

示例中出现的模式 S、C、SC 和 D 分别表示学生、课程、学生选课情况和院系。

S(SNO,SNAME,DNO,SEX, AGE); C(CNO,CNAME,CREDIT);

SC(SNO,CNO,SCORE); D(DNO,DNAME)。

其属性分别表示如下：SNO—学生编号，SNAME—学生姓名，DNO—院系编号，SEX—性别，AGE—年龄，CNO—课程编号，CNAME—课程名称，CREDIT—课程学分，SCORE—成绩，DNAME—院系名称。

➤ 第二章

1、查询王宏同学选修的所有课程名称

分析：本题目主要考察学生对于自然连接的理解，题目的已知条件是学生姓名，查询的是课程名称，因此需要三个表做自然连接。

$$\Pi_{\text{cname}}(\sigma_{\text{sname}='王宏'}(S) \bowtie SC \bowtie C)$$

2、查询至少选修了 c1 和 c2 号课程的学生号。

分析：本题目主要考察学生对于“包含”这一类查询的关系代数表达式，加强学生对于除运算的理解，另外本题目可以用下列两个表达式，让学生判断和选择，

表达式一：

$$\Pi_{\text{sno}, \text{cno}}(SC) \div \sigma_{\text{cno}='c1' \vee \text{cno}='c2'}(C)$$

表达式二：

$$\Pi_{\text{sno}}(SC \div \sigma_{\text{cno}='c1' \vee \text{cno}='c2'}(C))$$
（本表达式错误）

表达式二的错误在于，SC 关系中的属性 score 会影响除运算的计算结果，属性 score 在本题目中是不需要的，应该通过投影运算，现将去去除，然后再进行除运算。表达式二的错误分析，让学生理解除运算中，属性之间的关系。

3、查询缺考的学生的学号和姓名

分析：本题目考察学生对于 null 的理解，以及对于 null 的判断。缺考的学生在属性 score 上的取值为 null。

$$\Pi_{\text{sno}, \text{sname}}(\sigma_{\text{score is null}}(SC \bowtie S))$$

➤ 第三章

1、查询平均成绩最高的学生学号。

分析：子查询的作用是查询所有学生的平均成绩，父查询中，用每一个学生的平均成绩，和全体学生的平均成绩进行比较，如果某一个学生的平均成绩大于等于全体学生的平均成绩，这个学生的平均成绩最高。

```
select    sno

from      SC

group by  sno

having    avg(score) >= all

                (select    avg(score)

                  from      SC

                  group by  sno)
```

2、查询选修了全部课程的学生姓名

分析：选修了全部课程意味着：任意课程，所查询学生选之，即不存在任何一门课程，所查询学生没有选。

```
select    sname

from      S

where     not exists

                (select    cno

                  from      C

                  where     not exists

                        (select    *

                          from      SC

                          where     sc.cno = c.cno

                        and        sc.sno = s.sno ))
```

3、查询数学成绩比王红同学数学成绩高的学生姓名

姓名	课程	成绩
刘洪	物理	93
王红	数学	86
张军	数学	89

分析：将表 R 通过重命名，得到两个相同的表，将 a 看做是王红同学的数学成绩，b 是需要查询的学生，a 和 b 之间的关系式本题目需要关注的重点。

```
select b.sname
from R a, R b
where a.sname='王红' and a.cname='数学' and
      b.cname='数学' and b.score>a.score
```

➤ 第六章

一个工厂有若干仓库；每一仓库有若干职工作为仓库管理员，职工之间有领导与被领导的关系；仓库中保存工厂生产的多种零件。用 E-R 图表示上述内容，关注仓库面积、仓库中保存零件的种类、每种零件的入库时间及入库数量，职工的姓名、职称、职务及工资待遇，零件的颜色、成本及出厂价。并将 E-R 图转换成相应的关系模型

分析：1、确定实体及其属性

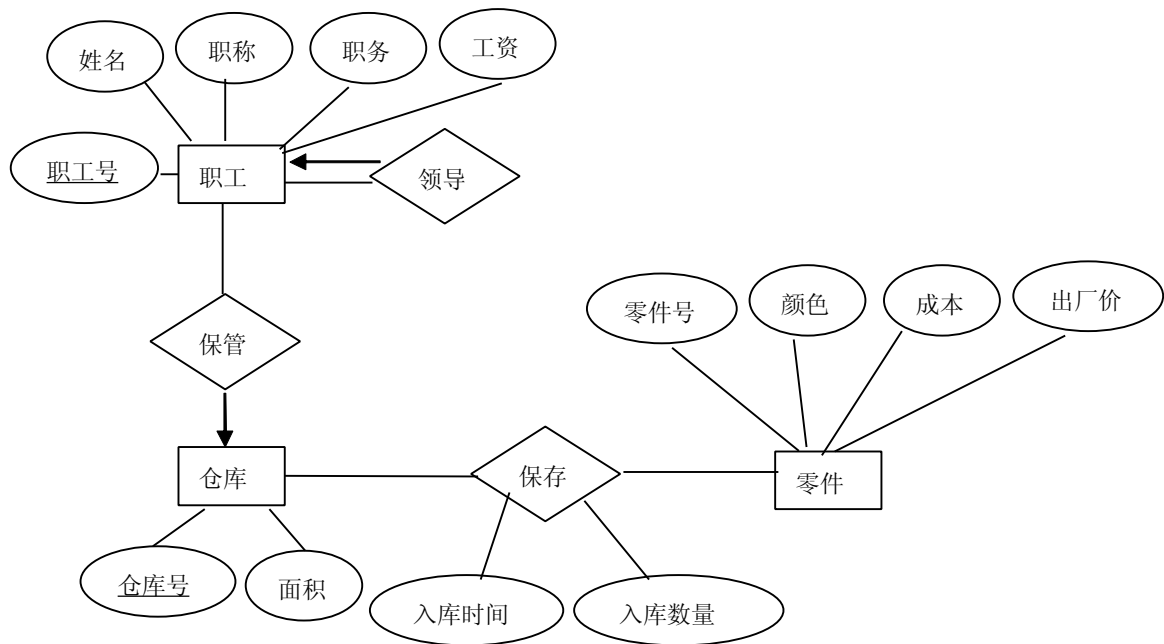
从上述需求文档中，我们首先分析实体，实体应该包括：仓库、职工和零件，其中仓库的属性包括仓库编号和面积，职工的属性包括：姓名、职称、职务和工资待遇，零件的属性包括：颜色、成本和出厂价。

2、确定实体之间的联系

通过分析需求文档，我们知道职工和职工之间有领导联系，职工和仓库之间有保管联系，零件和仓库之间有保存联系，其中保存联系还有自己的属性，包括：零件的入库时间和入库数量。

“每一仓库有若干职工作为仓库管理员”意味着职工和仓库之间的保管联系是二元一对多的联系，“职工之间有领导与被领导的关系”意味着职工之间的领导关系一元一对多联系，“仓库中保存工厂生产的多种零件”意味着仓库和零件之间的联系是二元多对多的联系。

综上，ER 图如下：



3、从 ER 图向关系模式的转变

从 ER 图中可以看出，领导和保管都是二元一对多的联系，可以将其与实体职工转换的关系模式合并，联系保存是二元多对多联系，必须独立成为一个关系模式。最终得到的关系模式如下：

职工(职工号、姓名、职称、职务、工资、领导、仓库号)

仓库(仓库号、面积)

零件(零件号、颜色、成本、出厂价)

保存(仓库号、零件号、入库时间、入库数量)

第七章

设有关系模式 $R(A, B, C, D)$ ，其上的函数依赖集为：

$F = \{A \rightarrow C, C \rightarrow A, B \rightarrow AC, D \rightarrow AC\}$

- (1) 计算 $(AD)^+$
- (2) 求 F 的最小等价依赖集 F_m
- (3) 求 R 的码
- (4) 将 R 分解使其满足 BCNF 且具有无损连接性
- (5) 将 R 分解成满足 3NF 并具有无损连接性与保持依赖性

解：(1) 令 $x = \{AD\}$, $x(0) = AD$, $x(1) = ACD$, $x(2) = ACD$, 故 $(AD)^+ = ACD$

(2) 将 F 中的依赖右部属性单一化：

$$F1 = \begin{bmatrix} A \rightarrow C \\ C \rightarrow A \\ B \rightarrow A \\ B \rightarrow C \\ D \rightarrow A \\ D \rightarrow C \end{bmatrix}$$

在 F1 中去掉多余的函数依赖:

$\because B \rightarrow A, A \rightarrow C \therefore B \rightarrow C$ 是多余的

又 $\because D \rightarrow A, A \rightarrow C \therefore D \rightarrow C$ 是多余的

$$F2 = \begin{bmatrix} A \rightarrow C \\ C \rightarrow A \\ B \rightarrow A \\ D \rightarrow A \end{bmatrix}$$

$\because F2$ 中所有依赖的左部都是单属性, 不存在依赖左部有多余的属性

\therefore

$$Fm = \begin{bmatrix} A \rightarrow C \\ C \rightarrow A \\ B \rightarrow A \\ D \rightarrow A \end{bmatrix}$$

函数依赖集的最小集不是惟一的, 本题还可以有其他答案。

(3) $\because B, D$ 在 F 中所有函数依赖的右部均未出现, \therefore 候选码中一定包含 BD, 而 $(BD)^+ = ABCD$, 因此, BD 是 R 惟一的候选码。

(4) 考虑 $A \rightarrow C$, $\because ABCD$ 不是 BCNF($A \rightarrow C$ 的左部不包含候选码 BD), 将 ABCD 分解为 AC 和 ABD。AC 已是 BCNF, 进一步分解 ABD, 选择 $B \rightarrow A$, 把 ABD 分解为 AB 和 BD。此时 AB 和 BD 均为 BCNF。

$\therefore \rho = \{AC, AB, BD\}$

(5) 由 (2) 可求出满足 3NF 的具有依赖保持性的分解为 $\rho = \{AC, BA, DA\}$ 。判断其无损连接性如下表所示, 由此可知 ρ 不具有无损连接性。令 $\rho = \rho \cup \{BD\}$, BD 是 R 的候选码,

$\therefore \rho = \{AC, BA, DA, BD\}$

R_i	A	B	C	D
AC	a_1		a_3	
BA	a_1	a_2	a_3	
DA	a_1		a_3	a_4

➤ 涉及“对所有的”的查询专题

本专题对于涉及“对所有的”的查询, 比如: 查询选择了学号是 s2 学生选修的所有课程的学生学号, 此类问题在关系代数中需要用到“除运算”, 在元组关系演算中需要用到“蕴含”, 在 SQL 中需要用到“exists”, 具体表达式如下:

1、关系代数

$$\Pi_{sno, cno}(SC) \div \Pi_{cno}(\sigma_{sno='s2'}(SC))$$

$$\Pi_{sno}(SC) - (\Pi_{sno}(SC) \times \Pi_{cno}(\sigma_{sno='s2'}(SC))) - \Pi_{sno, cno}(SC)$$

2、元组关系演算

$$\{t \mid (\exists x \in S(x[sno]=t[sno]) \wedge \forall u \in SC(u[sno]='s2' \Rightarrow \exists v \in SC(u[cno]=v[cno] \wedge t[sno]=v[sno])))\}$$

3、SQL

SELECT DISTINCT *sno*

FROM *sc scx*

WHERE NOT EXISTS

(SELECT *

FROM *sc scy*

WHERE *scy.sno = 's2' AND*

NOT EXISTS

(SELECT *

FROM *sc scz*

WHERE *scz.sno=scx.sno AND*

scz.cno=scy.cno))

SELECT DISTINCT *sno*

FROM *sc a*

WHERE NOT EXISTS

(SELECT *cno*

FROM *sc*

WHERE *sno = 's2'*

Except

Select *cno*

From *sc b*

Where *a.sno=b.sno)*

➤ 规范化过程示例

给定关系模式 S(sno , sname , dno ,dean , cno , score),

关系模式 S 的属性域都是原子的, 因此 S∈1NF, 但是其中的主码为(sno, cno), 函数依赖包括: sno →sname, sno →dno, sno →dean, dno→dean, (sno,cno) → score, 其中有非主属性对于主码的部分依赖, 因此 S∉1NF, 将 S 分解为 S1 和 S2, 使得 S1 和 S2 均为 2NF,

S1(sno , sname , dno ,dean), S2(sno , cno , score)

对于关系模式 S1 而言，主码为(sno)，其中的函数依赖包括：sno \rightarrow sname，sno \rightarrow dno，sno \rightarrow dean，dno \rightarrow dean，函数依赖 dno \rightarrow dean，不是平凡的函数依赖，决定因素不是超码，{dean} - {dno} = {dean}，也不包含在任何一个候选码中，因此，S1 \notin 3NF，将 S1 分解为 S11 和 S12，使得 S11 和 S12 均为 3NF。

S11(sno , sname , dno), S12(dno ,dean)