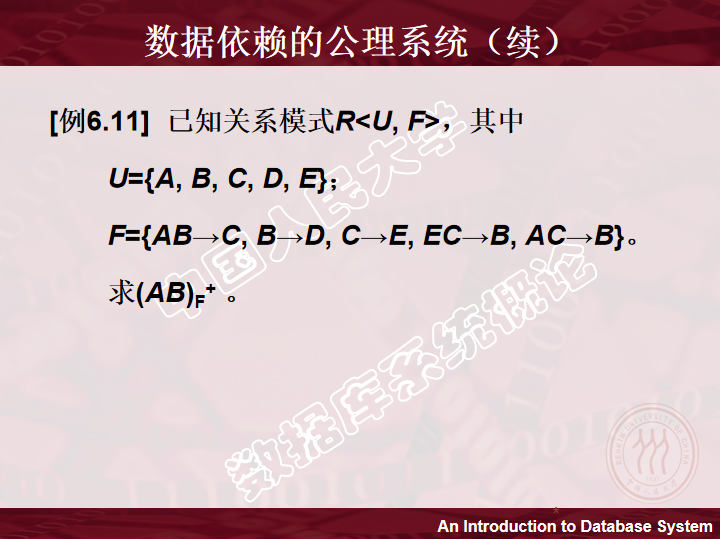
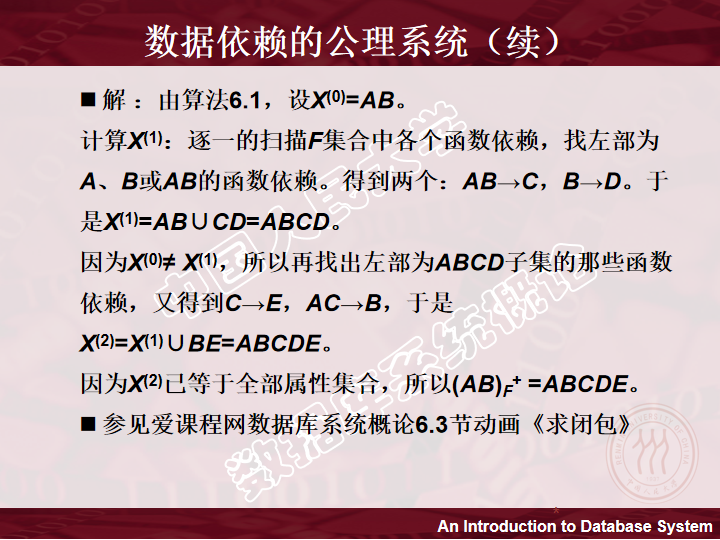
1.求属性集合X关于函数依赖集F的闭包X+





**注：闭包的标准写法是 (AB)F+**

1. **求最小函数依赖集算法：**

1）用分解法则，使F中的任何一个函数依赖的**右边仅含一个属性**

2）**去除多余的函数依赖**：从第一个函数依赖X→Y开始，将其从F中去掉，然后在剩下的函数依赖中求X的闭包X+，判断X+是否包含Y，如果包含，则去掉X→Y，否则保留。依次做下去，直到扫描所有函数依赖为止。

3）**去掉依赖左侧多余的冗余属性**：如XY→A，若要判断Y是冗余属性，则计算X+，判断A是否属于X+，如果属于，则Y是冗余属性，将其删除，同理判断X。   注意：若在第三步中修改了函数依赖集F，则需要重新做一次步骤二。

例：U=(A,B,C,D,E,G)

F={BG->C, BD->E, DG->C, ADG->**BC**, AG->B, B->D}，求F最小依赖集。

解：

**第一步：右边单一化。**拆分函数依赖右部，使F中的任何一个函数依赖的右边仅含一个属性。

F1={BG->C, BD->E, DG->C, **ADG->B, ADG->C**, AG->B,B->D}

**第二步：去除多余的函数依赖。**从第一个函数依赖X→Y开始，逐个尝试**将其从F中去掉**，然后在剩下的函数依赖中求X的闭包X+，判断X+是否包含Y，如果包含，则去掉X→Y，否则保留，依次做下去，直到扫描所有函数依赖为止。

在F1={BG->C, BD->E, DG->C, ADG->B, ADG->C, AG->B,B->D}中求解

1）尝试去掉BG->C：求(BG)+=B**C**DEG，包含右边属性C，所以去掉。

(BG)+ => BG =>BGCD=>BGCDE

(**注意**：求闭包的时候，不能使用用已经**去掉**的函数依赖)

{BD->E, DG->C, ADG->B, ADG->C, AG->B,B->D}

2）BD->E：(BD)+=BD，不包含右边E，所以不用去掉。

DG->C：(DG)+=DG，也不用去掉。

3）ADG->B：(ADG)+=U，包含B，所以去掉。

{BD->E, DG->C, ADG->C, AG->B,B->D}

4）ADG->C：(ADG)+包含C，去掉。{BD->E, DG->C, AG->B,B->D}

AG->B：(AG)+=AG，不用去掉。

B->D：(B)+=B，不用去掉。

所以F2={BD->E,DG->C,AG->B,B->D}

**第三步：去掉依赖左侧多余的冗余属性。**如XY→A，若要判断Y是冗余属性，则计算X+，判断A是否属于X+，如果属于，则Y是冗余属性，将其删除，同理判断X。

注意：若在第三步中修改了函数依赖集F，则需要重新做一次步骤二。

F2={BD->E,DG->C,AG->B,B->D}

1）BD->E：考查D，求(B)+=BDE，包含D，所以D冗余。

D->E，求(D)+=D，所以B不冗余。

用B->E代替BD->E。

2）DG->C：D->C，(D)+=D，所以G不冗余。

G->C，(G)+=G，所以D不冗余。

3）AG->B：A->B，(A)+=A，所以G不冗余。

G->B，(G)+=G，所以A不冗余。

4）Fmin={B->E,DG->C,AG->B,B->D}。

----------------------------- 题1 ------------------------------------------

设有关系模式：R（运动员编号，比赛项目，成绩，比赛类别，比赛主管）,存储运动员比赛成绩及比赛类别、主管等信息。

如果规定：每个运动员每参加一个比赛项目，只有一个成绩；每个比赛项目只属于一个比赛类别；每个比赛类别只有一个比赛主管。

试回答下列问题：

（1）试写出关系模式R的基本函数依赖和所有的候选码。

（2）说明R不是2NF模式的理由，并把R分解成2NF模式集。

（3）进而把R分解成3NF模式集，并说明理由。

有运动员参加比赛的关系模式R（运动员号码，比赛项目，成绩，项目类别，比赛主管），如果规定：每个运动员参加一个比赛项目只有一个比赛成绩，每个比赛项目只属于一个项目类别，每个项目类别只有一个比赛主管。

1．根据上述规定，写出模式R的基本FD和关键码。

2．R最高为第几范式？

3．将R规范为3NF。

**一、非标准解法**（第一节课解法，模式分解不按规范方法，凭对语义的理解）

1. 根据语义得出FD(函数依赖)：注意决定关系不要写反

1）每个运动员参加一个比赛项目只有一个比赛成绩：(运动员编号，比赛项目)->成绩

2) 每个比赛项目只属于一个项目类别：比赛项目->比赛类别

3) 每个项目类别只有一个比赛主管：比赛类别->比赛主管

候选码：运动员编号+比赛项目 PK

1）是非主属性对码完全函数依赖，2）、3）是对码的部分依赖

2、判别是否2NF

从2）函数依赖中看出，比赛类别对主码（运动员编号，比赛项目）存在部分依赖，因此R不是2NF，是1NF

投影分解成2NF：

R1(运动员编号，比赛项目，成绩)

R2(比赛项目，比赛类别，比赛主管)

3、判别是否3NF

R1已经是3NF，R2中存在传递依赖故不是3NF

R2分解成：

R21(比赛项目，比赛类别)

R22(比赛类别，比赛主管)

1. 规范为3NF：

R1(运动员编号，比赛项目，成绩) F1={(运动员编号，比赛项目)->成绩}

R2(比赛项目，比赛类别) F2={比赛项目->比赛类别}

R3(比赛类别，比赛主管) F3={比赛类别->比赛主管}

**二、规范分解法**

解：1、模式R的函数依赖关系

**分析语义，求FD集：**

1. 每个运动员参加一个比赛项目只有一个比赛成绩：**（运动员号码，比赛项目）->成绩**
2. 每个比赛项目只属于一个项目类别：**比赛项目->项目类别**

【注意：这里不能写反，一个比赛项目能唯一确定一个类别】

1. 每个项目类别只有一个比赛主管：**项目类别->比赛主管**

如果设关系模式R（运动员号码，比赛项目，成绩，项目类别，比赛主管）为R（A，B，C，D，E），则上述函数依赖可描述为F={AB->C，B->D，D->E} 【建议把属性符号化，不用汉字】

**确定主码**

猜测AB为候选码，求AB的**闭包**(AB) F+={C，D，E，A，B} =U，所以**R的关键码（primary key主码）**为AB，即（运动员号码，比赛项目），也唯一候选码

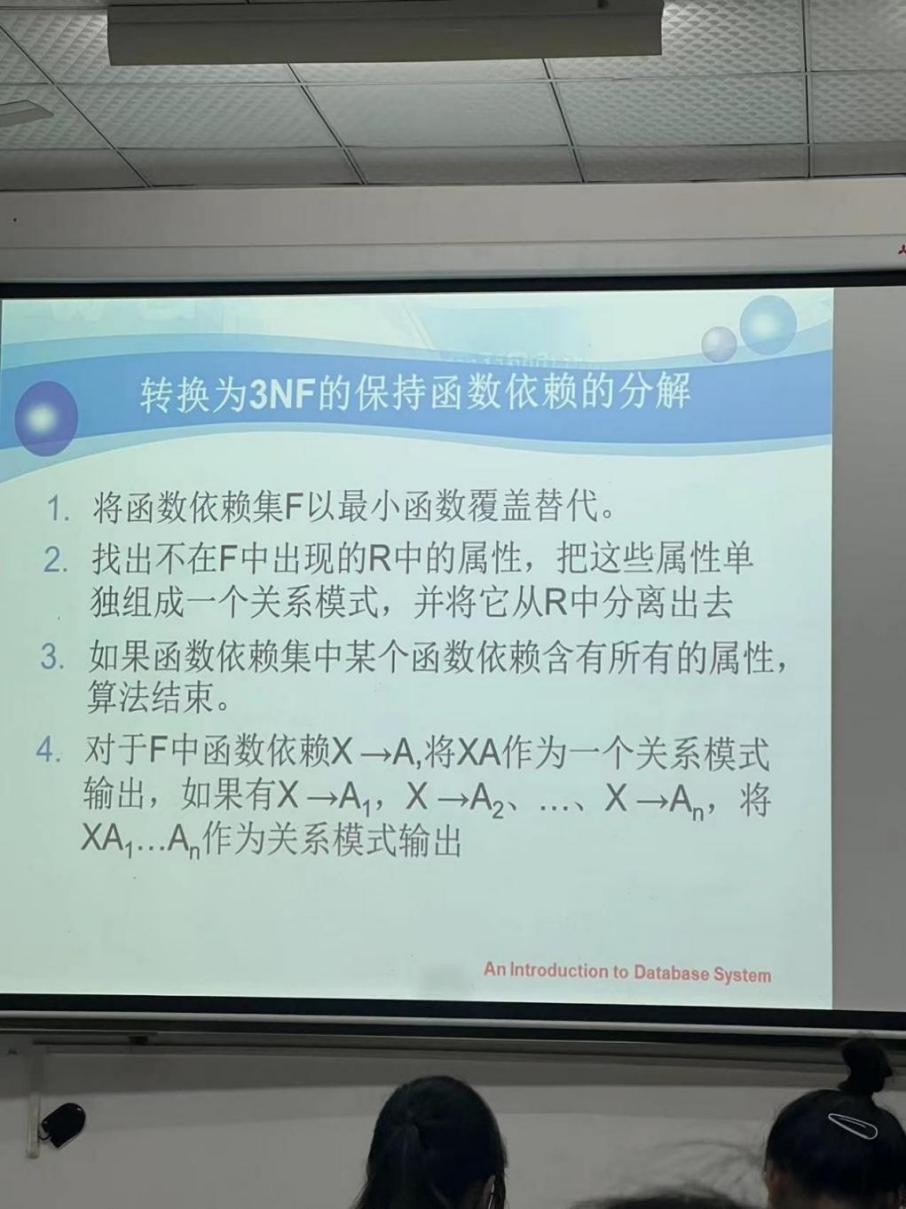
2、关系模式R中的主属性为A、B，非主属性C、D、E，因为存在非主属性D对主码AB的部分函数依赖B->D，所以R达不到2NF，最高为1NF。需要进行分解。

**求最小函数依赖：**

F={AB->C，B->D，D->E}

1. 所有函数依赖的右边均为单一属性；
2. F没有**冗余函数依赖（见前面求最小依赖集算法）**
3. A+={A}，B+={D,E}，说明C不属于A+和B+，即对A没有部分依赖，对B也没有依赖，C完全函数依赖于AB；所有函数依赖的左边没有**冗余属性（见前面求最小依赖集算法）**

综上，关系模式R的F已是**最小函数依赖**



算法6.3

1. 分解R为保持函数依赖的3NF【算法6.3】：

1）先对F极小化：F已是最小依赖集F={AB->C，B->D，D->E}

2）分离无关属性：R所有属性均出现在F中，无需分离。

3）分解：按**左部相同**的原则分组【依赖左部相同的所有属性分为一组】，得

R1(A，B，C)， R2(B，D)，R3(D，E)

4）每个关系分解的结束条件：此关系的某个函数依赖左右两端的属性构成关系属性全集。如R1(A，B，C) , AB->C包含U={A,B,C},因此不再进行分解；R2,R3同理

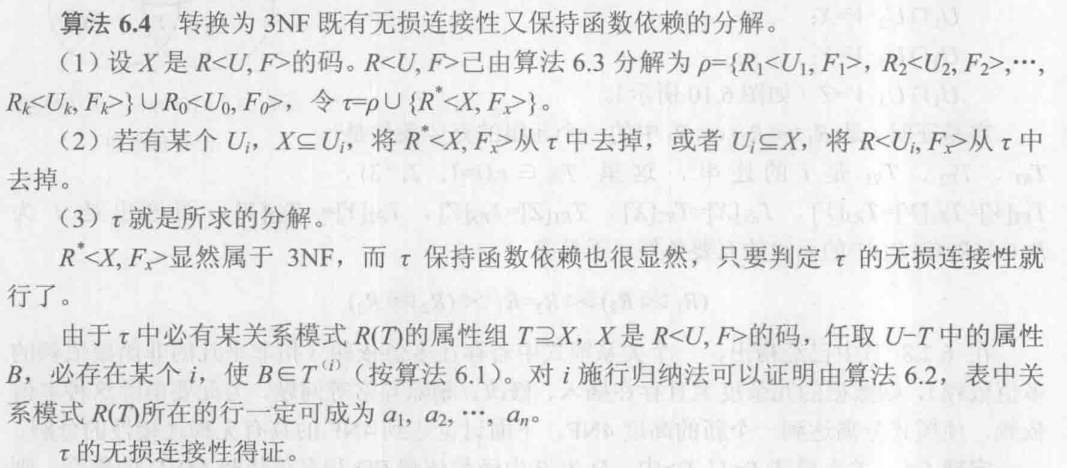
分解结果：

R1（运动员号码，比赛项目，成绩）

R2（比赛项目，项目类别）

R3（项目类别，比赛主管）

为一个**保持函数依赖且达到3NF**的分解。

6.4说明：

X 码

р 按算法6.3分解出的保持函数依赖的关系模型

R0 不在F中出现的属性，单独分离出来构成的关系模式

R\* 码属性与相关FD构成的关系模型

t р与R\*的并集

步骤(2) 去掉码关系模型与分解关系模型重复的部分

1. 既有无损连接性又保持函数依赖的3NF模式分解【算法6.4】

p={ R1(A，B，C){AB->C}， R2(B，D){B->D}，R3(D，E) {D->E}}

码 AB

R1的属性集U中包含码AB，因此p是所求模式分解

**题2**

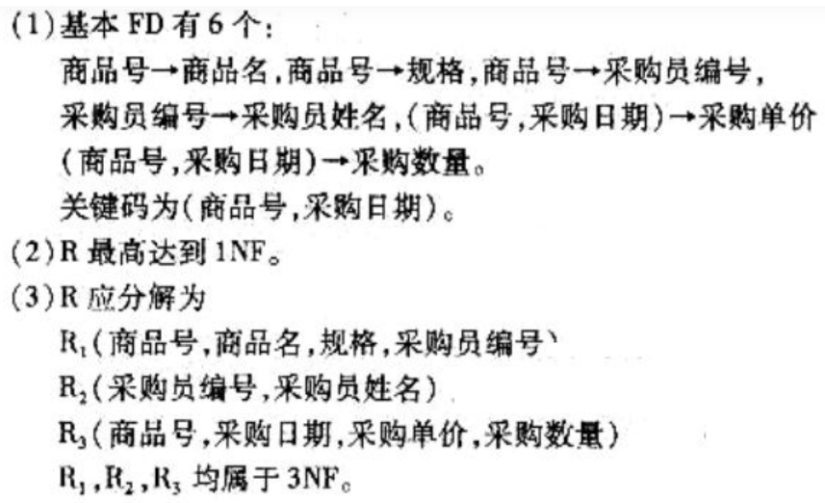
设有一个记录商品采购信息的关系模式：

R（商品号，商品名，规格，采购日期，采购单价，采购数量，采购员编号，采购员姓名）

如果规定：每名采购员可负责多种商品的采购，但每种商品只能由一个人员负责采购； 商品的采购单价随市场行情而定。

1. 根据上述规定，写出R的基本FD和关键码。
2. R最高达到第几范式。
3. 将R规范到3NF。

一、非标准解法

【百度答案】

解：

(1)R的基本FD：

商品号->商品名

商品号->规格

采购员编号->采购员姓名

每种商品只能由一个人员负责采购：商品号->采购员编号

商品的采购单价随市场行情而定：（商品号，采购日期）->(采购单价,采购数量）

关键码（主码）：（商品号，采购日期）

如果设关系模式R（商品号，商品名，规格，采购日期，采购单价，采购数量，采购员编号，采购员姓名）为R（A,B,C,D,E, F, G,H），则上述函数依赖可描述为

F={A->B，A->C, G->H, A->G, AD->EF} ；

主码为（A,D）

（2）R最高达到1NF。因为存在非主属性B,C,G对码的部份依赖

（3）规范到3NF

1）消除非主属性对码的部分依赖，分解为2NF

（A,D,E,F）

（A,B,C,G,H)

1. 消除非主属性对码的传递依赖,规范到3NF

(A,D,E,F）(商品号，采购日期，采购单价，采购数量)

(A,B,C,G)（商品号，商品名，规格，采购员编号）

(G,H)（采购员编号，采购员姓名）

1. 标准解法分解为3NF
2. 求最小函数依赖集

F={A->B，A->C, G->H, A->G, AD->EF} ；主码为（A,D）

**第一步：右边单一化。**

F1={A->B，A->C, G->H, A->G, AD->E, AD->F}

**第二步：去除多余的函数依赖。无**

F2={A->B, A->C,G->H,A->G,AD->E,AD->F}

**第三步：去掉依赖左侧多余的冗余属性。无**

F3={A->B, A->C,G->H,A->G,AD->E,AD->F}

**F已经是最小依赖集**

1. 关系模式R中存在非主属性B对码AD的部分函数依赖，所以R达不到2NF，只属于1NF。
2. 分解R：
3. 先对F1极小化：F1已是最小依赖集F1={A->B，A->C, G->H, A->G, AD->E, AD->F}
4. 分离无关属性：所有属性均出现，无需分离。

(3) 判是否需分解：需要。

(4) 分解：**按左部相同的原则分组【依赖左部相同的所有属性分为一组】**，则

R ={ R1(A，B，C，G)， R2(G，H)，R3(A,D,E,F) }，

1. 既有无损连接性又保持函数依赖的3NF模式分解【算法6.5】

R ={ R1(A，B，C，G)， R2(G，H)，R3(A,D,E,F) }

码 (A,D)

R3的属性集U中包含码AD，因此R是所求模式分解

题3

【作业题】设有关系模式 R(读者号，姓名，单位号，单位名，图书号，书名，借阅日期，还书日期)存储读者借阅图书等信息。

如果规定：每个读者只属于一个单位；每个读者可以借阅多本图书，每本图书也可以被多名 读者借阅，每个读者也可以对某本图书多次借阅，但每个读者每本图书每天最多借一次。

　　1. 根据上述条件，写出模式 R 的关键字。

　　2. R 最高属于第几范式，为什么?

　　3. 举例说明 R 在何种情况下会发生删除异常。

　　4. 将 R 规范到 3NF。

解法一：

（1）

1）基本函数依赖：

读者号->姓名

单位号->单位名

图书号->书名

每个读者属于一个单位：读者号->单位号【注意不要写成：单位号->读者号】

每个读者可以借阅多本图书，读者不能决定图书

每本图书也可以被多名读者借阅，图书不能决定读者

每个读者也可以对某本图书**多次**借阅：

**（读者，图书)不能决定一次借阅（借阅日期，还书日期）**

每个读者每本图书**每天（当天）**最多**借一次，**（读者号，图书号，借阅日期）决定一次借阅：

**(读者号，图书号，借阅日期）->还书日期**

说明：1.**决定->的含义**是，**属性值唯一确定或唯一对应**，而不一定是生活意义上的决定

1. 这是最后一个函数依赖，除了还书日期，其他可能的非主属性都已经被决定（已出现在依赖右部），所以**此处的依赖属性只写还书日期**。
2. （读者号，图书号，借阅日期）中任意两个属性的组合都不能唯一决定一个记录
3. 函数依赖右部可以是多个属性构成的属性集

2）R的关键字（主码）

（读者号，图书号，借阅日期）

闭包验证：(读者号，图书号，借阅日期)关于R函数依赖的闭包是全部属性的集合

（2）R属于1NF

R不属于2NF,因为R中存在非主属性对码的部分函数依赖，如：读者号->姓名

1. **R在如下情况下会发生删除异常：**

当存在传递依赖或部分依赖时，必然出现数据冗余现象，导致插入、删除、更新/修改异常。

**答案： 删除读者借阅图书信息时可能把读者的基本信息也删除了。**

解析： 1NF 是一个不含重复组的关系，1NF 存在数据冗余、插入和删除异常等问题。比如： 若数据库系统原理这本书被同一个单位的人借 100 次，那么这个单位的单位名被重复记录 100 次，造成大量数据冗余。 若数据库系统原理这本书还没有被借出，那么这本书的基本信息将无法存储到这个关系中去，就会出现插入异常的问题。 删除读者借阅图书信息时可能把读者的基本信息也删除了。

（4）规范到3NF

1）分解R，消除部分依赖，规范为2NF

R1(读者号，图书号，借阅日期，还书日期) 对码完全依赖

R2(读者号，姓名，单位号，单位名) 对码部份依赖

R3(图书号，书名)

2）消除传递依赖，R2仍存在非主属性【单位名】对码【读者号】的传递依赖：

读者号->单位号，单位号->单位名，单位号-\-> 读者号

消除传递依赖，R2模式分解如下：

R2(读者号，姓名，单位号)

R3(单位号，单位名)

1. 规范到3NF的关系模式为：

R1(读者号，图书号，借阅日期，还书日期)

R2(读者号，姓名，单位号)

R3(单位号，单位名)

R4(图书号，书名)

**解法二：（根据算法规范化）**

1. 前面步骤同解法一
2. 规范到3NF

读者号A，姓名B，单位号C，单位名D，图书号E，书名F，借阅日期G，还书日期H

{A->B, C->D, E->F, A->C, AEG->H}

**1.求最小函数依赖：**

1. 所有函数依赖的右边均为单一属性；
2. 去除多余的函数依赖: **无冗余的函数依赖**
3. A+={B}，E+={F}，G+=空集，说明H对A,E,G不存在部分依赖,完全函数依赖于AEG；所有函数依赖的左边没有**冗余属性**
4. 综上，FD集{A->B, C->D, E->F, A->C, AEG->H}是关系模式R的最小函数依赖

**2.分解R为保持函数依赖的3NF【算法6.3】：**

(1)先对F极小化：F={A->B, C->D, E->F, A->C, AEG->H} 已是最小依赖集

(2)分离无关属性：所有属性均出现，无需分离。

(3) 判是否需分解：存在对码的部分依赖，需要分解

(4) 分解：按左部相同的原则分组，则

**R1(读者号，姓名，单位号)**

**R2(单位号，单位名)**

**R3(图书号，书名)**

**R4(读者号，图书号，借阅日期，还书日期)**

3.既有无损连接性又保持函数依赖的3NF模式分解【算法6.5】

码AEG⊆R4的属性集U，因此R是所求模式分解

总结

