

정규화

정의

이상현상(anomaly)

정규화의 단계

함수 종속

정규화 상세단계

정의

관계형 데이터베이스 모델에서 데이터의 중복성을 제거하여 이상 현상을 방지
데이터의 일관성과 정확성을 유지하기 위해 무손실 분해하는 과정이다.

이상현상(anomaly)

이상 현상은 데이터의 중복성으로 인해 릴레이션을 조작할 때 발생하는 비합리적 현상이다.
삽입, 삭제, 갱신 이상이 있다.

- 삽입 이상: 정보 저장시 해당 정보의 불필요한 세부정보를 입력하는 경우
- 삭제 이상: 정보 삭제시 원치 않는 다른 정보가 같이 삭제되는 경우
- 갱신 이상 : 중복 데이터 중에서 특정 부분만 수정되어 중복된 값이 모순을 일으키는 경우

정규화의 단계

- 1정규형 : 원자값으로 구성
- 2정규형: 부분 함수 종속 제거
- 3정규형: 이행 함수 종속 제거
- BCNF: 결정자 후보 키가 아닌 함수 종속 제거

함수 종속

함수 종속 개념: 릴레이션에서 속성의 의미와 속성 간 상호 관계로부터 발생하는 제약조건이다.

결정자/종속자: $X \rightarrow Y$ 관계일때, X는 결정자, Y는 종속자이다.

함수 종속 종류:

- 부분 함수 종속: 릴레이션에서 기본 키가 복합 키일 경우 기본 키를 구성하는 속성 중 일부에게 종속된 경우
- 완전 함수 종속: 릴레이션에서 $X \rightarrow Y$ 관계가 있을때, Y는 전체 속성에 대해 종속하고, 부분 집합 속성에 종속하지 않는 경우
- 이행 함수 종속: 릴레이션에서 $X \rightarrow Y, Y \rightarrow Z$ 인 관계가 있을 때, $X \rightarrow Z$ 가 성립되는 경우

정규화 상세단계

1. 1차 정규화

고객ID	이메일
1	1@example.com,2@example.com

고객ID	이메일
1	1@example.com
1	2@example.com

테이블 내 속성값은 원자값을 가지고 있어야 한다.

속성 1개만 가지고 있도록 저장하면 1차 정규화를 만족한다.

2. 2차 정규화

부분 함수 종속성을 제거한다. 즉, 복합키의 일부에 종속된 속성을 제거한다.

고객명	서비스이름	서비스가격	서비스이용기간
김지용	헬스	70000	1달
김지용	수영	100000	2달
오유진	수영	100000	2달

→

고객명	서비스이름	서비스이용기간	서비스가격
김지용	헬스	1달	70000
김지용	수영	2달	100000
오유진	수영	2달	

<고객명,서비스 이름> 이 <서비스 이용기간>에 영향을 주고, <서비스 이름>이 <서비스가격>에 영향을 주는 관계를

부분함수 종속 관계라고 한다.

이렇게 부분함수 종속 관계를 한 테이블에 넣는 것은 2차 정규화를 만족하지 않는다.

부분 관계인 <서비스 이름, 서비스 가격> 관계를 별도의 테이블로 두면 부분함수 종속 관계가 제거되어 2차 정규화를 만족한다.

3. 3차 정규화

기본키에 의해 결정되지 않은 속성이 다른 비기본키 속성을 결정하지 않아야 한다.

책번호	도서이름	글쓴이	출판사	홈페이지
1	배터리	오유진	A사	O.COM
2	자바스크립트	최성호	B사	C.COM
3	국정원	김지용	B사	C.COM

→

책번호	도서이름	출판사	홈페이지
1	배터리	오유진	O.COM
2	자바스크립트	최성호	C.COM
3	국정원	김지용	

<책번호>가 <출판사>에 영향을 주고, <출판사>가 <홈페이지>에 영향을 주는, $A \rightarrow B$, $B \rightarrow C$ 이면서 $A \rightarrow C$ 관계가 같이 있는 경우를 이행함수 종속 관계라고 한다.

<책번호, 출판사, 홈페이지>를 한 테이블에 두는 것은 이행함수 종속성으로 인해 3차 정규화를 만족시키지 못한다.

<책번호>는 <홈페이지>에 직접 영향을 주는 관계가 되면 안되므로, 이행함수 종속 관계를 제거하여 3차 정규화를 만족시킨다.

4. 보이스-코드 정규화(BCNF)

모든 결정자가 후보키어야 한다.

학번	과목명	교수명
1	C언어	유승규
1	데이터베이스	송아람
2	데이터베이스	송아람

여기서 기본키는 <학번,과목명>으로 , 함수적 종속성을 따지면 과목명 → 교수명으로 종속성을 가지고 있다.

<학번>이 <교수명>을 종속하지 않으므로 제 3정규형을 만족하지만,
기본키가 아닌 <과목명>이 <교수명>의 결정자이므로 BCNF를 만족하지 않는다.

→

학번	과목명	과목명	교수명
1	C언어	C언어	유승규
1	데이터베이스	데이터베이스	송아람
2	데이터베이스		