

# 4장. 데이터 종속성과 정규화

# 1 정규화의 의미

- 정규화(normalization)

- 자료 저장 공간을 최소화하고 데이터베이스 내의 데이터가 불일치 되는 위험을 최소화하여 좋은 데이터베이스 스키마를 설계하는 것을 목적

## 1 정규화의 의미

### ● 데이터베이스 스키마가 잘못 설계된 경우

수 강

학 번	과 목 명	성적	이름
100	전자계산기구조	92	김사랑
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균

83 page

- 100번 학생이 전자계산기구조와 운영체제라는 두 과목을 신청하였기에 이 학생의 이름은 두 번 저장되는 문제점 발생
- 중복된 데이터의 저장으로 인해 저장 공간이 낭비된다는 문제점 이외에도 중복되어 저장된 데이터 때문에 릴레이션의 수정, 삽입, 삭제와 같은 조작을 할 때 여러 가지 곤란한 이상 (anomaly) 현상이 생김

## 1 정규화의 의미

### 1-1 수정 이상

#### ● 수정 이상

- 중복 데이터 중에서 일부만 갱신되어 정보의 모순이 발생하는 것

학 번	과 목 명	성적	이름
100	전자계산기구조	92	<del>김사랑</del> 김소연 ←
101	데이터베이스	82	오지호
100	운영체제	90	김사랑 ←
101	데이터 통신	76	오지호
102	운영체제	82	이선균

84 page

# 1 정규화의 의미

## 1-2 삽입 이상

### ● 삽입 이상

- 불필요한 정보를 함께 저장하지 않음
- 어떤 정보를 저장하는 것이 불가능하기에 원하지 않는 정보를 강제로 삽입해야 하는 것

기본키에는 널 값을 저장할 수 없기 때문에 한예슬이란 학생을 수강 릴레이션에 삽입하려고 한다면 수강 신청하지도 않은 가상의 과목명을 임시로라도 삽입해야 함

학 번 .....	과 목 명 .....	성적	이름
100	전자계산기구조	92	김사랑 김소연
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균
103	?	?	한예슬

85 page

## 1 정규화의 의미

### 1-3 삭제 이상

#### ● 삭제 이상

- 유용한 정보를 함께 삭제하지 않고는 어떤 정보를 삭제하는 것이 불가능한 것

학 번	과 목 명	성적	이름
100	전자계산기구조	92	김사랑 김소연
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균

102번 학생의 이름이  
이선균이라는  
정보까지도 연쇄적으로  
삭제되어 이 학생의  
정보가 모두 손실

86 page

# 1 정규화의 의미

## 1-3 삭제 이상

### ● 정규화(normalization)

- 갱신 이상이 생기지 않도록 불필요한 데이터가 중복되어 저장되지 않게 방지하여 바람직한 릴레이션 스키마로 만들어 가는 과정

### ● 정규화의 목적

- 반복적인 자료를 제거하여 다음과 같은 상태를 만들기 위한 것

1. 어떤 관계라도 데이터베이스 내에서 표현이 가능하도록 만드는 것
2. 관계에서 바람직하지 않은 삽입, 삭제, 갱신 이상이 발생하지 않도록 함
3. 새로운 형태의 데이터가 삽입될 때 관계를 재구성할 필요성을 줄일 수 있음
4. 보다 간단한 관계 연산에 기초하여 검색을 보다 효율적으로 할 수 있음

### ● 함수적 종속 정의

-  $X \rightarrow Y$

### ● 정의

- 릴레이션의 속성들의 부분 집합을  $x, y$ 라 할 때 임의 튜플에서 속성  $x$ 의 값에 의해서 속성  $y$ 의 값이 결정된다면  $y$ 가  $x$ 에 종속됨

### ● 결정자(determinant)

- 주어진 릴레이션에서 다른 속성(또는 속성들의 집합)을 고유하게 결정하는 하나 이상의 속성



## 2 함수종속

- 수강 릴레이션에서 결정자를 찾아보자

수 강

학 번 .....	과 목 명 .....	성적	이름
100	전자계산기구조	92	김사랑
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균

87 page

예

학번 → 이름  
{학번, 과목명} → 성적

88 page

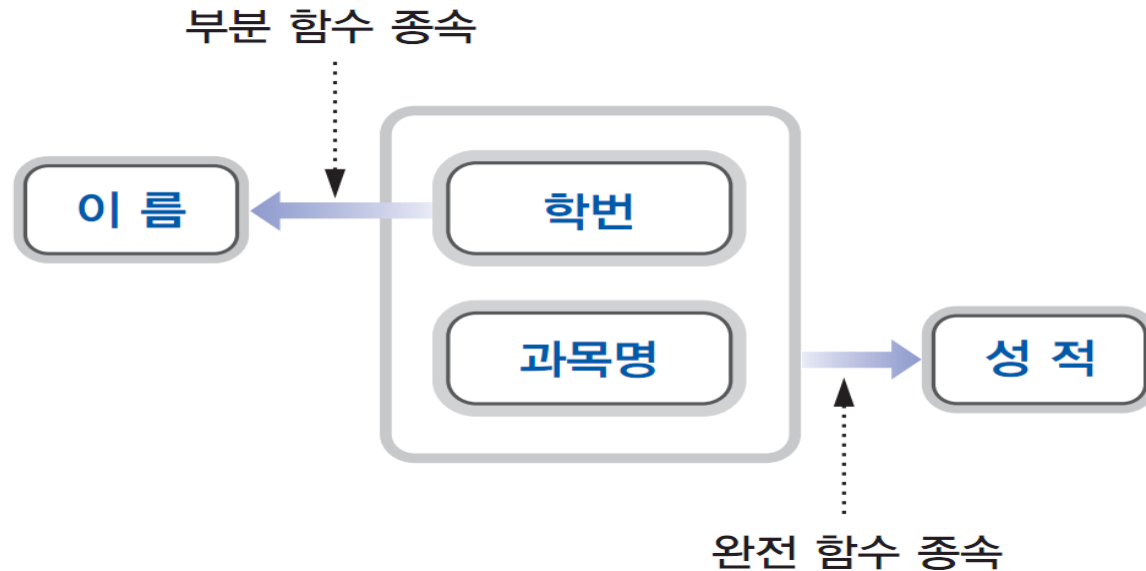
## 2 함수종속

### ● 함수적 종속의 성질

- 완전 함수 종속(FFD: Full Functional Dependency)
- 부분 함수 종속(PFD: Partial Functional Dependency)

### ● 함수 종속 다이어그램

- 릴레이션 내의 속성들의 종속 관계를 보다 쉽게 이해하기 위해서는 이를 도식적으로 표현



88 page

### 3 정규화 과정

#### ● 정규화(Normalization)

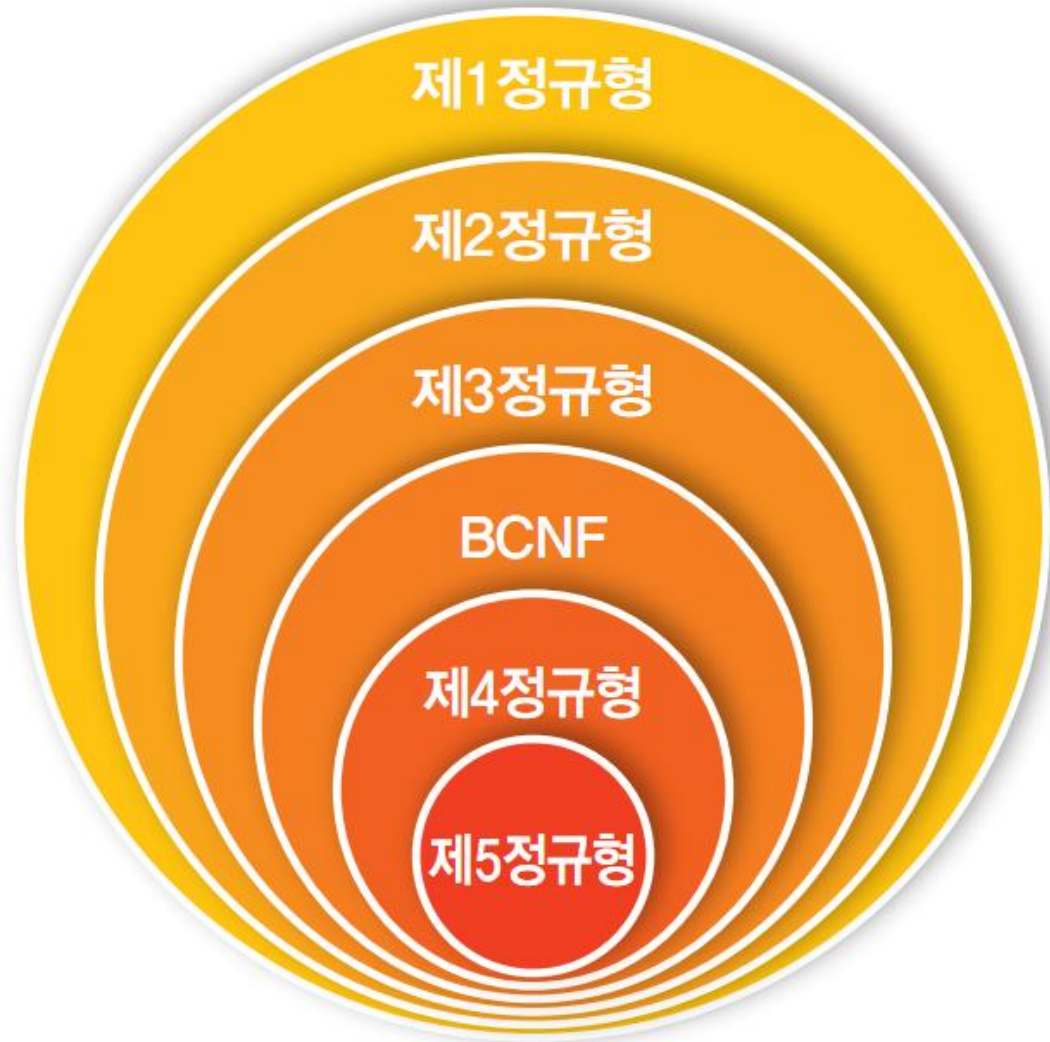
- 주어진 릴레이션 스키마를 함수적 종속성 등의 종속 이론을 이용하여 잘못된 릴레이션 스키마를 보다 더 작은 속성의 세트로 나누어서 갱신 이상이 발생하지 않는 바람직한 릴레이션 스키마로 만들어 가는 과정

#### ● 정규형(Normal Form)

- 테이블의 정규화된 정도

### 3 정규화 과정

- 정규형들 간의 관계



89 page

### 3 정규화 과정

## 3-1 제1정규형(1NF)

### ● 제1정규형에 만족하는 릴레이션

- 어떤 릴레이션 R에 속한 모든 도메인이 원자값(atomic value)
- 릴레이션의 속성값이 반복 집단이 없는 즉, 더 이상 분해될 수 없는 원자값으로만 구성

### ● 반복 집합이 있는 비정규 릴레이션

- 반복 집합은 한 개의 기본키 값에 대해서 두 개 이상의 값을 가질 수 있는 속성

수 강

학 번	과 목 명	성적	이름
100	{전자계산기구조, 운영체제}	{92, 90}	김사랑
101	{데이터베이스, 데이터 통신}	{82, 76}	오지호
102	운영체제	82	이선균

90 page

### 3 정규화 과정

## 3-1 제1정규형(1NF)

### ● 제1정규화 과정의 개념



### ● 제1정규화 과정

90 page

반복 집합이 있는 비정규 릴레이션

학 번	과 목 명	성적	이름
100	{전자계산기구조, 운영체제}	{92, 90}	김사랑
101	{데이터베이스, 데이터 통신}	{82, 76}	오지호
102	운영체제	82	이선균

반복 집합을  
제거함

원자값만으로 구성된 제1정규형 릴레이션

학 번	과 목 명	성적	이름
100	전자계산기구조	92	김사랑
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균

### 3 정규화 과정

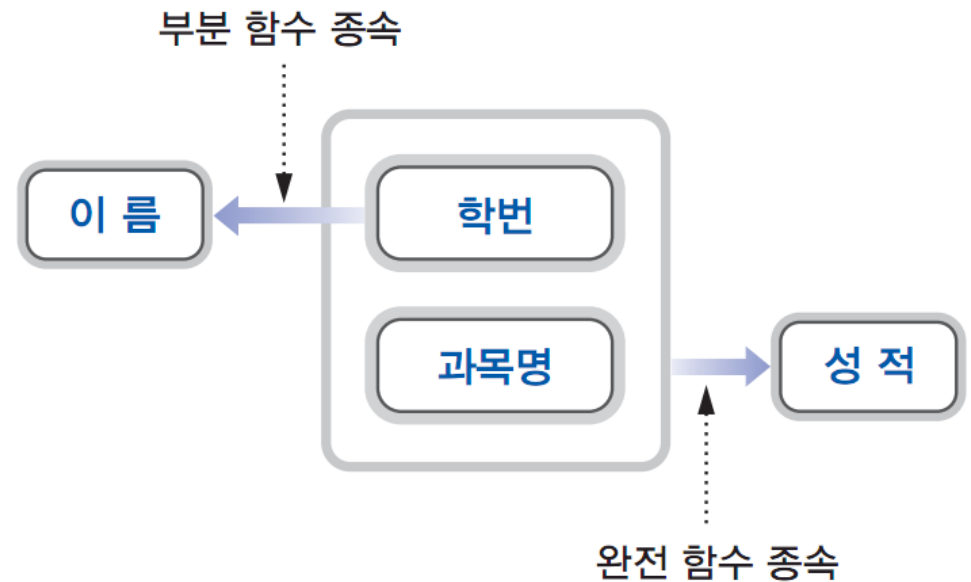
## 3-2 제2정규형(2NF)

### ● 제2정규형

- 어떤 릴레이션 R이 제1정규형이고, 기본 키에 속하지 않는 속성이 모든 (후보)키에 완전 함수 종속

수강

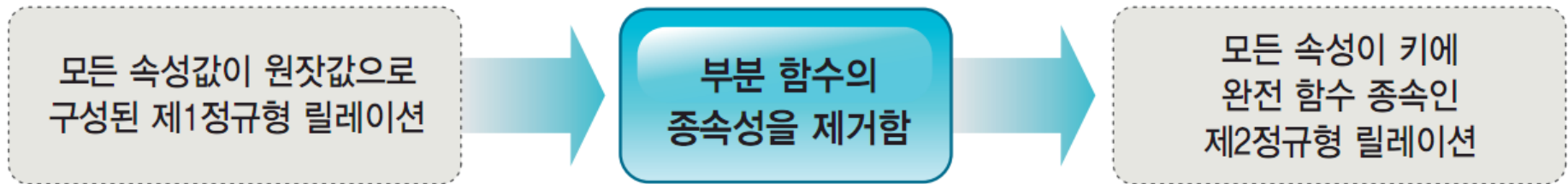
학번	과목명	성적	이름
100	전자계산기구조	92	김사랑
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균



### 3 정규화 과정

## 3-2 제2정규형(2NF)

- 제2정규화 과정의 개념



91 page



### 3 정규화 과정

## 3-2 제2정규형(2NF)

### ● 제2정규화 과정

수 강

학 번	과 목 명	성적	이름
100	전자계산기구조	92	김사랑
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균

부분 함수의  
종속성을 제거함

학 생

학 번	이름
100	김사랑
101	오지호
102	이선균

수 강 1

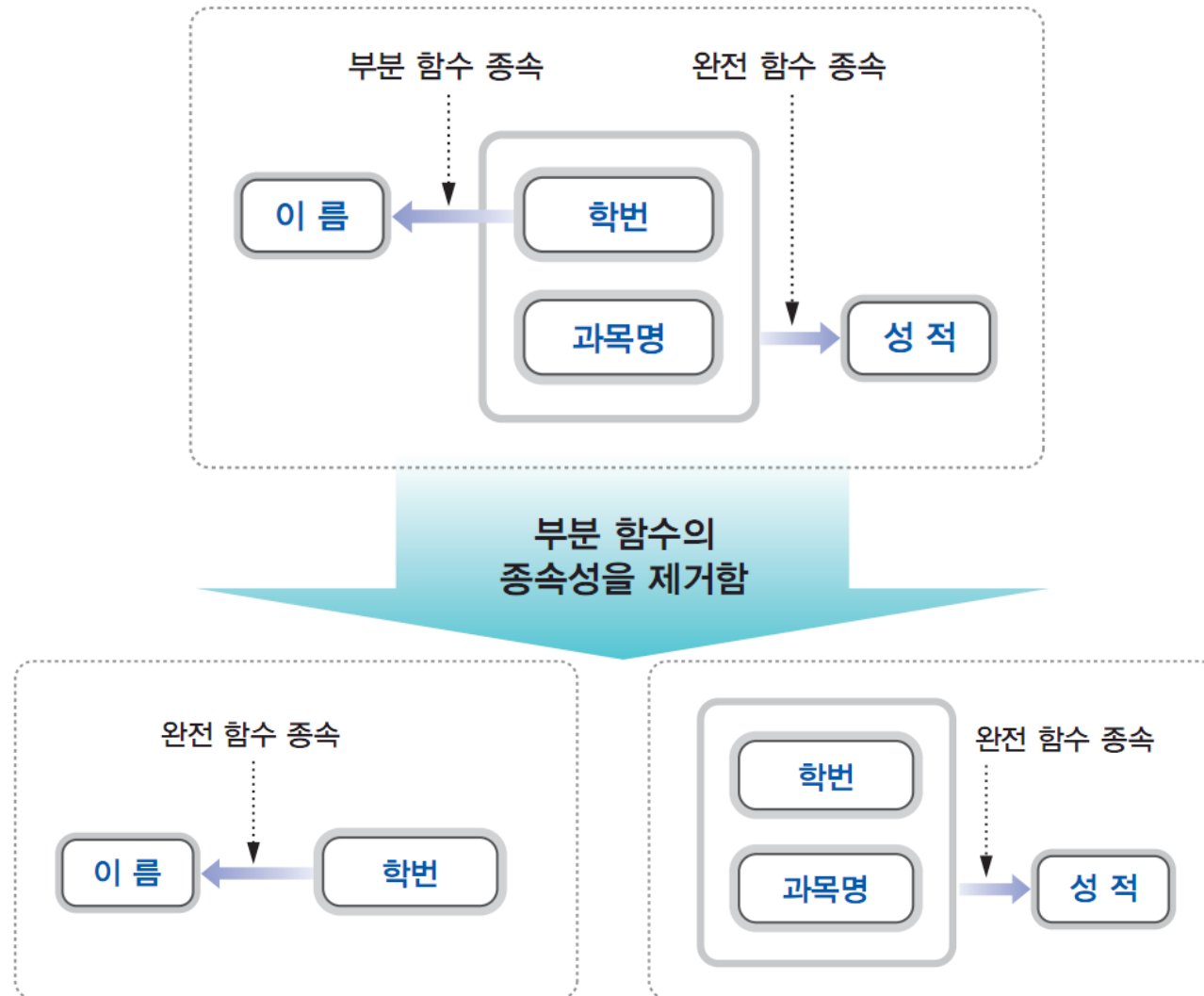
학 번	과 목 명	성적
100	전자계산기구조	92
101	데이터베이스	82
100	운영체제	90
101	데이터 통신	76
102	운영체제	82

92 page

### 3 정규화 과정

## 3-2 제2정규형(2NF)

- 제2정규화 과정의 함수 종속 다이어그램



### 3 정규화 과정

## 3-2 제2정규형(2NF)

#### ● 무손실 분해(nonloss decomposition)

- 자연 조인하였을 때 아무런 정보 손실 없이 다시 원래의 릴레이션으로 복귀된다면 2NF로 두 개의 릴레이션으로 분해하는 것

수 강

학 번	과 목 명	성적	이름
100	전자계산기구조	92	김사랑
101	데이터베이스	82	오지호
100	운영체제	90	김사랑
101	데이터 통신	76	오지호
102	운영체제	82	이선균

부분 함수의  
종속성을 제거함

학 생

학 번	이름
100	김사랑
101	오지호
102	이선균

수 강 1

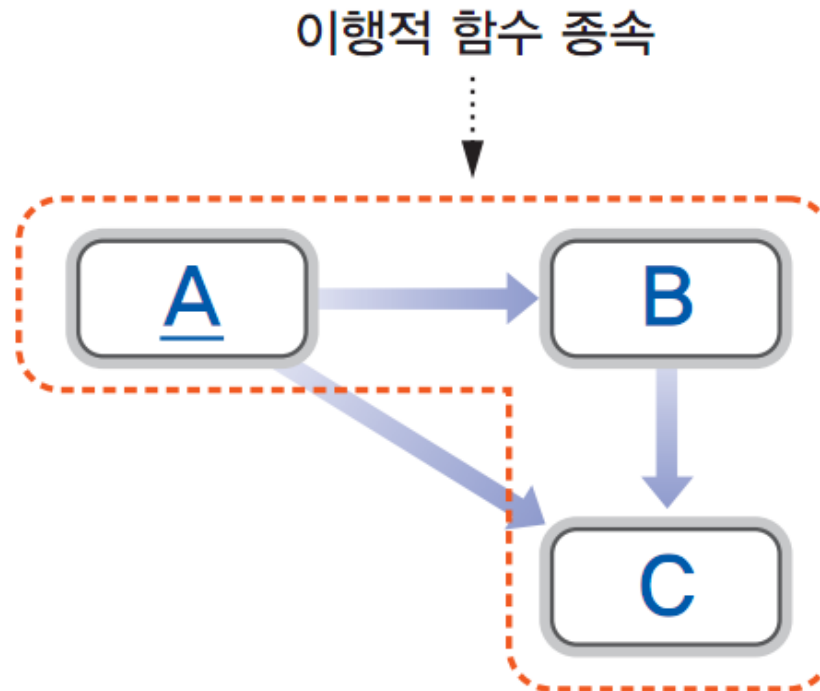
학 번	과 목 명	성적
100	전자계산기구조	92
101	데이터베이스	82
100	운영체제	90
101	데이터 통신	76
102	운영체제	82

### 3 정규화 과정

## 3-3 제3정규형(3NF)

- 이행적 함수 종속성

- 3개의 속성에 존재하는 함수의 종속성을 의미



### 3 정규화 과정

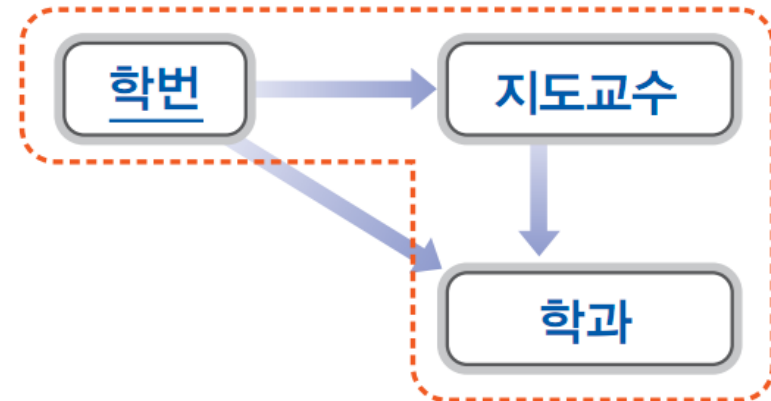
## 3-3 제3정규형(3NF)

### ● 이행적 함수 종속하는 지도 릴레이션

지도

학 번	지도교수	학과
100	이문세	컴퓨터 소프트웨어학과
101	김연아	멀티미디어학과
102	이문세	컴퓨터 소프트웨어학과
103	강승범	경영 정보학과
104	이문세	컴퓨터 소프트웨어학과
105	김연아	멀티미디어학과

이행적 함수 종속



94 page

예

학 번 → 지도교수 ∧ 지도교수 → 학과 → 학 번 → 학과

## 3 정규화 과정

### 3-3 제3정규형(3NF)

#### ● 제3정규형

- 어떤 릴레이션 R이 2NF이고, 모든 비주요 속성(nonprime attribute)들이 후보키(candidate key)에 이행적 함수 종속(transitive FD)하지 않는다.
- 단, 속성이 후보키를 통해 이행적 함수 종속 하는 것은 허용

### 3 정규화 과정

## 3-3 제3정규형(3NF)

- 이행적 함수 종속으로 인한 갱신 이상

- 수정 이상

학 번	지도교수	학과
100	이문세	컴퓨터 소프트웨어학과
101	김연아	멀티미디어학과
102	이문세	<del>컴퓨터 소프트웨어학과</del> 정보처리학과
103	강승범	경영 정보학과
104	이문세	컴퓨터 소프트웨어학과
105	김연아	멀티미디어학과

정보의 모순성이 발생

95 page

### 3 정규화 과정

#### 3-3 제3정규형(3NF)

- 이행적 함수 종속으로 인한 갱신 이상

- 삽입 이상

96 page

전혜영이란 교수가 모바일학과에 속한다는 사실만 기록해 놓으려고 하는데 이 교수의 지도를 받는 학생이 없다면 지도 릴레이션에 이 교수의 정보를 삽입할 수 없음

학 번 -----	지도교수	학과
100	이문세	컴퓨터 소프트웨어학과
101	김연아	멀티미디어학과
102	이문세	컴퓨터 소프트웨어학과
103	강승범	경영 정보학과
104	이문세	컴퓨터 소프트웨어학과
105	김연아	멀티미디어학과
?	전혜영	모바일학과



### 3 정규화 과정

#### 3-3 제3정규형(3NF)

- 이행적 함수 종속으로 인한 갱신 이상

- 삭제 이상

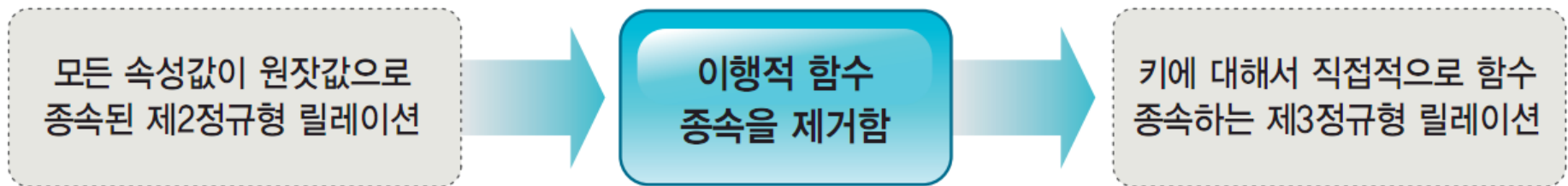
103번 학생이 강승범 교수의 지도를 받지 않겠다고 지도 관계를 취소하면 강승범 교수가 경영 정보학과 소속이라는 정보까지도 연쇄적으로 삭제 유지되어야 할 정보까지도 연쇄 삭제

학 번	지도교수	학과
100	이문세	컴퓨터 소프트웨어학과
101	김연아	멀티미디어학과
102	이문세	컴퓨터 소프트웨어학과
<del>103</del>	<del>강승범</del>	<del>경영 정보학과</del>
104	이문세	컴퓨터 소프트웨어학과
105	김연아	멀티미디어학과

### 3 정규화 과정

## 3-3 제3정규형(3NF)

- 제3정규화 과정의 개념



97 page

### 3 정규화 과정

## 3-3 제3정규형(3NF)

### ● 제3정규화 과정

지도(이행적 함수 종속 릴레이션)

학 번	지도교수	학과
100	이문세	컴퓨터 소프트웨어학과
101	김연아	멀티미디어학과
102	이문세	컴퓨터 소프트웨어학과
103	강승범	경영 정보학과
104	이문세	컴퓨터 소프트웨어학과
105	김연아	멀티미디어학과

이행적 함수  
종속을 제거함

지도1(이행적 함수 종속 제거한 3NF)

학 번	지도교수
100	이문세
101	김연아
102	이문세
103	강승범
104	이문세
105	김연아

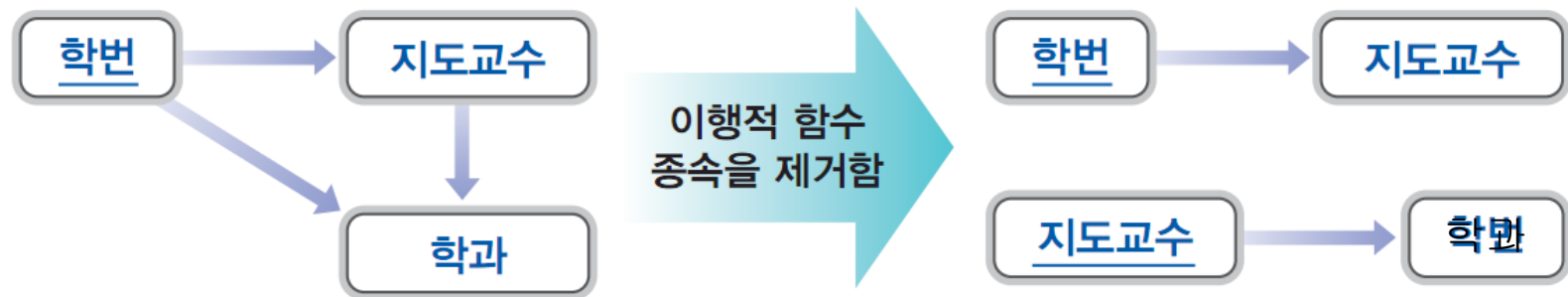
교수(이행적 함수 종속 제거한 3NF)

지도교수	학과
이문세	컴퓨터 소프트웨어학과
김연아	멀티미디어학과
강승범	경영 정보학과

### 3 정규화 과정

#### 3-3 제3정규형(3NF)

- 제3정규화 과정의 함수 종속 다이어그램



98 page

### 3 정규화 과정

## 정규형의 비공식적 정의

- 1<sup>st</sup> normal form
  - All attributes depend on **the key**
- 2<sup>nd</sup> normal form
  - All attributes depend on **the whole key**
- 3<sup>rd</sup> normal form
  - All attributes depend on **nothing but the key**

### 3 정규화 과정

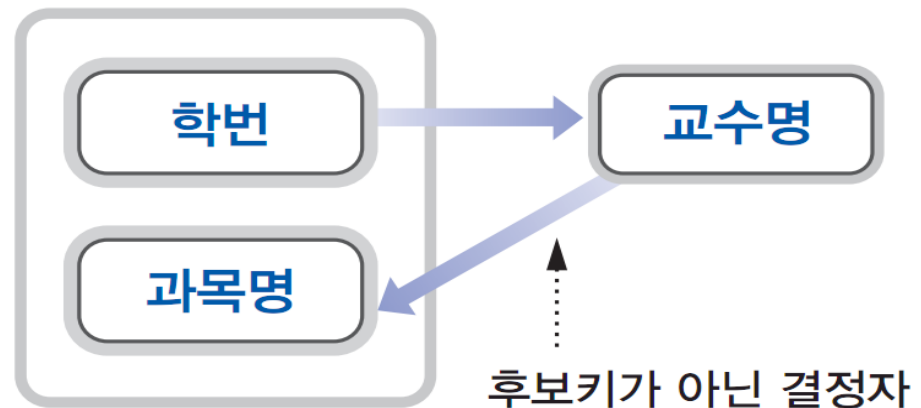
## 3-4 보이스/코드 정규형

### ● 보이스/코드 정규형

- 릴레이션 R이 제3정규형을 만족하고 모든 결정자가 후보키어야 한다.
- 복잡한 식별자 관계에 의해 발생하는 문제를 해결하기 위해서 제3정규형을 보완한 것
- 모든 결정자가 후보키
- 예: 만약 한 교수는 한 과목만을 강의한다고 가정하자. 하나의 과목을 가르치는 교수는 여러명 일 수 있다.

수 강

학 번	과목명	교수명
100	전자계산기구조	이문세
100	데이터베이스	김연아
100	운영체제	강승범
101	데이터베이스	김연아
101	운영체제	전혜영



### 3 정규화 과정

#### 3-4 보이스/코드 정규형

- 결정자가 후보키가 아닌 릴레이션에서의 갱신 이상

- 수정 이상

학 번	과목명	교수명
100	전자계산기구조	이문세
100	데이터베이스	<del>김연아</del> 김주현
100	운영체제	강승범
101	데이터베이스	김연아
101	운영체제	전혜영

정보의 모순성이 발생

99 page

### 3 정규화 과정

#### 3-4 보이스/코드 정규형

- 결정자가 후보키가 아닌 릴레이션에서의 갱신 이상

- 삽입 이상

서경진이란 교수가 웹 프로그래밍을 가르친다는 사실만 기록해 놓으려고 하는데 웹 프로그래밍을 수강 신청한 학생이 없다면 수강 릴레이션에 이 교수의 정보를 삽입할 수 없음

학 번 .....	과목명 .....	교수명
100	전자계산기구조	이문세
100	데이터베이스	김연아
100	운영체제	강승범
101	데이터베이스	김연아
101	운영체제	전혜영
?	웹 프로그래밍	서경진



### 3 정규화 과정

#### 3-4 보이스/코드 정규형

- 결정자가 후보키가 아닌 릴레이션에서의 갱신 이상

- 삭제 이상

101번 학생이 전해영 교수의 운영체제 수업을 듣지 않기로 하였다면 한 명의 학생만이 수강 신청한 과목을 삭제한 것이기에 운영체제를 전해영 교수가 가르친다는 정보마저 삭제되는 연쇄삭제 현상이 일어나게 되어 정보 손실이 발생

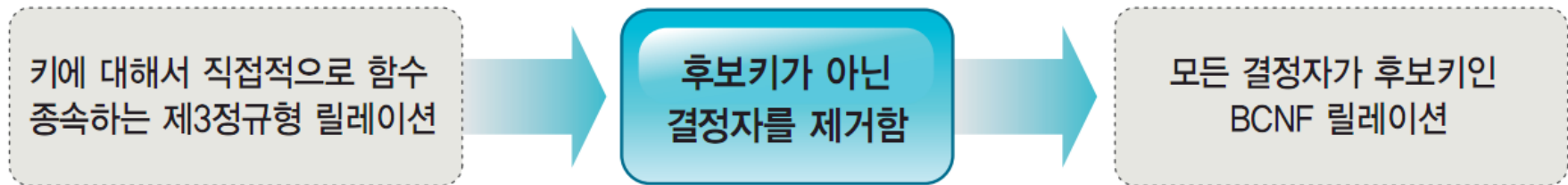
학 번	과목명	교수명
100	전자계산기구조	이문세
100	데이터베이스	김연아
100	운영체제	강승범
101	데이터베이스	김연아
<del>101</del>	<del>운영체제</del>	<del>전해영</del>

100 page

### 3 정규화 과정

## 3-4 보이스/코드 정규형

- 보이스/코드 정규화 과정 개념



100 page

### 3 정규화 과정

## 3-4 보이스/코드 정규형

- 보이스/코드 정규화 과정

수강(결정자가 키가 아닌 속성인 릴레이션)

학 번	과목명	교수명
100	전자계산기구조	이문세
100	데이터베이스	김연아
100	운영체제	강승범
101	데이터베이스	김연아
101	운영체제	전혜영

후보키가 아닌  
결정자를 제거함

지도교수(모든 결정자가 후보키인 BCNF 릴레이션)

학 번	교수명
100	이문세
100	김연아
100	강승범
101	김연아
101	전혜영

교수(모든 결정자가 후보키인 BCNF 릴레이션)

교수명	과목명
이문세	전자계산기구조
김연아	데이터베이스
강승범	운영체제
김연아	데이터베이스
전혜영	운영체제

101 page

### 3 정규화 과정

#### 3-4 보이스/코드 정규형

- 정규형의 특징과 정규화 과정

