



## 7장 정규화

김철학

# 목차

---

1. 이상현상
2. 함수 종속성
3. 함수 종속성과 이상현상
4. 정규화
5. 제1정규형
6. 제2정규형
7. 제3정규형
8. BCNF
9. 정규화 정리

# 1. 이상현상

- 이상현상<sup>Anomaly</sup>은 데이터베이스에서 데이터를 조작, 삽입, 삭제, 갱신할 때 발생할 수 있는 비정상적인 상태
- 이상현상에는 삽입이상, 삭제이상, 갱신이상이 있으면 이상현상은 데이터베이스의 일관성과 무결성을 저해

| 학생번호 | 학생이름 | 학과   | 주소       | 강좌이름   | 강의실     |
|------|------|------|----------|--------|---------|
| 501  | 박지성  | 컴퓨터과 | 영국 맨체스타  | 데이터베이스 | 공학관 110 |
| 401  | 김연아  | 체육학과 | 대한민국 서울  | 데이터베이스 | 공학관 110 |
| 402  | 장미란  | 체육학과 | 대한민국 강원도 | 스포츠경영학 | 체육관 103 |
| 502  | 추신수  | 컴퓨터과 | 미국 클리블랜드 | 자료구조   | 공학관 111 |
| 501  | 박지성  | 컴퓨터과 | 영국 맨체스타  | 자료구조   | 공학관 111 |
| 403  | 박세리  | 체육학과 | 대한민국 대전  | NULL   | NULL    |

**DELETE**

- 장미란 학생 삭제
- 연쇄삭제 문제

**UPDATE**

- 박지성 학생 주소 변경
- 불일치 문제

**INSERT**

- 박세리 학생 삽입
- NULL 값 문제

## 2. 함수 종속성

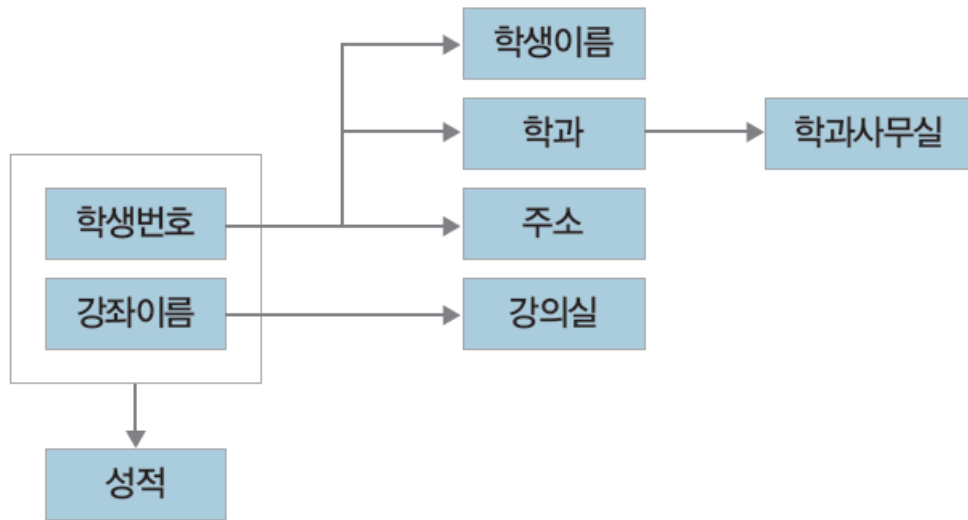
- 함수 종속성은 테이블 속성들 간의 **의존 관계**를 나타내는 개념
- 어떤 속성 A의 값을 알면 다른 속성 B의 값이 유일하게 정해지는 의존 관계를  $A \rightarrow B$ 로 표기하며, A를 B의 결정자
- 함수 종속성을 파악하기 위해서는 테이블의 각 튜플을 식별할 수 있는 데이터 확인

| 학생번호 | 학생이름 | 주소       | 학과   | 학과사무실  | 강좌이름   | 강의실     | 성적  |
|------|------|----------|------|--------|--------|---------|-----|
| 501  | 박지성  | 영국 맨체스타  | 컴퓨터과 | 공학관101 | 데이터베이스 | 공학관 110 | 3.5 |
| 401  | 김연아  | 대한민국 서울  | 체육학과 | 체육관101 | 데이터베이스 | 공학관 110 | 4.0 |
| 402  | 장미란  | 대한민국 강원도 | 체육학과 | 체육관101 | 스포츠경영학 | 체육관 103 | 3.5 |
| 502  | 추신수  | 미국 클리블랜드 | 컴퓨터과 | 공학관101 | 자료구조   | 공학관 111 | 4.0 |
| 501  | 박지성  | 영국 맨체스타  | 컴퓨터과 | 공학관101 | 자료구조   | 공학관 111 | 3.5 |

### 3. 함수 종속성과 이상현상

---

- 이상현상은 하나의 테이블에 두 개 이상의 정보, 즉 기본키를 제외한 다른 속성에 함수 종속이 포함될 때 나타남
- 기본키는 일반적으로 모든 속성을 함수적으로 결정하는 결정자이지만 함수 종속성을 판단할 때는 각 속성의 특성과 의미를 고려
- 함수 종속성 다이어그램을 통해 함수 종속 관계를 파악하고 정규화 진행



## 4. 정규화

- 정규화<sup>Normalization</sup>는 테이블을 적절하게 나누어서 중복을 최소화하고, 데이터의 일관성과 무결성을 유지하기 위한 설계 과정
- 정규화는 함수 종속성을 이용해 테이블을 연관성 있는 속성으로만 구성해 이상 현상이 발생하지 않는 테이블로 만들어가는 과정

| 정규화 유형 | 설명                            |
|--------|-------------------------------|
| 제1정규형  | 각 열은 원자값 가짐                   |
| 제2정규형  | 1NF + 부분적 종속성 제거              |
| 제3정규형  | 2NF + 이행적 종속성 제거              |
| BCNF형  | 3NF + 결정자가 후보키가 아닌 함수적 종속성 제거 |
| 제4정규형  | BCNF + 다치 종속성 제거              |
| 제5정규형  | 4NF + 조인 종속성 제거               |

## 5. 제1정규형

---

- 테이블의 모든 속성 값이 더 이상 분해되지 않는 원자값을 가지면 제1정규형(1NF)
- 제1정규형을 만족하기 위해서는 속성의 도메인이 원자 값으로만 구성되도록 테이블을 재구성

| 이름  | 취미     |
|-----|--------|
| 김연아 | 인터넷    |
| 추신수 | 영화, 음악 |
| 박세리 | 음악, 쇼핑 |
| 장미란 | 음악     |
| 박지성 | 게임     |



| 이름  | 취미  |
|-----|-----|
| 김연아 | 인터넷 |
| 추신수 | 영화  |
| 추신수 | 음악  |
| 박세리 | 음악  |
| 박세리 | 쇼핑  |
| 장미란 | 음악  |
| 박지성 | 게임  |

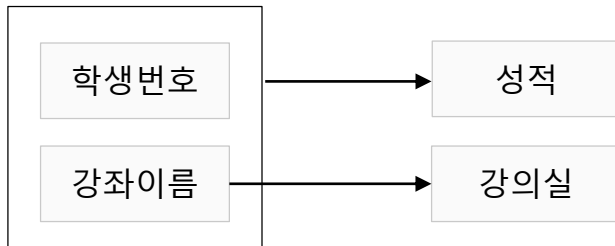
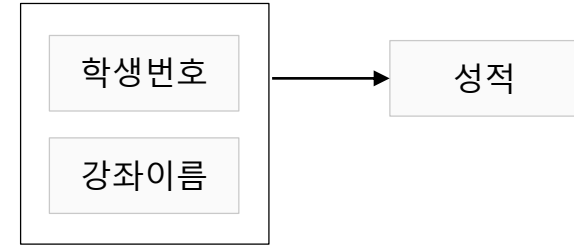
## 6. 제2정규형

- 제2정규형은 릴레이션이 제1정규형을 만족하고 기본키가 아닌 속성이 기본키에 완전 함수 종속일 때 2NF를 만족
- 완전 함수 종속은 모든 속성이 기본키 또는 기본키가 여러 속성으로 이루어진 복합키에만 종속된 경우
- 제2정규형을 만족하기 위해서는 부분 함수 종속을 제거

| 학생번호 | 강좌이름   | 강의실     | 성적  |
|------|--------|---------|-----|
| 501  | 데이터베이스 | 공학관 110 | 3.5 |
| 401  | 데이터베이스 | 공학관 110 | 4.0 |
| 402  | 스포츠경영학 | 체육관 103 | 3.5 |
| 502  | 자료구조   | 공학관 111 | 4.0 |
| 501  | 자료구조   | 공학관 111 | 3.3 |



| 학생번호 | 강좌이름   | 성적  |
|------|--------|-----|
| 501  | 데이터베이스 | 3.5 |
| 401  | 데이터베이스 | 4.0 |
| 402  | 스포츠경영학 | 3.5 |
| 502  | 자료구조   | 4.0 |
| 501  | 자료구조   | 3.3 |



| 강좌이름   | 강의실     |
|--------|---------|
| 데이터베이스 | 공학관 110 |
| 스포츠경영학 | 체육관 103 |
| 자료구조   | 공학관 111 |





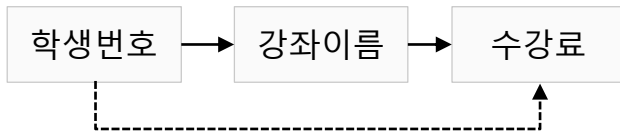
## 7. 제3정규형

- 제3정규형은 릴레이션이 제2정규형이고, 기본키가 아닌 모든 속성이 기본키에 이행적 함수 종속이 되지 않으면 3NF 만족
- 이행적 함수 종속이란  $A \rightarrow B, B \rightarrow C$ 가 성립할 때  $A \rightarrow C$ 가 성립되는 함수 종속성
- 제3정규형을 만족하기 위해서는 모든 속성이 이행적 함수 종속이 되지 않도록 분해

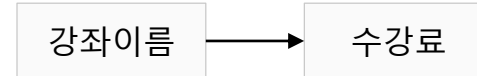
| 학생번호 | 강좌이름   | 수강료   |
|------|--------|-------|
| 501  | 데이터베이스 | 20000 |
| 401  | 데이터베이스 | 20000 |
| 402  | 스포츠경영학 | 15000 |
| 502  | 자료구조   | 25000 |



| 학생번호 | 강좌이름   |
|------|--------|
| 501  | 데이터베이스 |
| 401  | 데이터베이스 |
| 402  | 스포츠경영학 |
| 502  | 자료구조   |



| 강좌이름   | 수강료   |
|--------|-------|
| 데이터베이스 | 20000 |
| 스포츠경영학 | 15000 |
| 자료구조   | 25000 |



# 8. BCNF

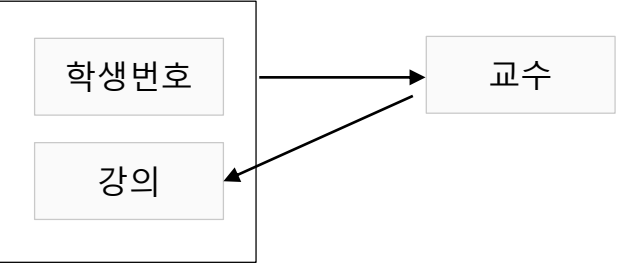
- BCNF<sup>Boyce Codd Normal Form</sup> 는 릴레이션에서 함수 종속성  $X \rightarrow Y$ 가 성립할 때 모든 결정자 X가 후보키인 정규형
- BCNF는 릴레이션을 엄격하게 규정하여 데이터 이상현상을 방지하지만, 데이터베이스의 성능과 설계 복잡성을 초래

| 학생번호 | 강의       | 교수  |
|------|----------|-----|
| 501  | 프로그래밍 언어 | 김유신 |
| 401  | 프로그래밍 언어 | 김유신 |
| 402  | 컴퓨터 구조론  | 김춘추 |
| 502  | 알고리즘     | 강감찬 |
| 501  | 알고리즘     | 이순신 |



| 학생번호 | 강좌       |
|------|----------|
| 501  | 프로그래밍 언어 |
| 401  | 프로그래밍 언어 |
| 402  | 컴퓨터 구조론  |
| 502  | 알고리즘     |
| 501  | 알고리즘     |

|      |    |
|------|----|
| 학생번호 | 강의 |
|------|----|



| 강좌       | 교수  |
|----------|-----|
| 프로그래밍 언어 | 김유신 |
| 컴퓨터 구조론  | 김춘추 |
| 알고리즘     | 강감찬 |
| 알고리즘     | 이순신 |



## 9. 정규화 정리

---

- 일반적으로 모든 릴레이션은 제3정규형까지 진행하면 실제적인 이상현상은 발생하지 않음
- 과도한 정규화로 인해 SQL 처리 성능이 떨어져 서비스 개발과 운영에 문제가 될 경우에는 반정규화 진행

