

1. 논리 데이터저장소 확인하기

7. 함수 종속(FD : Functional Dependency)

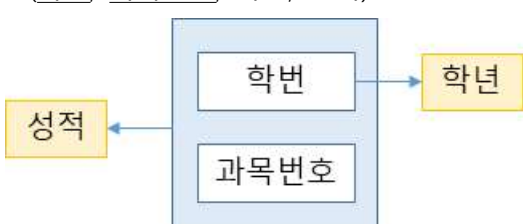
(1) 정의

- 어떤 테이블에 컬럼 X와 컬럼 Y가 있을 경우 컬럼 X의 값이 컬럼 Y의 오직 하나의 값과 연관되어 있다면 Y는 X에 함수 종속이라 하고, $X \rightarrow Y$ 로 표기한다.
- 또한 이때 X를 결정자라고 하고 Y를 종속자라고도 한다.

| | |
|--|---|
| <p>[예시1] 학생 (학번, 이름, 학년, 학과)</p> <p>학생 테이블에 학번, 이름, 학년, 학과 컬럼이 있다. 어떤 학생의 학번이 정해지면 그 학번에 대응하는 이름, 학년, 학과는 오직 하나의 값만 가지게 된다. 그러므로 이름, 학년, 학과 컬럼은 학번 컬럼에 함수 종속이라고 말할 수 있다.</p> <p>이름 기호로 표현하면, 다음과 같다.</p> | <p>학번 \rightarrow 이름</p> <p>학번 \rightarrow 학년</p> <p>학번 \rightarrow 학과</p> <p>또는</p> <p>학번 \rightarrow (이름, 학년, 학과)</p> |
|--|---|

(2) 함수 종속 다이어그램

- 한 테이블에서 컬럼들 간의 복잡한 함수 종속 관계를 쉽게 이해하기 위해 도식화하여 표현한다.

| | |
|---|---|
| <p>[예시2] 수강 (학번, 과목번호, 학년, 성적)</p>  <p><수강 테이블의 함수 종속 다이어그램></p> | <p>(학번, 과목번호) \rightarrow 성적</p> <p>학번 \rightarrow 학년</p> |
|---|---|

(3) 완전 함수 종속과 부분 함수 종속

- 성적 컬럼은 기본키인 (학번, 과목번호)에 함수 종속이며, 학년 컬럼은 학번 컬럼에만 함수 종속이 되어 있다. 그러나 학번 컬럼이 기본키인 (학번, 과목번호) 복합 컬럼에 포함되어 있으므로 당연히 학년 컬럼이 (학번, 과목번호) 컬럼에도 함수 종속되어 있다고 말할 수 있다.

(학번, 과목번호) \rightarrow 학년

- 이러한 경우 성적 컬럼은 기본키인 (학번, 과목번호) 복합 컬럼에 완전 함수 종속이라고 하며, 학년 컬럼은 기본키인 (학번, 과목번호) 복합 컬럼에 부분 함수 종속이라고 한다.

8. 정규화(Normalization)

(1) 정의

논리 모델링 단계에서 수행하며, 중복성을 최소화하고 정보의 일관성을 보장하기 위한 작업

(2) 목적

(가) 데이터 중복 배제로 데이터 관리 편의성 및 자료 저장 공간의 최소화

(나) 데이터 모형 단순화

(다) 데이터 구조의 안정성 및 무결성 유지

(라) 속성의 배열상태 검증

(마) 엔티티와 속성의 누락 여부 검증 수단

(바) 자료검색과 추출의 효율성을 추구

(3) 특징

(가) 어떠한 관계구조가 바람직한 것인지, 바람직하지 못한 관계를 어떻게 분해하여야 하는지에 관한 구체적인 판단기준을 제공

(나) 정규화된 데이터 모델은 정확성, 일치성, 단순성, 비중복성, 안정성 보장

(4) 유형

| | |
|-------|--|
| 제1정규화 | 1) 반복되는 속성이나 Group 속성 제거 2) 새로운 실체와 1:N의 관계 추가 3) 모든 속성은 반드시 하나의 값을 가져야 함(반복형태가 있어서는 안됨) |
| 제2정규화 | 1) 주식별자에 완전하게 종속되지 않는 속성 제거 2) 불완전 함수적 종속(Non Fully Dependency) 제거 3) 모든 속성은 반드시 UID전부에 종속되어야 함(UID일부에만 종속되어서는 안됨) |
| 제3정규화 | 1) 비식별자에 종속되는 속성 제거 2) 주식별자에 이행종속(Transitive Dependency) 되는 속성 제거 3) UID가 아닌 모든 속성 간에는 서로 종속될 수 없음(속성간 종속성 배제) |
| 제4정규화 | 1) 실제로 거의 고려되지 않는 정규화 2) 주식별자에 다가종속(Multi-Valued Dependency)되는 속성을 두가지 이상 두지 않음 3) 2차 정규화된 테이블은 다대다 관계를 가질 수 없음 4) 어떠한 관계구조가 바람직한 것인지, 바람직하지 못한 관계를 어떻게 분해하여야 하는지에 관한 구체적인 판단기준을 제공 |

(5) 제1정규형(1NF : First Normal Form)

어떤 테이블에 속한 모든 값들이 원자 값으로 되어 있다면 제1정규형에 속한다.

| 비정규 릴레이션 | | | |
|----------|------|--------------|--------|
| 학번 | 이름 | 과목명 | 성적 |
| 100 | King | Oracle, Java | 90, 85 |
| 101 | Abel | Java, Linux | 80, 95 |

→

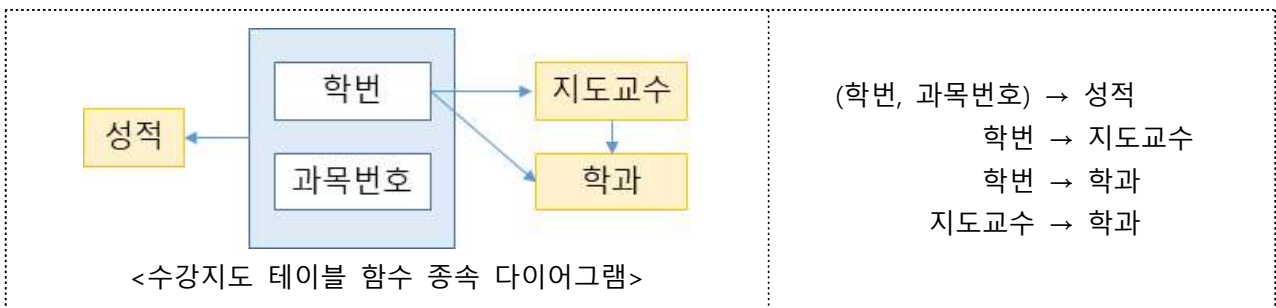
| 원자값으로 구성된 1NF 테이블 | | | |
|-------------------|------|--------|----|
| 학번 | 이름 | 과목명 | 성적 |
| 100 | King | Oracle | 90 |
| 101 | Abel | Java | 80 |
| 101 | Abel | Linux | 95 |
| 100 | King | Java | 85 |

2001020205_16v4 데이터 입출력 구현

- 대부분의 테이블은 제1정규형에 속한다. 그러나 제1정규형에만 속해 있는 테이블은 여러 가지 이유에서 바람직하지 않은 구조가 될 수 있다.

<수강지도 테이블>

| 학번 | 지도교수 | 학과 | 과목번호 | 성적 |
|-----|------|-----|------|----|
| 100 | P1 | 컴퓨터 | C413 | A |
| 100 | P1 | 컴퓨터 | E412 | A |
| 200 | P2 | 전기 | C123 | B |
| 300 | P3 | 컴퓨터 | C312 | A |
| 300 | P3 | 컴퓨터 | C324 | C |
| 300 | P3 | 컴퓨터 | C413 | A |
| 400 | P1 | 컴퓨터 | C312 | A |
| 400 | P1 | 컴퓨터 | C324 | A |
| 400 | P1 | 컴퓨터 | C413 | B |
| 400 | P1 | 컴퓨터 | E412 | C |



- 학번과 과목번호의 조합으로 학생이 수강한 과목의 성적을 식별할 수 있다.
- 한 학생은 한 지도교수를 가질 수 있고, 한 학과에만 속한다.
- 또한 각 지도교수는 한 학과에만 속한다.
- 수강지도 테이블의 저장된 값들을 보면 이 테이블은 지도교수와 학과 컬럼에 중복된 데이터가 많은 것을 볼 수 있다. 이로 인해 여러 가지 이상 현상이 발생할 수 있다.

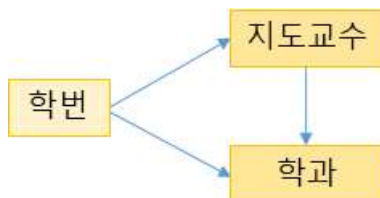
| | |
|-------|--|
| 삽입 이상 | - 예를 들어, 학번이 500인 학생의 지도교수가 P4라는 사실만 삽입할 수 없다. 왜냐하면 학번이 500인 학생이 어떤 교과목을 등록하지 않으면 기본 키에 속하는 과목번호가 널이 되므로 제약조건 위배로 인해 작업이 수행되지 않기 때문이다. |
| 삭제 이상 | - 만일 학번이 200인 학생이 교과목 C123의 등록을 취소하면 교과목(기본키)만 삭제가 되지 않으므로 학번이 200인 학생의 행 전체를 삭제해야 한다. 이렇게 되면 이 학생의 지도교수가 P2라는 정보까지도 손실되므로 삭제 이상이 발생된다. |
| 갱신 이상 | - 학번이 400인 학생의 지도교수가 P1에서 P3로 변경되었다면, 4개의 행에 대해 지도교수 데이터 값을 P3으로 변경해 주어야 한다. 만일 일부만 변경되면 400번 학생의 지도교수가 P1과 P3, 즉 둘이 되어 데이터의 일관성이 떨어지게 되는 갱신 이상이 발생된다. |

2001020205_16v4 데이터 입출력 구현

- 이러한 문제가 발생하는 원인은 수강지도 테이블에서 기본키가 아닌 컬럼들이 기본키에 완전 함수 종속되지 못하고 부분 함수 종속이 되기 때문이다. 따라서 이러한 문제들에 대한 해결은 수강지도 테이블을 두 개의 테이블로 분할하여 부분 함수 종속을 제거하는 것이다.
- 부분 함수 종속을 제거하기 위해 수강지도 테이블을 지도 테이블과 수강 테이블로 재구성하여 제1정규형에서 일어나는 문제를 해결하였다.
- 한 테이블에 부분 함수 종속이 존재한다는 것은 기본키로 식별되는 개체와 무관한 컬럼이 있다는 것이다. 즉, 두 가지 종류의 상이한 정보를 하나의 테이블에 혼합시켜 놓은 것이므로 정규화를 거쳐야 이상 현상을 막을 수 있다.

[테이블1] 지도 (학번, 지도교수, 학과)

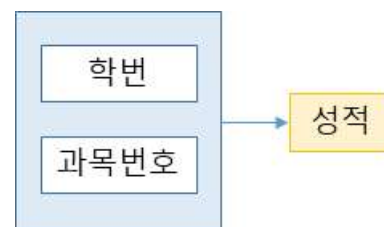
학번 → 지도교수
학번 → 학과
지도교수 → 학과



<지도 테이블 함수 종속 다이어그램>

[테이블2] 수강 (학번(FK), 과목번호, 성적)

(학번, 과목번호) → 성적



<수강 테이블 함수 종속 다이어그램>

(6) 제2정규형(2NF)

어떤 테이블이 1NF이고 기본키에 속하지 않은 컬럼 모두가 기본키에 완전 함수 종속이면 제2정규형에 속한다.

<지도 테이블>

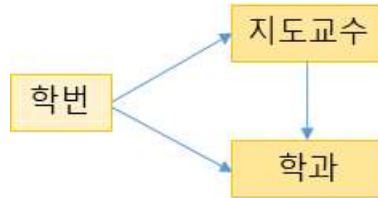
| 학번 | 지도교수 | 학과 |
|-----|------|-----|
| 100 | P1 | 컴퓨터 |
| 200 | P2 | 전기 |
| 300 | P3 | 컴퓨터 |
| 400 | P1 | 컴퓨터 |

<수강 테이블>

| 학번(FK) | 과목번호 | 성적 |
|--------|------|----|
| 100 | C413 | A |
| 100 | E412 | A |
| 200 | C123 | B |
| 300 | C312 | A |
| 300 | C324 | C |
| 300 | C413 | A |
| 400 | C312 | A |
| 400 | C324 | A |
| 400 | C413 | B |
| 400 | E412 | C |

2001020205_16v4 데이터 입출력 구현

- 지도 테이블과 수강 테이블은 모두 2NF에 속한다.
- 지도 테이블과 수강 테이블의 함수 종속 다이어그램을 살펴보면, 수강 테이블은 문제가 없지만 지도 테이블은 복잡한 상호 함수 종속 관계를 나타내고 있다.



<지도 테이블 함수 종속 다이어그램>

- 학과는 학번에 완전 함수 종속이면서 지도교수를 통해 이행적 함수 종속이 되고 있다. 즉, 학번은 지도교수를 결정하고 이 지도교수는 학과를 결정한다.

학번 → 지도교수, 지도교수 → 학과 =====> 학번 → 학과

- 학과는 학번에 이행적 함수 종속이라고 한다. 이러한 이행적 함수 종속을 가진 지도 테이블에 나타나는 이상 현상은 다음과 같다.

<지도 테이블>

| 학번 | 지도교수 | 학과 |
|-----|------|-----|
| 100 | P1 | 컴퓨터 |
| 200 | P2 | 전기 |
| 300 | P3 | 컴퓨터 |
| 400 | P1 | 컴퓨터 |

| | |
|-------|---|
| 삽입 이상 | - 지도 테이블에 어떤 지도교수가 어떤 특정 학과에 속한다는 사실만을 삽입하려고 할 때 삽입을 할 수 없다. 왜냐하면 지도교수의 지도를 받는 학생 정보가 없다면 기본키가 널이 되기 때문에 삽입 이상이 발생된다. |
| 삭제 이상 | - 학번이 300인 학생의 행이 삭제되는 경우 지도교수 P3이 컴퓨터 학과에 속한다는 정보도 손실이 되는 삭제 이상이 발생된다. |
| 갱신 이상 | - 지도교수 P1의 학과가 컴퓨터에서 전자로 바뀐다면 학번 100과 400에 있는 학과의 값을 모두 변경시켜야 한다. 그렇지 않으면 지도교수 P1은 두 개의 상이한 학과 값을 갖게 되어 데이터 모순이 발생된다. |

- 이러한 이상은 결국 두 개의 상이한 정보를 하나의 테이블로 혼합해서 표현하려고 하여 발생된다.
- 이행적 함수 종속은 복수의 관계를 하나의 테이블로 표현하는 데서 나타나는 것이다. 그러므로 해결 방법은 이행적 함수 종속을 제거하여 두 개의 테이블로 분해하는 것이다.

[테이블1] 학생지도 (학번, 지도교수(FK))



<학생지도 테이블 함수 종속 다이어그램>

[테이블2] 지도교수학과 (지도교수, 학과)



<지도교수 테이블 함수 종속 다이어그램>

2001020205_16v4 데이터 입출력 구현

(7) 제3정규형(3NF)

어떤 테이블이 2NF이고 기본키에 속하지 않은 컬럼 모두가 기본키에 이행적 함수 종속이 아니면 제3정규형에 속한다.

<학생지도 테이블>

| 학번 | 지도교수(FK) |
|-----|----------|
| 100 | P1 |
| 200 | P2 |
| 300 | P3 |
| 400 | P1 |

<지도교수학과 테이블>

| 지도교수 | 학과 |
|------|-----|
| P1 | 컴퓨터 |
| P2 | 전기 |
| P3 | 컴퓨터 |

- 학생지도 테이블과 지도교수학과 테이블은 3NF이다.

(8) 보이시/코드 정규형(Boyce/Codd Normal Form(BCNF), 강한 제3정규형)

어떤 테이블의 모든 결정자가 후보키(기본키)이면 보이시/코드 정규형(BCNF)에 속한다.
BCNF에 속하는 테이블은 모두 제3정규형에 속하지만 반대는 성립하지 않는다.
그런 의미에서 BCNF는 제3정규형보다 강력하므로 "강한 제3정규형"이라고도 한다.

- 앞에서 보았던 테이블 중 수강지도 테이블, 지도 테이블은 3NF도 아니고 BCNF도 아니다.

| <수강지도 테이블> | <지도 테이블> |
|---|--|
| 수강지도 (학번, 지도교수, 학과, 과목번호, 성적) (학번, 과목번호) → 성적 학번 → 지도교수 학번 → 학과 지도교수 → 학과 (학번, 과목번호), 학번, 지도교수가 결정자이지만 (학번, 과목번호)만 기본키이므로 BCNF가 아니다. | 지도 (학번, 지도교수, 학과) 학번 → 지도교수 학번 → 학과 지도교수 → 학과 학번, 지도교수가 결정자이지만 학번만 기본키이므로 BCNF가 아니다. |

- 그러나 수강 테이블, 학생지도 테이블, 지도교수학과 테이블은 3NF이면서 BCNF에 속한다.

| <수강 테이블> | <학생지도 테이블> | <지도교수학과 테이블> |
|--|--------------------------------------|------------------------------------|
| 수강 (학번(FK), 과목번호, 성적) (학번, 과목번호) → 성적 | 학생지도 (학번, 지도교수(FK)) 학번 → 지도교수 | 지도교수학과 (지도교수, 학과) 지도교수 → 학과 |
| 위 세 테이블은 결정자가 모두 기본키이므로 BCNF에 속한다. | | |

- 3NF이지만 BCNF가 아닌 테이블인 경우 여러 이상 현상이 발생할 수 있다.
- 3NF이지만 BCNF가 아닌 테이블은 다음과 같다.
- 수강과목 테이블은 한 학생은 여러 과목을 수강할 수 있으므로 학번 컬럼 하나로는 기본키가 될 수

2001020205_16v4 데이터 입출력 구현

없다. 그러므로 (학번, 과목) 또는 (학번, 교수)가 기본키가 될 수 있다.

- (학번, 과목)을 기본키로 결정하였다.
- 또한 교수는 한 과목만 담당할 수 있고, 하나의 과목은 여러 교수가 담당할 수 있다.

수강과목 (학번, 과목, 교수)

(학번, 과목) → 교수

교수 → 과목

```
graph LR; subgraph Table [수강과목]; direction TB; A[학번]; B[과목]; end; C[교수]; A --> C; C --> B;
```

<수강과목 테이블 함수 종속 다이어그램>



<수강과목 테이블>

| 학번 | 과목 | 교수 |
|-----|--------|----|
| 100 | 프로그래밍 | P1 |
| 100 | 데이터베이스 | P2 |
| 200 | 프로그래밍 | P1 |
| 200 | 데이터베이스 | P3 |
| 300 | 데이터베이스 | P3 |
| 300 | 프로그래밍 | P4 |

- 수강과목 테이블은 1NF, 2NF, 3NF이나 BCNF는 아니다. 왜냐하면 (학번, 과목)과 교수 컬럼이 결정자이지만 교수 컬럼이 기본키로 취급되고 있지 않기 때문이다.
- 수강과목 테이블에서 발생하는 이상 현상은 다음과 같다.

| | |
|-------|---|
| 삽입 이상 | - 교수 P5가 데이터베이스를 담당하게 되었을 때 이 사실만을 삽입할 수 없다. 적어도 수강 학생이 한 사람 있어야 가능하다. |
| 삭제 이상 | - 학번이 100인 학생이 데이터베이스 과목을 취소하여 테이블로부터 행이 삭제된다면 교수 P2가 데이터베이스 과목을 담당하고 있다는 정보마저 없어지게 되는 삭제 이상이 발생된다. |
| 갱신 이상 | - 교수 P1의 담당과목이 프로그래밍에서 데이터베이스로 변경된다면 P1에 해당되는 모든 행을 변경해 주어야 한다. 그렇지 않으면 교수는 한 과목만 담당하기로 되어 있으나 사실과 달리 P1이 두 과목을 담당한 것이 되어 데이터 모순이 발생한다. |

- 이와 같은 이상 현상의 원인은 교수 컬럼이 결정자이지만 기본키로 취급되고 있지 않기 때문이다.
- 문제 해결을 위해 다음과 같이 수강교수 테이블과 과목교수 테이블로 분해하였다.

| | |
|--|---|
| <p>[테이블1] 수강교수 (학번, 교수(FK))</p>  | <p>[테이블2] 과목교수 (교수, 과목)</p>  <p><과목교수 테이블 함수 종속 다이어그램></p> |
|--|---|

- 수강교수 테이블과 과목교수 테이블은 제3정규형에서 일어날 수 있는 이상 현상을 제거한 BCNF에 속한다.