《数据库系统原理》各章课外作业题目与答案，许卓明提供。【严禁外传！】

第1章

【主教材Page 11：习题1中的题3】

3. 试区分下列名词：

（1）数据模型与数据模式

（2）概念数据模型与概念数据模式

（3）数据库系统与数据库管理系统

参考答案：

（1）

数据模型（data model）是用来描述数据的一组概念和定义，这种描述包括三个要素/两种特性：数据的结构和数据间的联系/数据的静态特性；对数据的约束/数据的静态特性；数据上的操作/数据的动态特性。

数据模式（data schema）是运用某种数据模型（手段）对一个企业（enterprise）/组织（organization）的一组数据的结构、联系和约束的描述（结果）。

（2）

概念数据模型（conceptual data model）是面向现实世界和用户的、与DBMS无关的数据模型，用来描述一个企业/组织的概念化结构。

概念数据模式（conceptual data schema）更普遍的名称叫概念模式（conceptual schema），是用逻辑数据模型（logical data model）对一个企业/组织的一组数据的结构、联系和约束的描述。

（3）

数据库系统（database system）：由数据库（DB）、数据库管理系统（DBMS）、数据库应用程序和创建、维护与使用数据库的人（people）所组成的系统。

数据库管理系统（Database Management System, DBMS）是管理数据库的专用软件，是数据库系统的核心，对数据库系统的功能和性能有决定性影响。

第2章

【主教材Page 54：习题2中的题7(1)－(4), (6)】

7. 关系如图2-24所示，试求下列结果：

（1）П3,4 (*R*) ∪ *S*

（2）П3,4 (*R*) − *S*

（3）б*R*.1=a2 (*R*)

（4）*R* ⋈c *S*, c = (*R*.3=*S*.1) AND (*R*.4=*S*.2)

（6）(П1,2 (*R*) × *S*）− *R*

参考答案：

（1）(c1, d1); (c2, d2); (c3, d3)

（2）(c3, d3)

（3）(a2, b2, c1, d1); (a2, b2, c2, d2)

（4）(a2, b2, c1, d1); (a2, b2, c2, d2)

（6）(a1, b1, c1, d1); (a2, b2, c1, d1); (a3, b3, c1, d1);

(a1, b1, c2, d2); (a2, b2, c2, d2)

第3章

【主教材Page 86：习题3中的题2，但该题中只需做（1）、（3）、（4）、（5）、（7）】

2. 对例3-1中所定义的三个基表和例3-2所给的三个表的值，写出下列查询的SQL语句和查询的结果：

（1）查询身高大于1.80cm的男生和姓名；

（3）查询选修计算机系秋季所开课程的男生的姓名、课程号、学分数、成绩；

（4）查询至少选修一门电机系课程的女生的姓名；

（5）查询每位学生已修课程的门数和总平均成绩；

（7）查询所有课程的成绩都在80分以上的学生的姓名、学号，并按学号升序排列。

参考答案：

基表（详细定义见主教材Page 59例3-1）：

student (sno, sname, sex, bdate, height)

course (cno, lhour, credit, semester)

sc (sno, cno, grade)

（1）

SELECT sno, sname FROM student

WHERE height>1.8 AND sex = ‘男’;

（3）

SELECT sc.cno, credit, sname, grade FROM sc, course, student

WHERE course.cno = sc.cno AND student.sno = sc.sno AND

sc.cno LIKE ‘cs%’ AND semester = ‘秋’ AND sex = ‘男’

ORDER BY sc.cno;

（4）

SELECT sname FROM student

WHERE sex = ‘女’ AND

EXISTS ( SELECT \* //相关子查询

FROM sc WHERE sc.sno = student.sno AND sc.cno LIKE ‘ee%’ );

或者

SELECT sname FROM student

WHERE sex = ‘女’ AND

sno IN ( SELECT DISTINCT sno //不相关子查询

FROM sc WHERE cno LIKE ‘ee%’ );

或者

用“连接”（不好！）：

SELECT DISTINCT sname FROM student, sc

WHERE student.sno = sc.sno AND sex = ‘女’ AND cno LIKE ‘ee%’;

注：此处的DISTINCT，若加，则会丢失其他同姓名而又均选修了电机系课程的学生；若不加，则选多门电机系课程的某学生的姓名重复出现。因此，用“连接”的方法不好！

（5）

SELECT sc.sno, sname, COUNT(sc.cno), AVG(sc.grade) FROM sc, student

WHERE student.sno = sc.sno

GROUP BY sc.sno;

注：标准SQL或Oracle 7中SQL的SELECT列表仅可包含下列表达式类型：(1)常数；(2)组函数（也称聚集函数）；(3)与GROUP BY子句中相同的列表达式；(4)由GROUP BY子句中的列表达式所组成的表达式。因此，上述查询中不能再连接student表而获得sname。

（7）

SELECT sname, sno FROM student

WHERE 80 < ALL ( SELECT grade FROM sc //相关子查询

WHERE sc.sno = student.sno AND grade IS NOT NULL)

ORDER BY sno;

注：具体的RDBMS对NULL值的处理可能不同。

【主教材Page 86：习题3中的题3】

3. 在STUDENT和SC关系中，删去SNO以“01”开头的所有记录。

参考答案：

DELETE FROM student WHERE sno LIKE ‘01%’;

DELETE FROM sc WHERE sno LIKE ‘01%’;

【主教材Page 86：习题3中的题4】

4. 在STUDENT关系中，增加以下记录：

<0409101 何平 女 1987-03-02 1.62>

<0408130 向阳 男 1986-12-11 1.75>

参考答案：

INSERT INTO student VALUES (‘0409101’, ‘何平’, ‘女’, 1987-03-02, 1.62)；

INSERT INTO student VALUES (‘0408130’, ‘向阳’, ‘男’, 1986-12-11, 1.75)；

【主教材Page 86：习题3中的题5】

5. 将课程CS-221的学分数增为3，讲课时数增为60。

参考答案：

UPDATE course SET credit = 3, lhour = 60 WHERE cno= ‘CS-221’;

第4章

【补充作业】

1. 名词解释：

（1）事务及其ACID性质；

（2）三层体系结构；

（3）集中式数据库与分布式数据库；

（4）数据字典。

参考答案：

（1）

事务（transaction）是DBMS的（最小、完整的）执行单位，它由有限的数据库操作序列（如SQL语句）所组成，传统数据库理论要求事务必须满足ACID性质（ACID properties），即：A——执行的原子性；C——更新的一致性；I——彼此的隔离性；D——作用的持久性。

事务的ACID性质：

执行的原子性（**A**tomicity）：在一个事务中，所有的数据库操作是一个不可分割的操作序列，事务中的操作要么全做，要么全不做（nothing or all）。

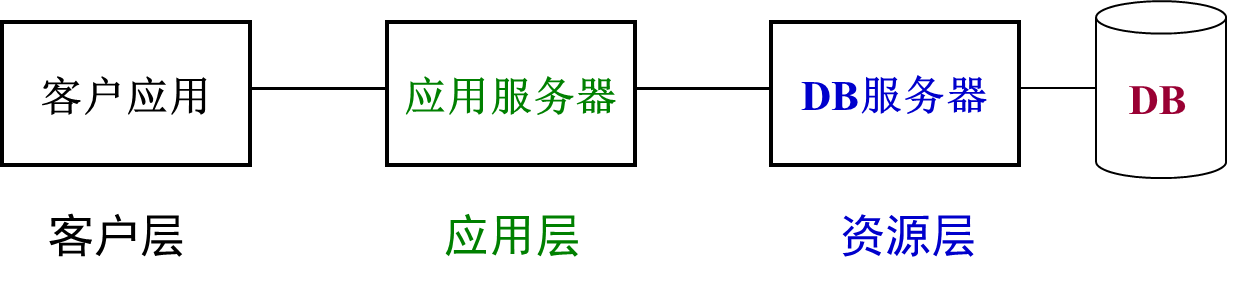
更新的一致性（**C**onsistency）：事务在功能上必须使数据库从当前的一致状态（consistent state）变成下一个一致状态。所谓一致状态就是数据库中的数据必须满足已定义的完整性约束（包括业务规则）。

彼此的隔离性（**I**solation）：多个事务并发执行时彼此不受影响，就好象各个事务独立执行一样。

作用的持久性（**D**urability）：事务一旦成功执行，其对数据库的影响是持久的，即使数据库发生故障也应能够恢复，即：维持这个事务的执行结果。

（2）

三层体系结构 （three-tier architecture）是对传统的Client/Server（C/S）结构的一种改进结构，目的是将任务合理分配到Client端和Server端来实现，降低了系统的通讯开销。后端分为应用服务器（application server）和数据库服务器（运行数据库系统）；前端客户机中不包含任何直接的数据库调用，而是通过表单界面与应用服务器进行通信，应用服务器再与数据库服务器进行通信以访问数据，应用程序的业务逻辑（business logic）被嵌入到应用服务器中而不是分布在多个客户机上。



（3）

集中式数据库（centralized database）是指数据集中存储，由DBMS集中管理的数据库系统。

分布式数据库（distributed database）是指数据分布存储但逻辑上相互关联的数据库系统，传统上认为应由分布式DBMS（DDBMS）统一管理。分布式数据库系统有两类结构：

1）物理上分布、逻辑上集中的分布式数据库系统：数据库中的数据在逻辑上是一个整体，但物理地分布在计算机网络的不同结点上。网络中的每个结点都可以独立处理本地数据库中的数据，执行局部应用。同时也可以同时存取和处理多个异地数据库中的数据，执行全局应用。这类分布式数据库系统的特点是：有全局数据模式；强调统一管理

2）物理上分布、逻辑上分布的分布式数据库系统，也称联邦式系统（federated system）。这类分布式数据库系统的特点是：无全局数据模式；强调结点自治。

（4）

数据字典（data dictionary, DD）也称数据目录（data catalog），更普遍地称为系统目录（system catalog），用于存储元数据（metadata），即关于数据的数据，包括数据库中的各种模式对象的定义、完整性约束（integrity constraints）、存储信息（storage information）和用户信息（user information）等。

第5章

【补充作业】

1. 试解释关系数据库系统中基表的四种典型存储机制（方法）：(1) 普通表；(2) 索引的表；(3) 索引簇表；(4) 散列簇表。

参考答案：

（1）

普通表就是没有建立索引、散列或簇集的基表。

（2）

索引的表是建立了索引的基表。索引（indexing）是与表（或簇集）相关的一种可选存储机制，它通过一棵有序树（如B树）将索引键（index key）的值与该值所对应的数据记录的物理地址（简称：记录地址）或地址集建立联系，以提高数据查找速度。

（3）

簇集（cluster）是存储表数据的一种可选方法。一个簇集是一个（组）表，其中具有同一公共列（组）值的所有表行（即元组）均存储在一起（即物理上相同或相邻的数据块中）。这些公共列（组）称簇集键（cluster key）。

索引簇集（Indexed Cluster）就是对簇表在簇集键上再建索引，每个簇集键值有一个索引项。

（4）

散列（Hash）是与表（或簇集）相关的一种可选存储机制，由于可通过一个Hash函数将散列键（Hash key）的值映射成一个物理块（数据块）的地址，因此，给出散列键的值Ki后立即可通过hash(Ki)得到其对应的物理块地址，从而可明显改进数据检索的性能。

散列簇集（Hash Cluster）就是对簇表的行在簇集键列上运用Hash函数进行散列，以决定相应物理块的地址。这样，具有同一散列值的行将存储在一起。

第6章

【补充作业】

1. 将主教材Page 64中例3-4查询语句的查询条件中去掉YEAR(BDATE)=1986，然后用语法树表示该修改后查询语句的代数优化过程【提示：这是连接3个基表的复杂查询，需运用“先做小关系间的连接／笛卡尔积”的变换策略。请参考主教材例6-1作答。】

参考答案：

为了简化，主教材例3-4中三个基表关系模式简化表示为：

s (sno, sname, sex, bdate, height)

c (cno, lhour, credit, seme)

sc (sno, cno, grade)

待优化的SQL查询为：

SELECT sname, c.cno, grade FROM s, c, sc

WHERE s.sno = sc.sno AND sc.cno = c.cno AND seme = ‘秋’;

该查询可表示成如下关系代数操作表达式，也可表示成一颗原始语法树：

∏sname, c.cno, grade(δt ((s×sc)×c))，其中：t = s.sno = sc.sno AND sc.cno = c.cno AND seme = ‘秋’

该查询的代数优化（即：语法树等价变换）过程如下：

优化后的查询可表示成如下关系代数操作表达式：

∏sname, c.cno, grade ((∏sno, sname s) ⋈ s.sno = sc.sno(∏cno (δseme=‘秋’c) ⋈ c.cno=sc.cno sc)))

先做δ，后做×

合并×与δ，成为⋈

(原始语法树)

∏sname, c.cno, grade

δt

×

c

s

sc

×

∏sname, c.cno, grade

δs.sno = sc.sno AND sc.cno = c.cno

×

δseme = ‘秋’

s

sc

×

c

∏sname, c.cno, grade

sc.cno = c.cno

s.sno = sc.sno

δseme = ‘秋’

s

sc

c

∏sname, c.cno, grade

sc.cno = c.cno

s.sno = sc.sno

δseme = ‘秋’

s

sc

c

加必要的投影，消除无用属性

∏sname, c.cno, grade

sc.cno = c.cno

s.sno = sc.sno

δseme = ‘秋’

s

sc

c

∏sno, sname

∏cno

(优化后语法树)

先做小⋈，后做大⋈

设|δseme = ‘秋’ c|<<|s|<|sc|

第7章

【主教材Page 169：习题7中的第6题】

6. 试述各种失效情况下应采取的恢复措施。

参考答案：

**事务故障（transaction failure）**是指事务因不可预知的原因而夭折，其特征是发生在事务提交完成前。**事务故障的恢复措施/步骤包括：**如果需要，就进行undo操作（因为前像（BI）已写入日志）；从活动事务表（ATL）中删除故障事务的ID，释放事务所占的资源。

**系统故障（system failure）**是指掉电或系统崩溃（system crash）而导致数据库系统无法继续正常运行下去，此时必须重新启动，从而导致一切事务均无条件中止运行，其特征是内存数据全部丢失，但外存上的数据库未遭破坏。**系统故障的恢复**，首先要使系统恢复正常运行；其次在故障发生时，可能有已提交事务的更新丢失了，可能有未提交的事务夭折了，**故恢复措施/步骤包括：**重新启动OS和DBMS；利用日志中的前像（BI）对未提交的事务进行undo操作（称向后恢复），利用日志中的后像（AI）对已提交的事务进行redo操作（称向前恢复）

**介质故障（disk/media failure）**是指数据库存储介质（通常是磁盘）发生故障而导致不可读/写盘或盘中数据丢失，其特征是外存上的数据库已遭破坏，一切已提交的事务对数据库的影响全部丢失。**介质故障的恢复**必须首先利用后备副本（backup）来恢复数据库，然后，再利用日志（log）进行向前恢复，**故恢复措施/步骤包括：**修复系统，必要时更换磁盘；重启系统OS和DBMS；加载最近后备（backup）；利用日志中的后像（AI）重做（redo）自该backup后的所有已提交事务的更新操作。

【主教材Page 169：习题7中的第7题】

7. 试回答下列有关数据库系统并发控制的技术问题。

（1）什么叫并发？

（2）为什么要并发运行？

（3）并发会引起什么问题？

（4）什么样的并发执行才是正确的？

（5）如何避免并发所引起的问题？

参考答案：

（1）

并发（concurrency）是指对数据库的并发访问（concurrent access），也称事务的并发运行，即：DBMS同时接纳多个事务，事务在时间上重叠地被执行。

（2）

为了改善数据库系统的性能，提高资源利用率，数据库中的事务总是并发运行的。

（3）

并发会引起的数据库的不一致问题，包括：丢失更新（lost update）；读脏数据（dirty read）；读值不可复现（unrepeatable read）。

（4）

并发运行事务的可串行化调度（serializable schedule）能保持数据库的一致状态，因此，一般DBMS都以“可串行化调度”作为并发控制的正确性准则。在实际系统中，通过让事务遵守“加锁协议”（locking protocol）的方法来确保并发控制遵循正确性准则。

（5）

一旦引入“加锁”机制，就会面临死锁（dead lock）问题，因此产生了如何检测以及如何预防死锁的问题。

对付死锁通常有两种办法：1）检测死锁的发生，发现后处理死锁；2）防止死锁发生。

【主教材Page 169：习题7中的第8题】

8. 什么是两段封锁协议？怎样实现两段封锁协议？

参考答案：

一个事务如果遵守“先加锁，后操作”的原则，则此事务称为合式事务（well-formed transaction）。合式事务是保证并发事务正确执行的基本条件。在一个事务中，如果所有加锁动作都在所有释放锁动作之前，则称这样的事务称为两段事务（two-phase transaction）。上述加锁限制称为两段封锁协议（Two-Phase Locking Protocol, 2PL协议）。如果所有事务都是合式、两段事务，则它们的任何调度都是可串行化的。

第8章

【主教材Page 182：习题8的第1题】

1. 对于例3-1所定义的三个关系，如果DBA要对用户U1,U2,U3,U4分别规定下列权限，试用SQL写出授权语句。

（1）U1只能读STUDENT关系中除HEIGHT以外的属性；

（2）U2可以读、增、删SC关系，但不得修改SC关系，并且可把读SC关系的特权转授给其他用户；

（3）U3可以读、增、删COURSE关系，且可修改除LHOUR和CREDIT以外的属性，但无权转授这些特权；

（4）授予数据库新用户U4 RESOURCE特权，并规定其口令为newuser。

参考答案：

例3-1所定义的三个关系如下：

student (sno, sname, sex, bdate, height)

course (cno, lhour, credit, semester)

sc (sno, cno, grade)

（1）

GRANT select(sno, sname, sex, bdate) ON student TO U1;

（2）

GRANT select, insert, delete ON sc TO U2;

GRANT select ON sc TO U2 WITH ADMIN OPTION;

（3）

GRANT select, insert, delete, update(cno, semester) ON course TO U3;

（4）

GRANT RESOURCE to U4 IDENTIFIED BY newuser;

【主教材Page 182：习题8的第2题】

2. 在例3-1的三个关系的定义中，指出哪些是域完整性约束？哪些是实体完整性约束？哪些是引用完整性约束？在DBMS中，何时、如何实现这些隐含完整性约束检查？当发现违反这些完整性约束时，DBMS将分别采取何种措施？

参考答案：

域完整性约束：student中定义的4个NOT NULL；course中定义的4个NOT NULL；sc中定义的2个NOT NULL。

实体完整性约束：student中定义的PRIMARY KEY (SNO)；course中定义的PRIMARY KEY (CNO)；sc中定义的PRIMARY KEY (SNO, CNO)。

引用完整性约束：sc中定义的FOREIGN KEY (SNO)和FOREIGN KEY (CNO)。

隐含完整性约束（implicit constraints）是指隐含在数据模式中的约束，用DDL来申明，约束定义存于数据字典中。对关系数据库，包括：域完整性约束（domain integrity constraints），实体完整性约束（entity integrity constraints），引用完整性约束（referential integrity constraints）。每当数据库发生更新时（包括数据的初始载入），DBMS对隐含完整性约束进行检查，一旦发现违反完整性约束的数据库操作，就按照预定义的策略做出相应反应，以维护数据库的完整性，具体包括：

定义了NOT NULL的属性值不允许取null值；每个PRIMARY KEY的键值应唯一，且不能取null值；定义了ON DELETE CASCADE的外键实施级联删除；定义了ON DELETE RESTRICT的外键实施禁止删除。

第9章

【补充作业】

1. 职员（emp）基表定义同课件：emp (empno, ename, job, mgr, sal, comm, deptno)。试用SQL:1999/SQL3语法（即课件中介绍的SQL语法）定义一个名为empBandh的触发器，来实现以下功能：一旦某个员工的数据从emp表中被删除，只要此员工的工种不是“bandh”（停职），就在emp表中恢复（即重新插入）此员工的数据，并将其工种（job）置为“bandh”，月薪（sal）置为2000.0，佣金（comm）置为null，其余属性值均不变。

参考答案：

注：职员（emp）基表定义：emp (empno, ename, job, mgr, hiredate, sal, comm, deptno)。

触发器定义如下：

CREATE TRIGGER empBandh

AFTER DELETE ON emp //事件

REFERENCING

OLD ROW AS oldrow,

NEW ROW AS newrow

FOR EACH ROW

WHEN oldrow.job <> “bandh” //条件

BEGIN //动作：通常不管有几条SQL语句，均可以用begin...end界定

INSERT INTO emp

VALUES (oldrow.empno, oldrow.ename, “bandh”, oldrow.mgr, 2000.0, NULL, oldrow.deptno)

END;

第10章

【主教材Page 221：习题10中的第8题(1) - (7)】

8. 试回答下列关系模式最高属于第几范式？并解释其原因。

（1）R(ABCD), F = {B→D, AB→C}。

（2）R(ABCDE), F = {AB→CE, E→AB, C→D}。

（3）R(ABCD), F = {B→D, D→B, AB→C}。

（4）R(ABC), F = {A→B, B→A, A→C}。

（5）R(ABC), F = {A→B, B→A, C→A}。

（6）R(ABCD), F = {A→C, D→B}。

（7）R(ABCD), F = {A→C, CD→B }。

【补充作业】进一步将上述(7)中的关系模式R无损分解到BCNF，并说明此分解不能保持依赖的理由。

参考答案：

（1）R(ABCD), F={B→D, AB→C};

KEY={A, B}=AB 【注：大多数教材中这样来简化表示属性集】

非平凡函数依赖B→D中：

决定子B不是超键，所以R不属于BCNF；

被决定子D不是主属性，所以R不属于3NF；

而且，由B→D知：D部分依赖于KEY，所以R不属于2NF；

结论：R属于1NF。

（2）R(ABCDE), F={AB→CE, E→AB, C→D};

KEY=AB或E 【注：该关系模式有两个键】

非平凡函数依赖C→D中：

决定子C不是超键，所以R不属于BCNF；

被决定子D不是主属性，所以R不属于3NF；

F中不存在非主属性部分依赖于KEY；

结论：R属于2NF。

（3）R(ABCD), F={B→D, D→B, AB→ C };

KEY=AB或AD 【因为：由D→B知AD→AB，故AD也是KEY】

非平凡函数依赖B→D中：

决定子B不是超键，所以R不属于BCNF；

但被决定子D是主属性；

非平凡函数依赖D→B中被决定子B是主属性；

非平凡函数依赖AB→C中决定子AB是超键；

结论：R属于3NF。

（4）R(ABC), F={A→B, B→A, A→C};

KEY=A或B

非平凡依赖A→B中决定子A是超键；

非平凡依赖B→A中决定子B是超键；

非平凡依赖A→C中决定子A是超键；

由B→A, A→C有B→C，其中决定子B是超键；

即找不到任何冒犯BCNF的非平凡依赖；

结论：R属于BCNF。

（5）R(ABC), F={A→B, B→A, C→A};

KEY=C

非平凡依赖A→B中：

决定子A不是超键，所以R不属于BCNF；

被决定子B又不是主属性，所以R不属于3NF；

而且，F中不存在非主属性部分依赖于KEY；

结论：R属于2NF。

（6）R(ABCD), F={A→C, D→B};

KEY=AD

非平凡依赖A→C中：

决定子A不是超键，所以R不属于BCNF；

被决定子C又不是主属性，所以R不属于3NF；

非主属性C（及B）部分依赖于KEY，所以R不属于2NF；

结论：R属于1NF。

（7）R(ABCD), F={A→C, CD→B};

KEY=AD

非平凡依赖A→C中：

决定子A不是超键，所以R不属于BCNF；

被决定子C又不是主属性，所以R不属于3NF；

因A→C，即非主属性C部分依赖于KEY，所以R不属于2NF；

结论：R属于1NF。

【补充】：将R无损分解到BCNF。

对A→C，其冒犯BCNF条件，运用“细胞分裂法”将R无损分解成：

R1(AC) 和 R2(ABD)。

因为R1(AC)上仅有函数依赖A→C，故R1(AC)的函数依赖集F1={A→C}，于是R1的KEY=A（注：A也是超键），所以R1属于BCNF。

又因为函数依赖A→C和CD→B逻辑蕴含AD→B，即：{A→C, CD→B} ⊨ AD→B。因此，R2(ABD) 的函数依赖集F2={AD→B}，于是R2的KEY＝AD（注：AD也是超键），所以R2属于BCNF。

但是，F1∪F2={A→C}∪{AD→B} ⊭ CD→B，而F={A→C, CD→B}，于是：F1∪F2 ⊭ F，所以ρ={R1(AC), R2(ABD)}不是一个对R的保持依赖的分解。【注：⊨表示逻辑蕴含；⊭表示不逻辑蕴含】

所以说：一个关系模式可以找到分解到BCNF的一个无损分解，但不一定能找到分解到BCNF的一个保持依赖的分解。

第11章

【补充作业】

1. 试对四个实体：学生、班级、课程、教师，以及它们之间的四个联系：学生与班级之间的组成、学生与课程之间的选课（有成绩属性）、教师与课程之间的任课、教师与班级之间的班导师（有任职年度属性）进行E-R建模（自己决定合适的实体属性和联系语义），画出E-R图，然后把E-R图转换成关系数据库模式，并写出创建相应基表的SQL DDL语句，同时定义PK与FK、以及你认为必要的其他完整性约束。

参考答案：

E-R图：

学号

姓名

性别

入学年度

编号

名称

学分数

编号

所属院系

职工号

姓名

职称

学生

课程

选课

组成

班级

班主任

任课

教师

成绩

N

M

N

1

N

1

N

1

任职年度

从E-R图转换成的关系数据库模式：

班级（编号，所属院系，班主任职工号，任职年度）；

学生（学号，姓名，性别，入学年度，所在班级编号）；

课程（编号，名称，学分数，任课教师职工号）；

选课（学号，课程编号，成绩）；

教师（职工号，姓名，职称）；

（完）