

数据库技术基础

基本概念

数据库和数据库系统

一般来说,数据库系统是由 **数据库,硬件,软件和人员**组成的

数据库管理系统(DBMS)的功能

1. 数据定义. 使用数据定义语言(DDL)对数据库的结构进行描述.包括外模式,模式和内模式定义.完整性定义,安全和存取权限等.这些定义存储在数据字典中.是DBMS运行的基本依据
2. 数据操作. 使用数据操作语言(DML)实现对数据的增删改查索引等操作.DML分为宿主型和自含型
3. 数据库运行管理
4. 数据的组织,存储和管理.
5. 数据库的维护

数据库系统特征和分类

特点

1. 数据结构化切统一管理.
2. 拥有较高的数据独立性
3. 数据控制功能. 安全性,完整性,并发控制.故障恢复

分类

1. 关系型数据库.
2. 面向对象型数据库
3. 对象关系型数据库

数据库体系结构

1. 集中式数据库,数据和数据管理都是集中的
2. 客户端/服务器结构
 - 事物服务器 也称查询服务器.
 - 数据服务器
3. 并行数据库系统
 - 共享内存式多处理器
 - 无共享式并行体系
4. 分布式数据库系统

数据库的三级模式

三级模式指的是是巨款管理系统的内部结构: **内模式,模式和外模式**

- 内模式 存储模式. 关注的物理结构的存储.
- 模式也称概念模式. 关注的是数据的逻辑结构. 比如库,表,
- 外模式 对应的是数据库视图.关注的是用户对数据的操作.

- 模式/内模式 可以看作数据的存储形式和表之间的关系
- 模式/外模式 可以看作是表和用户操作之间的关系

数据模型

数据模型是对现实世界数据特征的抽象.常见的有 概念数据模型(实体-联系模型)和基本数据模型(层次模型,网状模型,面向对象模型等)

数据模型三要素










1. 数据结构
2. 数据操作
3. 数据的约束条件

E-R模型

在ER图中:

- 实体用矩形表示,关系用菱形表示.属性用椭圆表示.
- 子类和父类的连线中间有一个圆圈.
- 双边矩形边框表示子类
- 超类到圆圈有双线连接表示全特殊化,单线连接表示部分特殊化.
- 有符号U的线表示特殊化,其中圆圈里面是d的表示不相交特殊化.o表示重叠特殊化.

表 9-1 E-R 图中的主要构件

构 件	说 明
矩形 	表示实体集
双边矩形 	表示弱实体集
菱形 	表示联系集
双边菱形 	表示弱实体集对应的标识性联系
椭圆 	表示属性
线段 	将属性与相关的实体集连接，或将实体集与联系集相连
双椭圆 	表示多值属性
虚椭圆 	表示派生属性
双线 	表示一个实体全部参与到联系集中

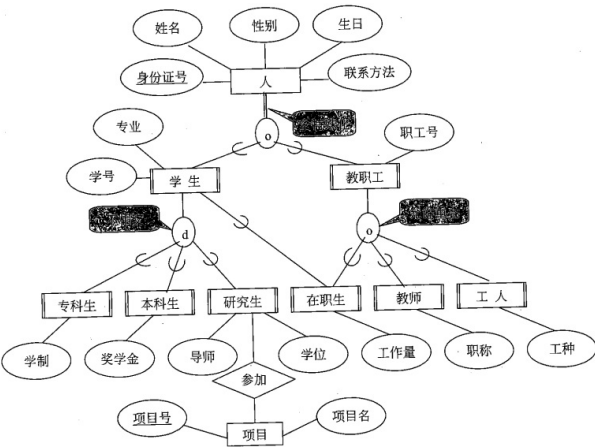


图 9-11 特殊化应用实例

关系代数

- \cup 并: 关系R和S**必须有相同的模式**,列/属性必须相同才能进行并运算.是记录的运算,结果去重.
- \cap 交: 关系R和S**必须有相同的模式**,列/属性必须相同才能进行交运算..是记录的运算
- $-$ 差: 关系R和S**必须有相同的模式**,差运算有方向,是记录的运算,.R-S就是在R中但是不在S中的记录.
- \times 笛卡尔积: $C = A \times B$ 笛卡尔积的元组/记录,一半来自关系A,一半来自关系B.笛卡尔积的元数是原来两个关系的元数的和.笛卡尔积的基数是原来两个关系的基数的乘积.
- π 投影: 从关系中选择列的操作,是关系的垂直操作,广义投影允许使用算术运算符进行扩充
- σ 选择: 从关系中选择行的操作,是关系的水平操作.
- \bowtie 连接: 从笛卡尔积的结果用以一定的条件进行筛选.连接有3种, θ 等值和自然连接:
 1. $\bowtie_{x\theta y}$ θ 连接.,相当与 where $x\theta y$
 2. $\bowtie_{x=y}$ 等值连接.相当与 where $x=y$
 3. \bowtie 自然连接.相当与把两个关系中相等的属性进行比较.并在结果中去掉重复的属性列.
- **外连接**: 为了保留因为连接而丢失的记录.还有三种外连接: 左外连接,右外连接和全外连接.
- \div 除: 除是同时进行水平和垂直方向的操作.关系R(X,Y)和关系S(Y,Z)必须有相等的属性组

A	B	C	D
a	b	c	d
a	b	e	f
b	c	e	f
e	d	c	d
e	d	e	f
a	b	d	e

R

C	D
c	d
e	f

S

A	B
a	b
e	d

$R \div S$

解题方法是:

1. 先求关系S在Y上的投影,本例中,Y是C和D.那么关系S在Y上的投影就是(c,d)和(e,f) $\pi_{CD} = \{(c, d), (e, f)\}$
2. 再求关系组X的分量的象集.关系组X在本里中是A和B,X在R上分量x就是{(a,b), (b,c), (e, d)}.那他们的象集(象集是指x中的每一个元素在R关系模式下,对应的Y的部分).这些象集是: $CD_{(a,b)} = \{(c, d), (e, f), (d, e)\}; CD_{(b,c)} = \{(e, f)\}; CD_{(e,d)} = \{(c, d), (e, f)\}$,这三个象集中,包含(必须包含关系S在Y上的投影 π_{CD} 的全部元素) $\pi_{CD} = \{(c, d), (e, f)\}$ 有{(a,b), (e,d)}. 因此, $R \div S$ 的结果就是{(a,b), (e,d)}

提醒:

- **除**常用来构建从表A中选择属性c等于m或者n的b和c属性这样的问题. $\pi_{b,c}(A) \div \pi_c(\sigma_{c=m \vee c=n}(A))$,其中 $\pi_c(\sigma_{c=m \vee c=n}(A))$ 是从A表中选择c等于m或者等于n的记录的全挑出去然后做c的投影.然后再从表中筛选c属于等于这个子查询结果的b和c的投影.

请仔细看教材课本.关系代数是必考,特别是关系代数和sql查询语句的对应关系

关系型数据库SQL

常用命令

创建表

```
create table table_name(
    column_name1 data_type limit...,
    ...
    PRIMARY KEY(column_name), # 主键
    FOREIGN KEY (column_name) REFERENCES table_name2(column_name), # 外键
)
```

修改和删除表

```
alter table table_name add new_column_name data_type 约束条件; # 添加列
alter table table_name drop column_name; # 删除列
alter table table_name modify column_name data_type; # 修改列
drop table table_name;
```

求两个查询结果的交集 intersect

关系型数据库的规范化

函数依赖

- 部分函数依赖, 对于 $A \rightarrow B$ 如果存在A的一部分a, $a \rightarrow B$, 则称存在部分函数依赖. 通俗的来说, 就是A能决定B, 如果部分A也能决定B就是B部分依赖A
- 完全函数依赖. A能决定B, 不存在部分A能决定B的情况就是B完全依赖A
- 传递函数一来 如果A能决定B, B能决定C, 则称C对A有传递函数依赖

规范化

候选键要求能后唯一标识实体. 主键是从候选键里选出的
主属性是指在候选键里出现过的属性.

- 第一范式** 属性都是原子的, 不可再分. (能决定唯一性的复合主键也是原子的)
- 第二范式** 非主属性对候选键没有部分函数依赖.
- 第三范式** 非主属性对候选键没有传递函数依赖. 3NF是对字段冗余性的约束, 即任何字段不能由其他字段派生出来, 它要求字段没有冗余
- BC范式** 存在任何字段对任一候选关键字的传递函数依赖则符合第三范式. BC范式既检查非主属性, 又检查主属性. p判断的标准是这个关系模式的函数以来的决定因素(左边)必定包含关系模式的候选键.

规范化理论习题

某公司的部门（部门号，部门名，负责人，电话）、商品（商品号，商品名称，单价，库存量）和职工（职工号，姓名，住址）三个实体之间的关系如表1、表2和表3所示。假设每个部门有一位负责人和一部电话，但有若干名员工；每种商品只能由一个部门负责销售。部门关系不属于第三范式的原因是__ (1) __。如果用户要求得到表4所示的结果，需要__ (2) __，并增加关系模式__ (3) __。

A. 没有消除非主属性对码的部分函数依赖
如：部门名→负责人

B. 没有消除非主属性对码的部分函数依赖
如：负责人→电话

C. 只消除了非主属性对码的部分函数依赖而未消除传递函数依赖

D. 没有消除非主属性对码的部分函数依赖和传递函数依赖

A. 修改表1的结构，在表1中增加一个职工号

B. 修改表2的结构，在表2中增加一个职工号

C. 修改表2的结构，在表2中增加一个部门号

D. 修改表3的结构，在表3中增加一个部门号

A. 销售（职工号，商品号，日期，数量）

B. 销售（职工号，商品名称，商品号，数量）

C. 销售（职工号，部门号，日期，数量）

D. 销售（职工号，部门号，商品号，日期）

部门号	部门名	负责人	电话
001	家电部	E002	1001
002	百货部	E026	1002
003	食品部	E030	1003

商品号	商品名称	单价	库存量
30023	微机	4900	26
30024	打印机	1650	7
...
30101	毛巾	10	106
30102	牙刷	3.8	288

职工号	姓名	住址
E001	王 军	南京路
E002	李晓域	淮海路
E021	杨 烨	江西路
E026	田 渡	西藏路
E028	李晓域	西藏路
E029	刘丽华	淮海路
E030	李青彬	唐山路
E031	胡慧芬	昆明路

职工号	姓名	部门名	月销售额
E001	王 军	家电部	528900
E002	李晓域	家电部	368000
E021	杨 烨	百货部	12500
E028	李晓域	百货部	82500
E031	胡慧芬	食品部	282608

解题

- 第一题, 由于属性都是单值的, 所以满足第一范式. 部门表(1)的主键是部门号. 因而这个主键是单值属性. 所以非主属性不存对

主属性的部分函数依赖.满足第二范式.这样如果部门表不属于第三范式,那一定是有传递函数依赖(虽然我们不知道这个传递函数依赖是指什么).使用排除法.第一题选3

- 第二题.想得到部门名称,在职工表添加一个部门号即可.选c
- 第三题.表4中缺少的是商品销售信息.那么需要添加的是和商品有关的信息.c被排除.由于职工信息里有部门号,d中的部门号是多余的,所以d也被排除.而b中的商品名称对商品号来说是冗余信息.所以b也被排除.选a.

模式分解

- 无损分解 分解具有无损连接特性.
- 保持函数依赖 分解要保持函数依赖关系
- 既要无损又要保持函数依赖.

模式分解例题

思考题:

有关系模式: 成绩 (学号, 姓名, 课程号, 课程名, 分数)

函数依赖: 学号→姓名, 课程号→课程名, (学号, 课程号) →分数

若将其分解为:

成绩 (学号, 课程号, 分数)

学生 (学号, 姓名)

课程 (课程号, 课程名)

请思考该分解是否为无损分解?

由于有: 学号→姓名, 所以:

成绩 (学号, 课程号, 分数, 姓名)

由于有: 课程号→课程名, 所以:

成绩 (学号, 课程号, 分数, 姓名, 课程名)

解题:

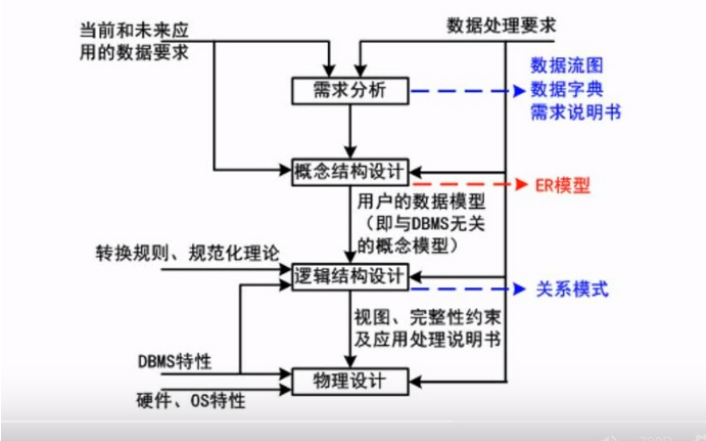
- 无损且保持函数依赖

并发控制

并发控制主要是封锁.基本的封锁手段有排他锁(X锁/写锁),和共享锁(S锁/读锁)

数据库设计应用题

数据库设计过程



数据库设计主要考察的是E-R模型和关系模式以及两者之间的转换.

试题1

❁ 试题1

希赛公司拟开发一个宾馆客房预订子系统，主要是针对客房的预订和入住等情况进行管理。

【需求分析结果】

(1) 员工信息主要包括：员工号、姓名、出生年月、性别、部门、岗位、住址、联系电话和密码等信息。岗位有管理和服务两种。岗位为“管理”的员工可以更改（添加、删除和修改）员工表中的本部门员工的岗位和密码，要求将每一次更改前的信息保留；岗位为“服务”的员工只能修改员工表中本人的密码，且负责多个客房的清理等工作。

(2) 部门信息主要包括：部门号、部门名称、部门负责人、电话等信息；一个员工只能属于一个部门，一个部门只有一位负责人。

(3) 客房信息包括：客房号、类型、价格、状态等信息。其中类型是指单人间、三人间、普通标准间、豪华标准间等；状态是指空闲、入住和维修。

(4) 客户信息包括：身份证号、姓名、性别、单位和联系电话。

(5) 客房预定情况包括：客房号、预定日期、预定入住日期、预定入住天数、身份证号等信息。一条预定信息必须仅对应一位客户，但一位客户可以有多个预定信息。

【概念模型设计】

根据需求阶段收集的信息，设计的实体-联系图（不完整）如图12-1所示。

❁ 试题1

【问题1】（3分）

根据问题描述，填写图12-1中（1）~（3）处联系的类型。联系类型分为一对一、一对多和多对多3种，分别使用1:1、1:n或1:*、m:n或*:*表示。

【问题2】（2分）

补充图12-1中的联系并指明其联系类型。

【问题3】（7分）

根据需求分析结果和图12-1，将逻辑结构设计阶段生成的关系模式中的空（4）~（8）补充完整。（注：一个空可能需要填多个属性。）

【问题4】（3分）

若去掉权限表，并将权限表中的操作权限属性放在员工表中（仍保持管理和服务岗位的操作权限规定），则与原有设计相比有什么优缺点（请从数据库设计的角度进行说明）。

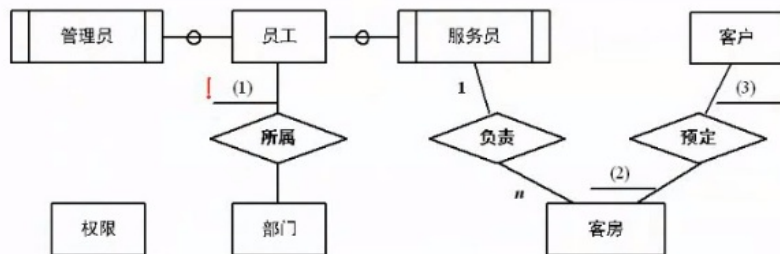


图12-1 实体-联系图

【逻辑结构设计】

逻辑结构设计阶段设计的部分关系模式（不完整）如下：

员工（（4），姓名，出生年月，性别，岗位，住址，联系电话，密码）

权限（岗位，操作权限）

部门（部门号，部门名称，部门负责人，电话）

客房（（5），类型，价格，状态，入住日期，入住时间，员工号）

客户（（6），姓名，性别，单位，联系电话）

更改权限（员工号，（7），密码，更改日期，更改时间，管理员号）

预定情况（（8），预定日期，预定入住日期，预定入住天数）

分析题干：

- 员工是一个关系,可能包含信息: 员工号,姓名,出生年月,性别,部门,岗位,住址,联系电话,密码等.
- 岗位有管理和服务2种.管理和服务
- 管理负责管理服务的信息.
- 服务负责客房清理
- 部门信息包含 部门号,部门名称,负责人,电话等. 部门和员工是1:n,和负责人是1:1
- 客房信息包含 客房号,类型,价格,状态.
- 客户信息包含 身份证号,姓名,性别, 单位,联系电话.
- 客房预订信息 客房号,预订入住日期,预订入住天数,身份证号码等.
- 客户可以预订多个客房,客房也可以在不同时间被不同的客户预订.两者是m:n的关系

解题

- 第一问: 1是1:m,2和3是m:n
- 第二问: 缺少权限和员工的联系, 权限和员工是1:m的联系(无需注明联系名)
- 第三问: 4 员工号,部门号;5 客房号;6 身份证号码;7 (原来的)岗位; 8 身份证号码,客房号
- 第四问: 操作权限属性放入员工共表中是一种反范式化的行为,优点是减少一次join操作,性能更高.缺陷是有数据的冗余.

试题2

阅读下列说明，回答问题1至问题3，将解答填入答题纸的对应栏内。

【说明】
 某集团公司拥有多个大型连锁商场，公司需要构建一个数据库系统以方便管理其业务运作活动。
 【需求分析结果】
 1.商场需要记录的信息包括商场编号（编号唯一），商场名称，地址和联系电话。某商场信息如表12-1所示。
 表12-1 商场信息表

商场编号	商场名称	地址	联系电话
PS2101	淮海商场	淮海中路918号	021-64158818
PS2902	西大街商场	西大街时代盛典大厦	029-87283220
PS2903	东大街商场	碑林区东大街239号	029-87450287
PS2901	长安商场	雁塔区长安中路38号	029-85264953

2. 每个商场包含有不同的部门，部门需要记录的信息包括部门编号（集团公司分配），部门名称，位置分布和联系电话。某商场的部门信息如表12-2所示。
 表12-2 部门信息表

部门编号	部门名称	位置分布	联系电话
DT002	账务部	商场大楼六层	82504342
DT007	后勤部	商场地下副一层	82504347
DT021	安保部	商场地下副一层	82504358
DT005	人事部	商场大楼六层	82504446
DT001	管理部	商场裙楼三层	82504668

3. 每个部门雇用多名员工处理日常事务，每名员工只能隶属于一个部门（新进员工在培训期不隶属于任何部门）。员工需要记录的信息包括员工编号（集团公司分配），姓名，岗位，电话号码和工资。员工信息如表12-3所示。

3. 每个部门雇用多名员工处理日常事务，每名员工只能隶属于一个部门（新进员工在培训期不隶属于任何部门）。员工需要记录的信息包括员工编号（集团公司分配），姓名，岗位，电话号码和工资。员工信息如表12-3所示。

表12-3 员工信息表

员工编号	姓名	岗位	电话号码	工资
XA3310	周 超	理货员	13609257638	1500.00
SH1075	刘 飞	防损员	13477293487	1500.00
XA0048	江雪花	广播员	15234567893	1428.00
BJ3123	张正华	部门主管	13345698432	1876.00

4. 每个部门的员工中有一名是经理，每个经理只能管理一个部门，系统需要记录每个经理的任职时间。

【概念模型设计】
根据需求阶段收集的信息，设计的实体联系图和关系模式（不完整）如下：



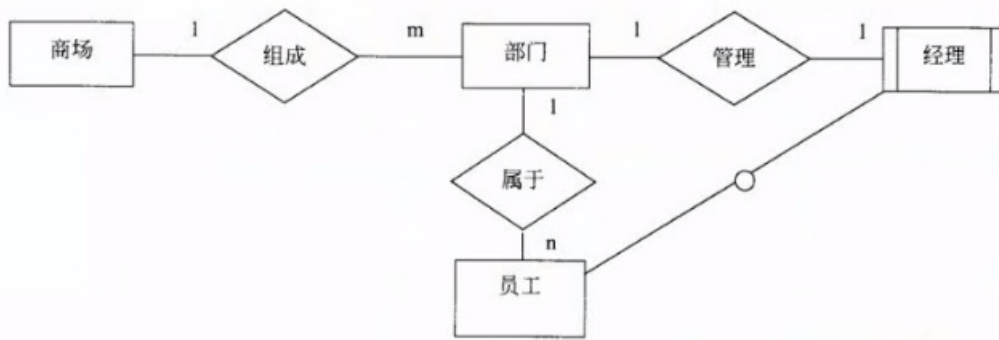
【关系模式设计】
商场（商场编号，商场名称，地址，联系电话）
部门（部门编号，部门名称，位置分布，联系电话，（a））
员工（员工编号，员工姓名，岗位，电话号码，工资，（b））
经理（（c），任职时间）

【问题1】（6分）
根据问题描述，补充四个联系，完善图12-2的实体联系图。联系名可用联系1、联系2、联系3和联系4代替，联系的类型分为1:1、1:n和m:n。

【问题2】（6分）
根据实体联系图，将关系模式中的空（a）~（c）补充完整，并分别给出部门、员工和经理关系模式的主键和外键。

【问题3】（3分）
为了使商场有紧急事务时能联系到轮休的员工，要求每位员工必须且只能登记一位紧急联系人的姓名和联系电话，不同的员工可以登记相同的紧急联系人。则在图12-2中还需添加的实体是（1），该实体和图12-2中的员工存在什么联系（填写联系类型）。给出该实体的关系模式。

解题
第一问: 完整的E-R模型如下图所示,补上了实体, 关系, 和联系的类型.注意员工和经理间的特殊化画法(连线中间的圆圈)



第二问: a 商场编号, b 部门编号, c 员工号

第三问: 紧急联系人, 一个员工只有1个紧急联系人.一个紧急联系人对应多个员工,紧急联系人:员工 1:n. 紧急联系人姓名,电话,员工号