

# DB

## 9 정규화

## 주요 질문

# 9.

## 정규화

- ✓ 데이터 이상현상(Anomaly)란 무엇이며 어떤 것들이 있나요?
- ✓ 정규화의 개념과 목적에 대해서 설명해 보세요.
- ✓ 2차 정규화에 대해서 설명해 보세요.

## 이상현상(Anomaly)

- 데이터의 중복으로 인해 릴레이션을 조작할 때 발생하는 비합리적 현상
- 하나의 릴레이션에 두 개 이상의 속성들을 혼합하여, 즉 정규화 되지 않은 경우 데이터의 중복이 원인이 되어 발생함

## 이상현상 종류

| 종류                          | 주요개념   |
|-----------------------------|--|
| 갱신이상<br>(Update Anomaly)    | 릴레이션R에서 특정 속성값 갱신시 - 중복 저장되어 있는 속성값 중 하나만 갱신하고, 나머지는 갱신하지 않아서 발생하는 데이터의 불일치 현상(Data Inconsistency) |
| 삭제이상<br>(Deletion Anomaly)  | 릴레이션R에서 특정 튜플을 삭제할 경우 - 원하지 않는 정보까지도 삭제되는 현상   |
| 삽입이상<br>(Insertion Anomaly) | 릴레이션R에서 특정 튜플을 삽입할 경우 - 원하지 않는 불필요한 정보까지도 삽입해야 하는 현상   |

## 이상현상(Anomaly)

9.

정규화

| 학번    | 과목코드 | 학생이름 | 과목명    | 학점 |
|-------|------|------|--------|----|
| 96404 | B102 | 홍길동  | 자료구조   | A+ |
| 96404 | C203 | 홍길동  | 데이터베이스 | A  |
| 98515 | C234 | 김영희  | 알고리즘   | B  |
| 99316 | B102 | 김철수  | 자료구조   | B+ |

| 종류                           | 설명   |
|------------------------------|--|
| 삽입 이상<br>(Insertion Anomaly) | 학번 key 추가 시, 수강한 과목이 없어도 임의의 과목코드를 만들어 삽입해야 함                          |
| 갱신 이상<br>(Update Anomaly)    | 학번 96404과 과목코드 B102의 과목명을 변경 시(운영체제), {99316, B102}의 과목명과 다른 이름을 갖게 됨. |
| 삭제 이상<br>(Deletion Anomaly)  | 98515 학번을 삭제하면(자퇴 등), 알고리즘 과목 까지 삭제됨                                   |

## 정규화(Normalization)의 개념

## 9. 정규화

- 이상현상을 발생시키는 속성 간의 종속성, 중복성을 제거하고 무결성을 보장하기 위해 릴레이션을 분해하는 과정
- 이상 현상(Anomaly)을 야기하는 Attribute 간의 종속 관계를 제거하기 위해 Relation을 작은 Relation으로 무손실 분해하는 과정

### 정규화의 목적

| 목적                  | 내용                   |
|---------------------|----------------------|
| 데이터의 중복 최소화         | 데이터 저장공간의 최소화        |
| 이상현상(Anomaly) 발생 방지 | 정보의 불일치 및 손실의 위험 최소화 |
| 데이터 구조의 안정성 유지      | 향후 새로운 요구에 유연하게 대처   |
| 종속성 제거              | 데이터의 일관성과 무결성 보장     |

## 정규화(Normalization)의 종류

## 9. 정규화

| 정규화 종류  | 함수적 종속성          | 키워드  |
|---------|------------------|--|
| 1차 정규화  | 완전함수 종속성         | <ul style="list-style-type: none"> <li>속성의 원자화. 다중값 및 반복되는 속성을 제거</li> </ul>   |
| 2 차 정규화 | 부분함수 종속성<br>제거   | <ul style="list-style-type: none"> <li>1차 정규화를 만족하고 Relation의 기본키가 아닌 속성 들이 완전함수적으로 종속할 경우</li> <li>복합키에서 발생, 개별키에 의한 종속 발생 시</li> </ul> |
| 3차 정규화  | 이행함수 종속성<br>제거   | <ul style="list-style-type: none"> <li>2NF를 만족하고 기본키 외의 속성들 간에 함수적 종속성을 가지지 않는 경우</li> <li>일반 속성 들에서 종속성이 발생 시</li> </ul>                |
| BCNF    | 결정자 함수 종속성<br>제거 | <ul style="list-style-type: none"> <li>Relation의 모든 결정자가 후보키 일 경우</li> <li>일반 속성에 의해 기본키가 종속성 발생 시</li> </ul>                            |

## 정규화(Normalization)의 종류

9.

정규화

### 부분함수종속성

| 성적        |  |
|-----------|--|
| <u>학번</u> |  |
| 과목코드      |  |
| 학점        |  |
| 과목명       |  |

### 이행함수종속성

| 학생        |  |
|-----------|--|
| <u>학번</u> |  |
| 학생이름      |  |
| 학과        |  |
| 학과사무실     |  |

### 결정자함수종속성

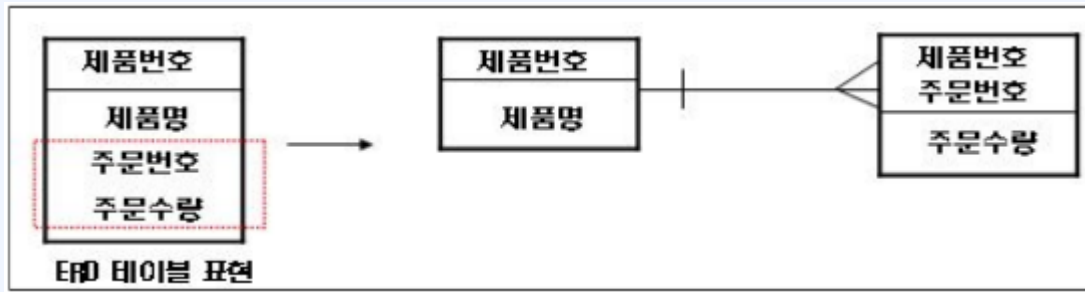
| 학생        |  |
|-----------|--|
| <u>학번</u> |  |
| 과목        |  |
| 학점        |  |
| 교수        |  |

## 1차 정규화

## 9. 정규화

제품번호 1001 이 반복(중복) 되는 현상 제거

| 제품번호 | 제품명 | 주문번호        | 주문수량 |
|------|-----|-------------|------|
| 1001 | 컴퓨터 | 20040101001 | 1    |
| 1001 | 컴퓨터 | 20040101002 | 1    |
| 1001 | 컴퓨터 | 20040101003 | 2    |





## 1차 정규화

## 9. 정규화

데이터가 원자값으로 존재하지 않을 때

| 학번    | 학생이름 | 과목명          |
|-------|------|--------------|
| 96404 | 홍길동  | 운영체제, 데이터베이스 |
| 96506 | 김기동  | 운영체제, 알고리즘   |
| 98515 | 김영희  | 네트워크, 자료구조   |
| 99316 | 김철수  | 컴퓨터구조, SW공학  |

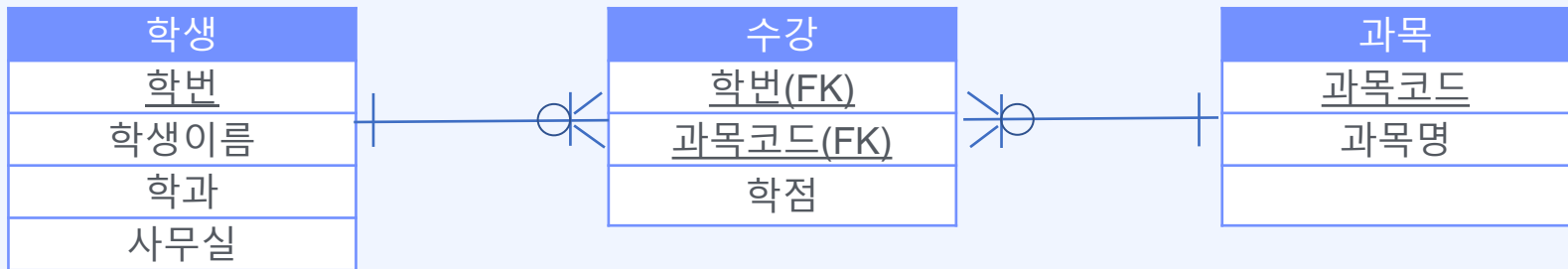
## 2차 정규화

9.  
정규화

## 부분함수 종속성 제거

| 학번    | 과목코드 | 학생이름 | 과목명    | 학과    | 학과사무실 | 학점 |
|-------|------|------|--------|-------|-------|----|
| 96404 | B102 | 홍길동  | 자료구조   | 컴퓨터공학 | 공학관   | A+ |
| 96404 | C203 | 홍길동  | 데이터베이스 | 컴퓨터공학 | 공학관   | A  |
| 98515 | C234 | 김영희  | 알고리즘   | SW공학  | 자연관   | B  |
| 99316 | B102 | 김철수  | 자료구조   | 멀티미디어 | 학사관   | B+ |

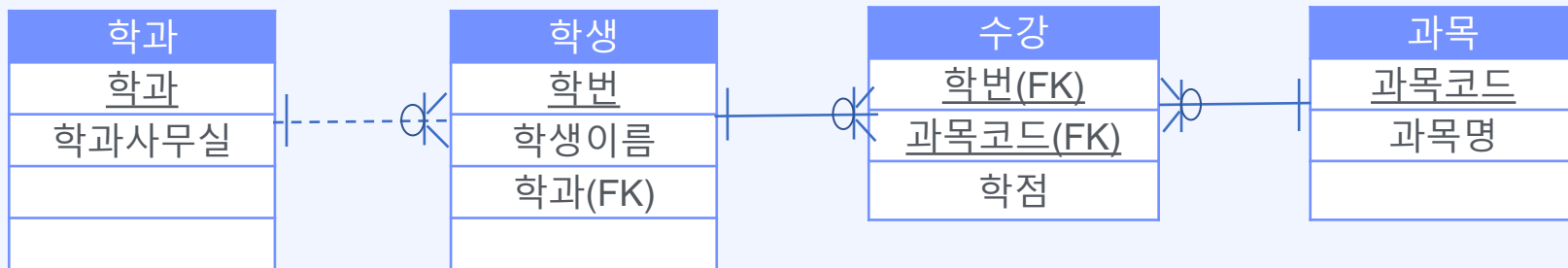
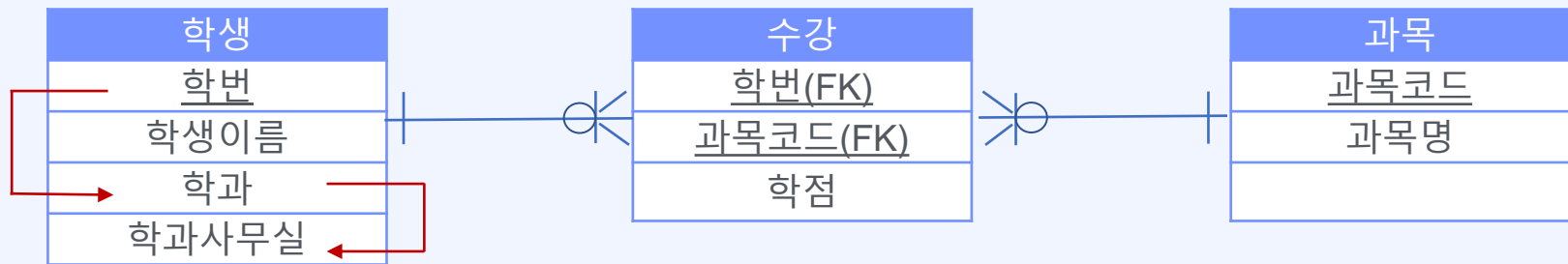
학번+과목코드 -> 학점 / 학번 -> 학생이름 / 과목코드 -> 과목명



### 3차 정규화

## 9. 정규화

#### 이행함수 종속성 제거



BCNF  
(Boyce-Code Normalization)

## 결정자 함수적 종속성 제거

| 학번    | 과목   | 학점 | 교수    |
|-------|------|----|-------|
| 96404 | 운영체제 | A+ | P2032 |
| 96404 | 자료구조 | B+ | P2039 |
| 98515 | 알고리즘 | A  | P3017 |
| 99316 | 자료구조 | B  | P2039 |

- 1) 각 과목에 대한 한 학생은 오직 한교수의 강의만 수강한다
- 2) 각 교수는 한과목만 담당한다
- 3) 한 과목은 여러 교수가 담당할 수 있다



## BCNF (Boyce-Code Normalization)

### 결정자 함수적 종속성 제거

