	会计原理	哲学
93110	赵平	50	85	75	90
93111	钱丽	85	78	68	80
93112	赵平	70	90	80	75
93113	李冬	65	45	80	80
93224	周英	90	95	96	86

表 1-14 R_4

专业	籍贯
会计	上海
统计	北京

[解]

1. $R_1 \times R_4$ 为乘积运算,运算的结果是将它们的属性按笛卡尔乘积组合在一起,如表 1-15 所示。

表 1-15 $R_1 \times R_4$

学号	姓名	性别	年龄	政治面貌	专业	籍贯
93110	赵平	男	20	党员	会计	上海
93110	赵平	男	20	党员	统计	北京
93111	钱丽	女	21	团员	会计	上海
93111	钱丽	女	21	团员	统计	北京
93112	赵平	女	22	群众	会计	上海
93112	赵平	女	22	群众	统计	北京
93113	李冬	男	19	团员	会计	上海
93113	李冬	男	19	团员	统计	北京
93114	周英	女	18	群众	会计	上海
93114	周英	女	18	群众	统计	北京

2. $R_1 \bowtie_{2=2} R_3$ 为连接运算,即将 R_3 的第二属性值等于 R_1 的第二属性值的元组连接在 R_1 的相应元组之后,如表 1-16 所示。

3. $R_1 \bowtie R_3$ 为自然连接。自然连接是按行(或元组)对两个关系中相同属性名(即学号,姓名)的值进行比较,如果相等,则从后一个关系(R_3)的这一行属性中去掉被比较的属性,并把剩下的属性按原序接在前一个关系(R_1)的同行属性的右边,如表 1-17 所示。

表 1-16 $R_1 \bowtie R_3$

学号	姓名	性别	年龄	政治面貌	高等数学	程序设计	会计原理	哲学
93110	赵平	男	20	党员	50	85	75	90
93111	钱丽	女	21	团员	85	78	68	80
93112	赵平	女	22	群众	70	90	80	75
93113	李冬	男	19	团员	65	45	80	80
93114	周英	女	18	群众	90	95	95	86
93224	周英	女	18	群众	90	95	95	86

表 1-17 $R_1 \bowtie R_3$

学号	姓名	性别	年龄	政治面貌	高等数学	程序设计	会计原理	哲学
93110	赵平	男	20	党员	50	85	75	90
93111	钱丽	女	21	团员	85	78	68	80
93112	赵平	女	22	群众	70	90	80	75
93113	李冬	男	19	团员	65	45	80	80

4. $\sigma_{[a] < 60 \vee [c] < 60 \vee [d] < 60 \vee [e] < 60 \vee [f] < 60 \vee [g] < 60}(R_3)$ 是选择运算, 即选择出成绩不及格的学生, 如表 1-18 所示。

表 1-18 $\sigma_{[a] < 60 \vee [c] < 60 \vee [d] < 60 \vee [e] < 60 \vee [f] < 60 \vee [g] < 60}(R_3)$

学号	姓名	高等数学	程序设计	会计原理	哲学
93110	赵平	50	85	75	90
93113	李冬	65	45	80	80

5. $\Pi_{2,3}(R_3)$ 为投影运算, 即在关系 R_3 中保留第二和第三属性, 去掉其它属性及其相应的属性值, 其运算结果如表 1-19 所示。

[1.9] 试述数据库设计的特征和基本步骤。

[解] 数据库设计和其它工程设计一样, 具有下列特征:

1. 反复性。数据库设计不可能“一气呵成”, 需要反复推敲和修改才能完成。

2. 试探性。数据库设计的过程往往是试探的过程, 有时多种方案并存, 供设计者选择, 而这种选择不是完全客观的, 多少取决于设计者的偏爱和观点, 因而数据库设计的结果往往不是唯一的。

3. 分步进行。设计数据库应分阶段进行, 前一阶段的设计结果可以作为后一阶段设计的依据, 后一阶段也可以向前面的设计阶段反馈其要求, 反复修改, 逐步完善。

表 1-19 $\Pi_{2,3}(R_3)$

姓名	高等数学
赵平	50
钱丽	85
赵平	70
李冬	65
周英	90

数据库设计的基本步骤如图 1-3 所示。

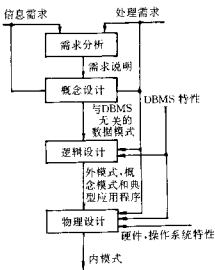


图 1-3

第一步是需求分析。调查分析一个单位的信息需求和处理需求，了解用户的业务情况和数据流向，并整理成需求说明。需求说明既是需求分析阶段的成果，又是数据库设计的依据，它包括数据库所涉及数据范围和各数据特征的描述以及数据量大小和使用频率的估计。例如数据名、属性及其类型、长度、取值范围、属性之间的联系和约束条件、主关键字、保密要求、使用频率、更改要求、数据量估计等。

第二步是概念设计。即用概念数据模型表示数据及其相互关系。在概念设计阶段，可以吸收用户参加和评价数据库设计。可以先设计各个用户的视图，即每个用户所看到的数据及其结构，然后进行视图的集成，以形成一个单位的数据模式。初步形成的数据模式再由设计者和用户审查修改，以保证所设计的数据模式能正

确地、完全地反映一个单位的数据及其相互关系，又便于进行所需的各种处理。

第三步是逻辑设计。把第二步所得到的与 DBMS 无关的数据模式，按照一定方法转化成用 DBMS 逻辑数据模型表示的概念模式（即整个数据库的逻辑结构，也称逻辑模式），同时也将第二步所得到的用户视图转化成外模式（即用户见到的那一部分数据的逻辑结构）。应该指出，逻辑设计不单纯是转化，而是在转化（翻译）过程中进一步落实需求说明，满足 DBMS 的各种限制。数据库逻辑设计的结果将以 DBMS 所提供的数据库定义语言（DDL）表示。

第四步是物理设计。主要根据 DBMS 提供的各种手段，设计数据库的内模式（物理结构），即设计数据的存储形式和存取路径，例如文件结构、索引的设置等。

以上四步不是完全顺序地进行的，后面的设计步骤可以反馈其要求到前面的各步，反复进行修改。

[1.10] 什么叫逐步规范化？规范化的目的何在？是否规范化程度越高模式就越好？

【解】逐步规范化就是用更单纯、更规则的关系来取代原有关系的过程。这个过程是可逆的，一组关系转换为另一组关系后，仍然可以恢复到原有的那种关系，规范化过程并不丢失任何信息，并保持属性间的依赖性。

规范化的目的：(1) 将关系的“概念”单一化，使每一个数据项是一个简单的数或字符串，而不是一个数组或重复组，从而方便用户使用；(2) 解决冗余度问题；(3) 解决多义性问题；(4) 解决操作可行性和提高操作方便性；(5) 使关系模式更灵活、易于实现接近自然语言的查询方式。

规范化的程度不是越高越好。这是因为规范化程度越高虽然关系简捷明了，有方便的一面，但分解过分会引起新的问题，如表格的层次一旦被拆散，将导致：①层次特性消失，用户很难看出层次关系；②数据处理困难，往往检索开销大；③存储相关语义特性无法描述，以致

删除一个文件中某项,而其它文件中相关项没有删去。所以在数据库设计中,必须根据应用环境和具体情况,合理地选择数据库模式。在一般情况下,关系达到 3NF 就能满足数据库应用的需要。

[1.11] 给出规范关系的判定条件,并指出下列关系中的决定因素、候选关键字、关键字、主属性,各关系属于哪种范式?

- (1) 船 1(船号,船名,部件(部件号,部件名,使用量))
- (2) 学生(学号,姓名,班级,班主任,课程号,成绩)
- (3) 职工(编号,姓名,工资级,工资)
- (4) 船 2(船号,船名,吨位,马力,颜色)
- (5) 学生档案(学号,姓名,年龄,所在系,在校住址)

[解]

1. 规范关系的判定条件

(1) 若关系中每个属性都是不可再分的基本数据项,即不出现“表中有表”或不出现“关系中含有重复组或另外一个关系”,则该关系属于第一范式(1NF)。

(2) 对于一个 1NF 的关系 R,如果 R 的每一个非主属性都完全函数依赖于关键字,则 R 就是一个第二范式(2NF)的关系。

(3) 如果 R 是一个 2NF 的关系,并且它的每一个非主属性都不传递函数依赖于关键字,即非主属性是相互独立的,相互之间不存在函数依赖,则 R 是一个第三范式(3NF)的关系。

2. 指出关系中的决定因素、候选关键字、关键字、主属性及属于哪种范式

(1) 在关系:船 1(船号,船名,部件(部件号,部件名,使用量))中,船号决定船名和船使用的部件;部件号决定部件名;(船号,部件号)决定部件使用量。因此,船号,部件号,(船号,部件号)是决定因素。(船号,部件号)既为候选关键字又为关键字。主属性是船号,部件号。该关系模式由于存在“表中有表”,因此是非规范化关系。

(2) 在关系:学生(学号,姓名,班级,班主任,课程号,成绩)中,学号,班级,课程号是决定因素。(学号,班级,课程号)是候选关键字。(学号,课程号)是关键字。学号,课程号是主属性。该关系属于 1NF。

(3) 在关系:职工(编号,姓名,工资级,工资)中,编号,工资级是决定因素。候选关键字、关键字、主属性均为编号。该关系属于 2NF。

(4) 在关系:船 2(船号,船名,吨位,马力,颜色)中,船号,颜色是决定因素(因为颜色决定吨位)。候选关键字、关键字、主属性均为船号。该关系属于 2NF。

(5) 在关系:学生档案(学号,姓名,年龄,所在系,在校住址)中,决定因素、候选关键字、关键字、主属性均为学号。该关系属于 3NF。

[1.12] 有如表 1-20、表 1-21 所示的两张表,试写出对应的关系模式并规范化为 3NF。

表 1-20 图 书

书号	书名	出版社	出版社地址	借书证号	读者姓名	读者单位	单位电话	借阅日期

表 1-21 ××产品存货明细帐

××年		凭证		摘要	收入			发出			结存		
月	日	字	号		数量	单价	金额	数量	单价	金额	数量	单价	金额

[解]

1. 表 1-20 的关系模式为:

图书(书号,书名,出版社,出版社地址,借书证号,读者姓名,读者单位,单位电话,借阅日期)
 根据语义可以找出这个关系的全部函数依赖,画出函数依赖如图 1-4 所示。该关系中各属性之间的对应关系也可用图 1-5 表示。

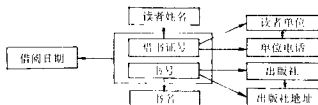


图 1-4

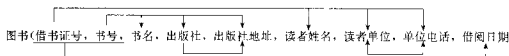
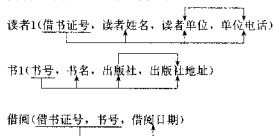


图 1-5

图书关系模式中,非主属性“书名”、“出版社”、“出版社地址”函数依赖于主属性“书号”,非主属性“读者姓名”、“读者单位”、“单位电话”函数依赖于“借书证号”,因此这些非主属性均为部分函数依赖主关键字(借书证号,书号),只有“借阅日期”完全函数依赖于(借书证号,书号),因此图书关系属于 1NF。为了消除部分函数依赖,对图书关系模式进行分解,即用投影运算将图书关系拆成三个关系:



相应的函数依赖图如图 1-6(a)、(b)、(c)所示:

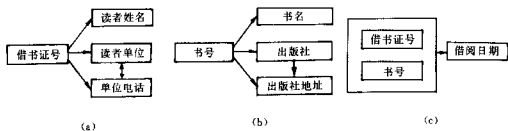


图 1-6

在这三个关系中,都不存在非主属性对主关键字的部分函数依赖,其中读者1关系和书1关系中存在传递函数依赖,故属于 2NF,借阅关系已达到 3NF。

再进行投影运算,消除非主属性对主关键字的传递函数依赖,便可将 2NF 变换成 3NF。

读者2(借书证号, 读者姓名, 读者单位)

单位(读者单位, 单位电话)

书2(书号, 书名, 出版社)

出版社(出版社, 出版社地址)

借阅(借书证号, 书号, 借阅日期)

这五个关系联合起来表达了图书关系要表达的内容。

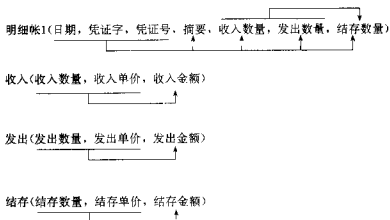
2. 表 1-21 的关系模式为:

产品存货明细帐(日期,凭证(凭证字,凭证号),摘要,收入(数量,单价,金额),发出(数量,单价,金额),结存(数量,单价,金额))

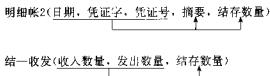
由于关系中,存在“表中有表”,所以是非规范化的关系,消除“表中表”使之变换为规范化形式:



进行投影运算,消除部分函数依赖,得出:



其中,收入、发出、结存三个关系已属于 3NF,而关系明细帐 1 还存在传递函数依赖,属于 2NF,再进行投影运算,得:



明细帐 2、结一收发、收入、发出、结存这五个关系均属于 3NF,可以表示产品存货明细帐的内容。

[1.13] 简述 FoxBASE+ 的特点、运行环境及主要性能指标。

[解] FoxBASE+ (FoxBASE PLUS) 是美国 Fox software 公司 80 年代中期开发成功的多用户关系数据库管理系统。现以 FoxBASE+ 2.10 版本为例加以说明。

FoxBASE+ 的特点及运行环境:

1. 和 dBASE III PLUS 完全兼容,通过系统提供的命令,可以将 dBASE III 环境下运行的程序不作任何修改就可转到 FoxBASE 环境下运行。
2. 比 dBASE III 增加了数十条命令和函数。在性能和功能上有许多扩充,运行速度比 dBASE III 快 5.9 倍,如配置 8087/80387 协处理器,则运算速度可进一步提高。
3. 可运行于多个操作系统 (如 MSDOS, UNIX...) 及多机器环境 (IBM/PC, VAX11/780...)。多用户版本与单用户版本兼容。

FoxBASE+ 的主要性能指标:

1. 数据库文件最大记录个数 10 亿,字段数 128 个,记录长度 4000 字节。
2. 同时访问的数据库文件个数 10 个,可同时打开的各类文件数 48 个。
3. 内存变量 (包括数组) 最大个数 3600,字符型变量可用字节数 64K,每个数组的元素数 3600 个。
4. 数值计算中有效数 16 位,字符串长度 254 字节,命令行长度 254 字节。支持单用户 FoxBASE+, 最低 DOS 版本为 2.00,支持多用户 FoxBASE+, 最低 DOS 版本为 3.10 版本。

[1.14] 试回答下列问题

1. 数据与信息区别和联系。

② 简述说明记录型与记录值的区别、联系。

③ 数据库的三大数据模型是什么?

1. 举例说明函数依赖、部分函数依赖及传递函数依赖的概念。

2. 举例说明物流、信息流、实体、属性等概念的含义。

④ 数据库与文件有何联系与区别?

⑤ 解释什么是字节、数据项、记录、文件、物理记录。

【1.15】 试问下列关系模式最高属于第几范式,并解释原因。

1. $R(A, B, C, D)$ 函数依赖的集合 $F = \{B \rightarrow D, AB \rightarrow C\}$

2. $R(A, B, C, D, E)$ $F = \{AB \rightarrow CE, E \rightarrow AB, C \rightarrow D\}$

3. $R(A, B, C, D)$ $F = \{B \rightarrow D, D \rightarrow B, AB \rightarrow C\}$

4. $R(A, B, C)$ $F = \{A \rightarrow B, B \rightarrow A, C \rightarrow A\}$

【1.16】 已知关系 R, S 分别如表 1-22, 表 1-23 所示, 求 $R \cup S, R - S, R \cap S$

R	M	N	P
	c	3	d
	b	4	f
	r	3	e

S	M	N	P
	b	1	f
	r	3	e
	a	3	i

【1.17】 设关系 U, V 如表 1-24, 表 1-25, 表 1-26 所示, 求 $U \bowtie_{N>S+T} V, U \bowtie V, \Pi_{N,S}(U)$

【1.18】 设有一部件关系, 其关系模式及部分值如下:

表 1-24

M	N	P
1	6	
3	3	
5	2	

表 1-25

V	Q	M	S	T
	d	a	2	1
	d	b	3	4
	b	d	4	2

表 1-26

Z	W	X
	c	h
	i	j

【1.18】 设有一部件关系, 其关系模式及部分值如下:

部件(部件号, 部件名, 现有数量, 工程项目(项目编号, 项目内容, 项目负责人号, 已提供部件数量))

12	车床	007	5
73	铣床	086	7
12	车床	007	33
29	冲床	012	25
36	减速器	024	16

【1.19】 试找出(1)~(4)题中(1)、(2)、(3)、(4)共四个关系分别规范化为 3NF。

【1.19】 试找出(1)~(4)题中(1)、(2)、(3)、(4)共四个关系分别规范化为 3NF 的关系。

[1.20] 现金日记帐如表 1-27 所示,试写出现金日记帐的借贷方、对方科目、类页、借方、贷方、余额的关系。

表 1 27 现金日记帐

××年		凭证		摘要	对方科目	类页	借方	贷方	余额
月	日	类别	号数						

[1.21] 表 1-28 是总分分类帐,试设计总分分类帐数据库。

表 1 28 总分分类帐

会计科目:

××年		凭证号	摘要	借方	贷方	核对号	借或贷	余额
月	日							

第二章 dBASE III /FoxBASE⁺ 语法基础

[2.1] 根据命令的结构格式,试对以下各命令说明每部分的含义。

1. DISPLAY NEXT 3 姓名,政治面貌,入学分 FOR 性别='男'
2. LIST FOR 性别='男'.AND.'华'\$姓名.OR. 入学分=470
3. REPLACE ALL 工龄工资 WITH 工龄工资+1

【解1】 DISPLAY 是命令动词,表示显示当前数据库文件的记录;NEXT 3 指出显示记录的范围是包括当前记录及其以下的共三个记录;姓名,政治面貌,入学分是字段名表,表示仅显示这三个字段的值,FOR 性别='男'是显示条件,表示使表达式性别='男'值为真的记录才予以显示。该命令的完整表达是:显示当前数据库文件中包括当前记录及其后面的共三个记录中性别字段值为男的姓名、政治面貌、入学分三个字段值。

【解2】 LIST 是命令动词,表示显示当前数据库文件的记录;FOR 后面的内容是条件表达式,只有使条件表达式值为真的记录才显示;范围项缺省,系统隐含显示全部记录,表达式表缺省,表示显示全部字段值。该命令的完整表达是:显示当前数据库文件中所有性别字段值为男,并且姓名字段中含有'华'的记录全部字段值,或入学分字段值为470的全部记录。

【解3】 REPLACE 是命令动词,表示用表达式值替换字段原来的内容;ALL 表示命令的操作范围是全部记录;工龄工资 WITH 工龄工资+1 是表达式表。该命令的完整表达是:对当前数据库文件的全部记录的工龄工资字段用表达式工龄工资+1 的值替换。

[2.2] 将数值0赋给内存变量A,B,C;字符串YES赋给内存变量A1;逻辑数真赋给内存变量A2;日期型数92年12月1日赋给内存变量A3。

【解】

方法一:

- STORE 0 TO A,B,C
- STORE 'YES' TO A1
- STORE .T. TO A2
- STORE CTOD ('12/01/92') TO A3

方法二:

- A=0
- B=0
- C=0
- A1='YES'
- A2=.T.

• A3=CTOD ('12/01/92')

方法三:

• INPUT '请输入数据:' TO A

请输入数据: 0 (下面划线部分是键入信息, ↵ 表示回车,下同)

• INPUT '请输入数据:' TO B

请输入数据: 0

• INPUT '请输入数据:' TO C

请输入数据: 0

• INPUT '请输入字符串:' TO A1

请输入字符串: 'YES'

• INPUT '请输入逻辑型数据:' TO A2

请输入逻辑型数据: .T.

• INPUT '请输入日期型数据:' TO A3

请输入日期型数据: CTOD ('12/01/93') ↵

方法四:

• ACCEPT '请输入数据:' TO A

请输入数据: 0 ↵

• ? TYPE ('A')

C

• ACCEPT '请输入字符串:' TO A1

请输入字符串: YES ↵

• ACCEPT '请输入逻辑型数据:' TO A2

请输入逻辑型数据: .T. ↵

• ? TYPE ('A2')

C

• ACCEPT '请输入日期型数据:' TO A3

请输入日期型数据: CTOD ('12/01/93') ↵

• ? TYPE ('A3')

C

方法五:

• WAIT '请输入数据:' TO A

请输入数据: 0

• ? TYPE ('A')

C

• WAIT '请输入字符串:' TO A1

请输入字符串: Y

说明:

1. 方法一、方法二是同一赋值命令的两种形式,STORE 适合给多个变量赋同一值。赋值的类型可以是 N,C,L,D 型。