

# 第四章 关系规范化理论

1971 E.F.Codd 提出

1NF 2NF 3NF BCNF 4NF 5NF

规范化

函数依赖

模式分解

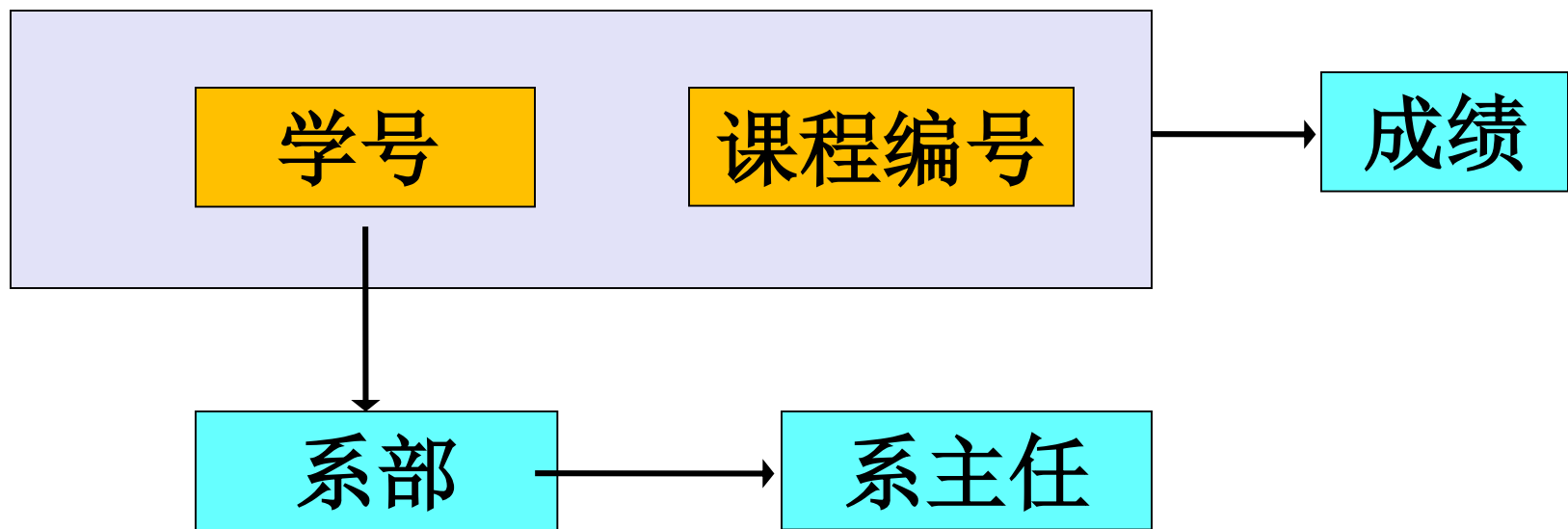
# 4.1 问题的提出

## 关系模式

$R(U, D, \text{dom}, I, F)$

数据依赖：关系中属性间互相依存、互相制约的关系。

[函数依赖、多值依赖、连接依赖、分层依赖和相互依赖]



例：  $U = \{\text{学号}, \text{系部}, \text{系主任}, \text{课程编号}, \text{成绩}\}$   
 $F = \{\text{学号} \rightarrow \text{系部}, \text{系部} \rightarrow \text{系主任},$   
 $[\text{学号}, \text{课程编号}] \rightarrow \text{成绩}\}$

学号	系部	系主任	课程编号	成绩
02101	CS	X	01	85
02101	CS	X	02	74
02101	CS	X	03	70
02102	MA	M	01	85
02102	MA	M	04	75
02103	MA	M	02	60
02104	IS	J	01	75
02105	IS	J	01	89

## 缺点

- 1、冗余太大
- 2、操作异常
  - 1) 插入异常
  - 2) 删除异常
  - 3) 修改异常

# 4.2 规范化理论

## 一、函数依赖：

属性或属性组之间可能存在的依赖性。

### 1、定义

**定义4.1：** 设 $R(U)$ 是属性集 $U$ 上的关系模式。 $X, Y$ 是 $U$ 的子集。若对于 $R(U)$ 的任意一个可能的关系 $r$ ，当且仅当 $r$ 中任意一个给定的 $X$ 的值， $r$ 中存在唯一的 $Y$ 值与之对应。也就是说，如果 $X$ 相等， $Y$ 也相等，则称 $Y$ 函数依赖于 $X$ ，或者 $X$ 函数确定 $Y$ ，记作 $X \rightarrow Y$ 。

**定义4.2:**  $R(U)$  的属性子集  $X$ ,  $Y$  之间的函数依赖用  $X \rightarrow Y$  表示, 它在构成关系  $R$  的任意元组  $r$  上指定了一个约束。这个约束是: 如果对于  $r$  中的任何两个元组  $t_1$  和  $t_2$  有  $t_1[X] = t_2[X]$ , 则必须也有  $t_1[Y] = t_2[Y]$ 。

设  $R(U)$  是属性集  $U$  上的关系模式,  $X, Y$  是  $U$  的子集。若对于  $R(U)$  的任意一个可能的关系  $r$ ,  $r$  中不可能存在两个元组在  $X$  上的属性值相等, 而在  $Y$  上的属性值不等, 则称  $X$  函数确定  $Y$  或  $Y$  函数依赖于  $X$ , 记作  $X \rightarrow Y$ 。

例：  $U = \{\text{学号}, \text{系部}, \text{系主任}, \text{课程编号}, \text{成绩}\}$   
 $F = \{\text{学号} \rightarrow \text{系部}, \text{系部} \rightarrow \text{系主任},$   
 $[\text{学号}, \text{课程编号}] \rightarrow \text{成绩}\}$

**注意：**函数依赖不是指关系模式R的某个或某些关系满足的条件，而是指R的一切关系均要满足的约束条件

**由定义可以导出下列基本概念：**

**1 . 决定因素：**若  $X \rightarrow Y$ ，则  $X$  叫做决定因素

**2 . 互相依赖：**若  $X \rightarrow Y$ ，  $Y \rightarrow X$ ，

则记作  $X \longleftrightarrow Y$ 。

**3 . 若  $Y$  不函数依赖于  $X$ ，则记作  $X \nrightarrow Y$ 。**



### 定义4.3：平凡（非平凡）函数依赖

在 $R(U)$ 中，一个函数依赖如果满足 $Y \subsetneq X$ ，则称此函数依赖是非平凡函数依赖，否则称为平凡函数依赖。

### 定义4.4：完全函数依赖

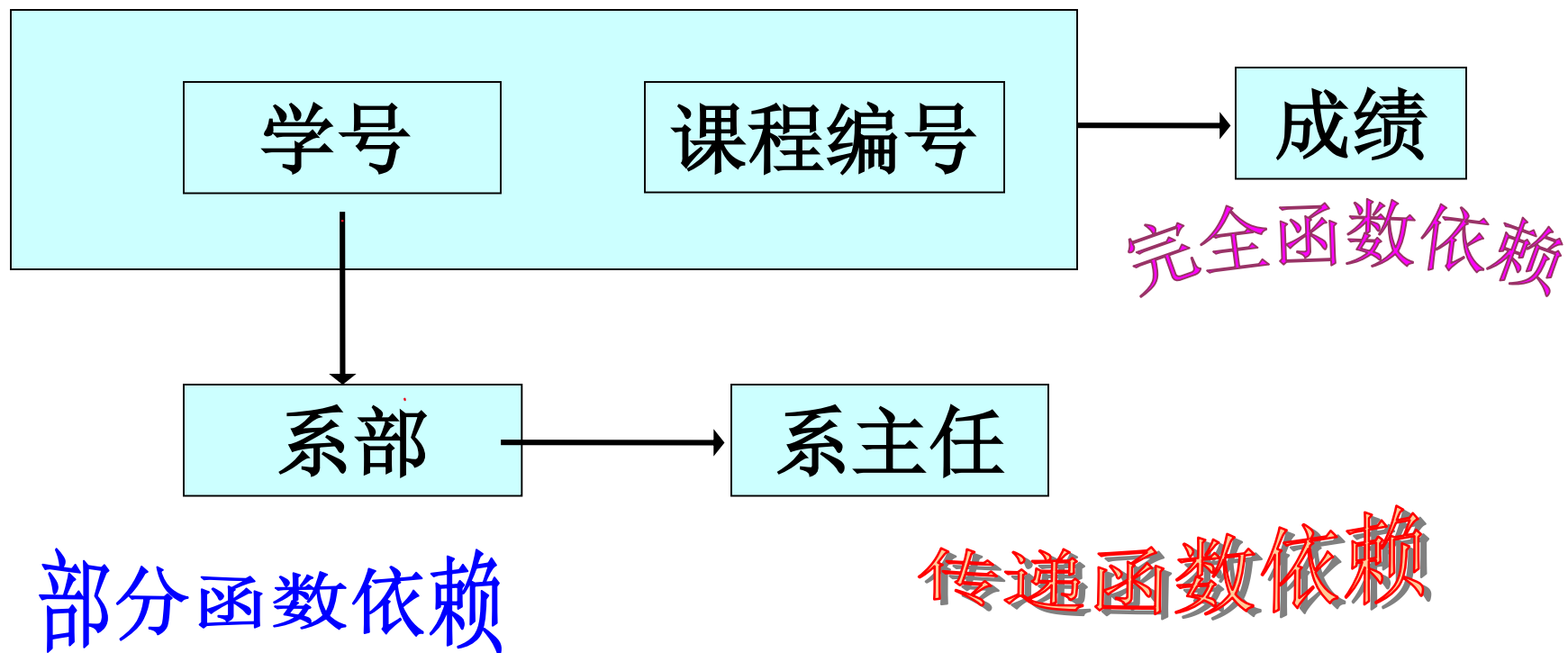
在 $R(U)$ 中，如果 $X \rightarrow Y$ ，并且对于 $X$ 的任何一个真子集 $X'$ ，都有 $X' \not\rightarrow Y$ ，则称 $Y$ 对 $X$ 完全函数依赖。记作： $X \xrightarrow{F} Y$

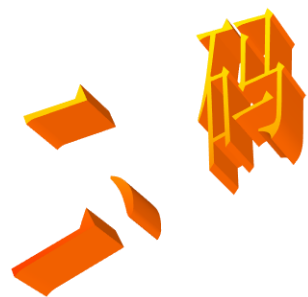
### 定义4.5：部分函数依赖

在 $R(U)$ 中，如果 $X \rightarrow Y$ ，存在真子集 $X'$ ，有 $X' \rightarrow Y$ 成立，则称 $Y$ 对 $X$ 部分函数依赖。记作： $X \xrightarrow{P} Y$

## 定义4.6: 传递函数依赖

在 $R(U)$ 中, 如果 $X \rightarrow Y$ , ( $Y \subsetneq X$ ),  
 $Y \twoheadrightarrow X$ ,  $Y \rightarrow Z$ , 则称 $Z$ 对 $X$ 传递函数依赖。





**定义4.7:** 设 $K$ 为 $R(U, F)$ 中的属性或属性组, 若  $K \xrightarrow{F} U$ , 则 $K$ 为 $R$ 的**候选码**。

**主码:** 若候选码多于一个, 则选定其中的一个为主码。

**主属性:** 包含在任何一个候选码中的属性。

**非主属性:** 不包含在任何码中的属性。

**全码:** 整个属性组是码。

**定义6.6:** 关系模式 $R$ 中属性或属性组 $X$ 并非 $R$ 的码, 但 $X$ 是另一个关系模式的码, 则称 $X$ 是 $R$ 的**外码**。

主码与外码提供了一个表示关系间联系的手段。

## 4.2.3、范式

### 第一范式(1NF)

定义4.9：满足关系的每一个分量是不可分的数据项这一条件的关系模式就属于第一范式(1NF)。

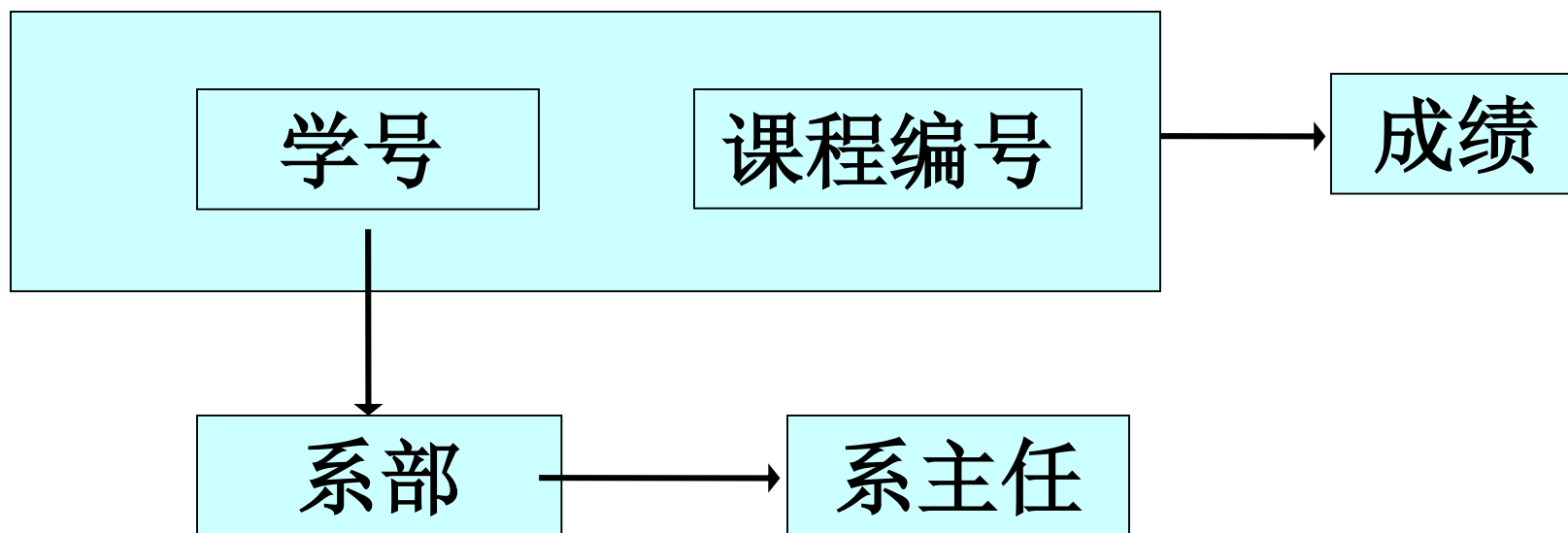
缺点：

插入异常、删除异常、冗余太大、修改复杂

学号	系部	系主任	课程编号	成绩
02101	CS	X	01	85
02101	CS	X	02	74
02101	CS	X	03	70
02102	MA	M	01	85
02102	MA	M	04	75
02103	MA	M	02	60
02104	IS	J	01	75
02105	IS	J	01	89

## 第二范式 (2NF)

定义4.9: 若 $R \in 1NF$ , 且每一个非主属性完全函数依赖于码, 则 $R \in 2NF$ 。



学号



系部



系主任

学号

课程编号



成绩

学号，系部，系主任

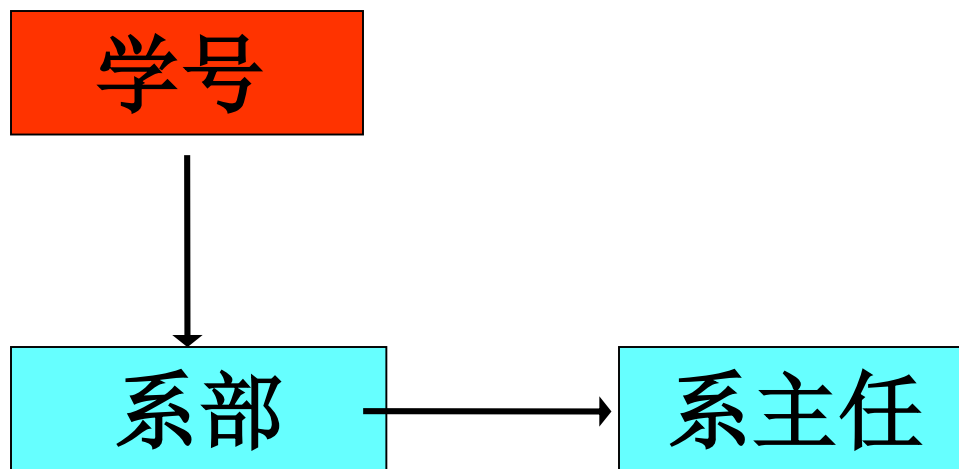
学号，课程编号，成绩



## 第三范式 (3NF)

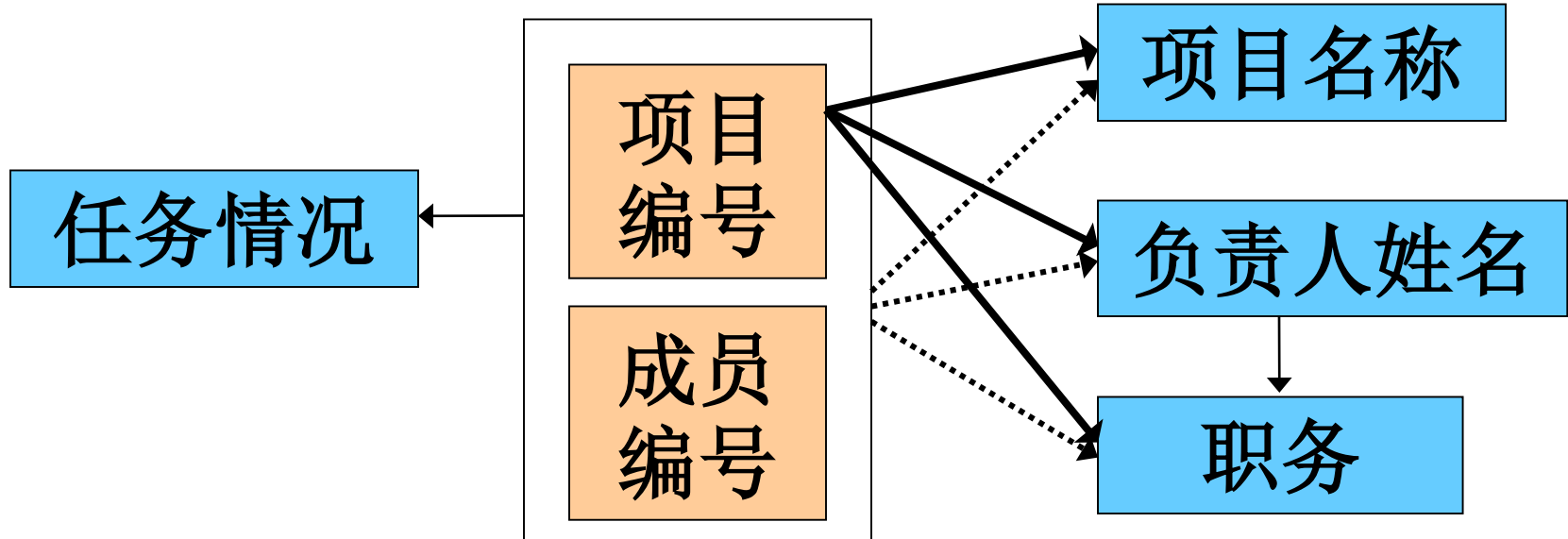
定义4.10: 关系模式 $R(U, F)$ 中若不存在这样的码 $X$ , 属性组 $Y$ 及非主属性组 $Z(Z \subsetneq Y)$ 使得 $X \rightarrow Y$ ,  $(Y \not\rightarrow X)$   $Y \rightarrow Z$ 成立, 则称 $R(U, F) \in 3NF$ 。

即: 若 $R \in 3NF$ , 且每一个非主属性既不部分依赖于码也不传递依赖于码。



若 $R \in 2NF$ , 且每一个非主属性不传递依赖于码, 则 $R \in 3NF$ 。

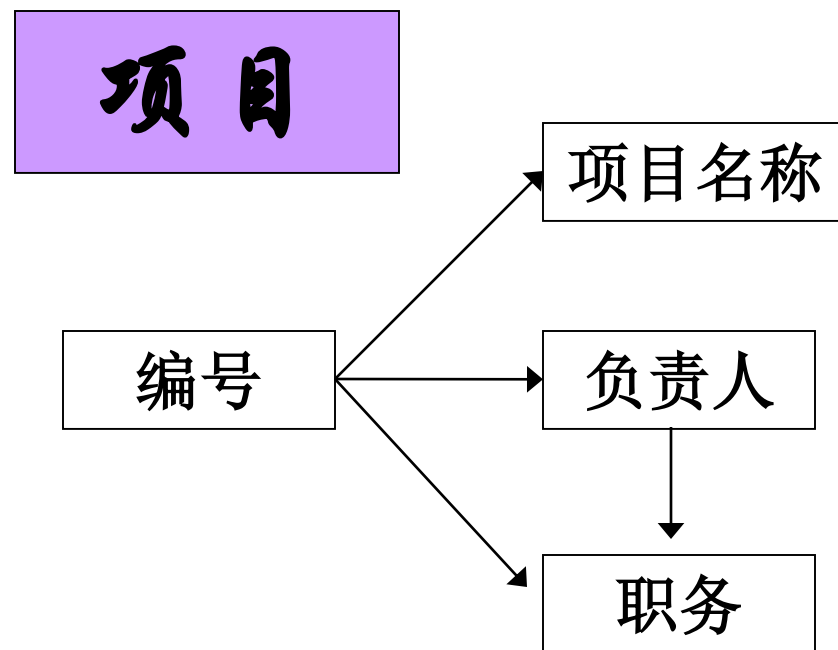
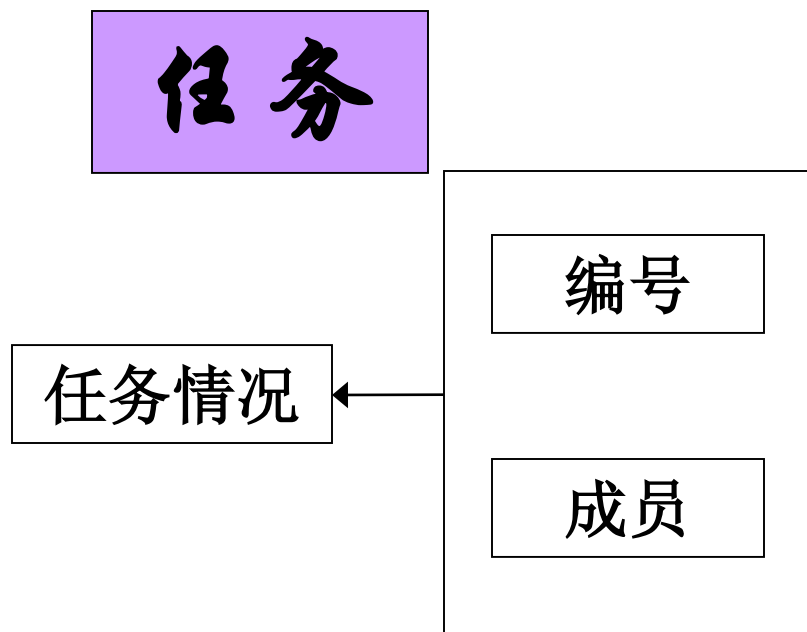
**例：项目（项目编号，项目名称，负责人姓名，职务，成员编号，任务情况）**



# 根据2NF的要求

任务（项目编号，成员编号，任务情况）

项目（项目编号，项目名称，负责人姓名，职务）



# 根据3NF的要求

任务

任务情况

编号

成员

负责人职务

负责人

职务

项目

项目名称

编号

负责人

**例：分析下列关系属于第几范式**

学生学习情况：

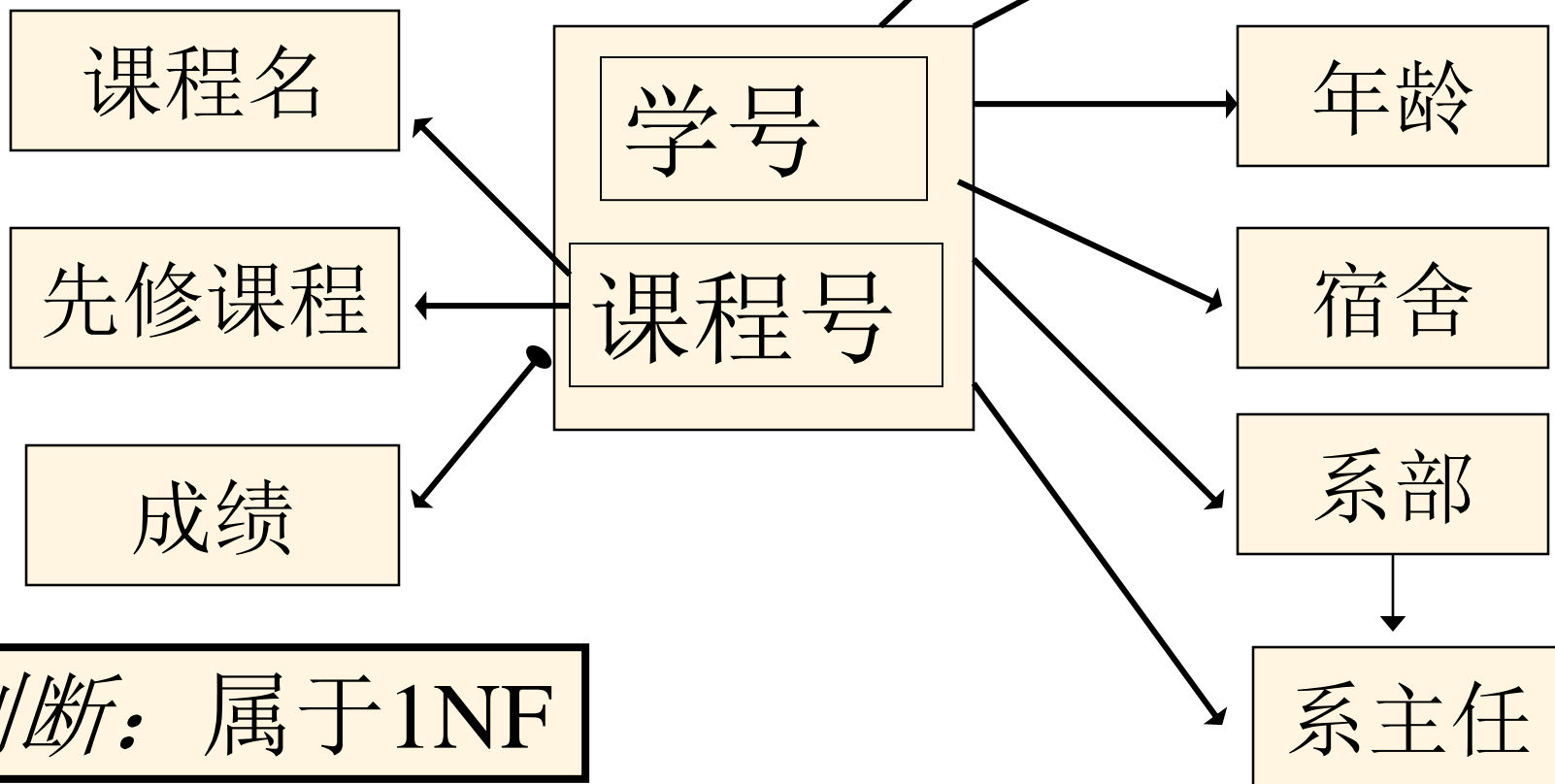
（学号，姓名，班级，年龄，宿舍，系部，系主任，课程号，课程名，先修课程，成绩）

分析:

关键字:

学号  
课程号

分析函数依赖关系:



判断: 属于1NF

# 分解:

学号+课程号

成绩

系部

系主任

课程号

先修课程

课程名

学号

姓名

班级

年龄

宿舍

系部