

16강 정규화

정규화의 목적

• 정규화(Normalization)

1. 관계형 데이터베이스를 설계할 때
2. 데이터의 중복을 최소화하여 저장 효율을 높이고 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

정규화의 장점

- 데이터베이스의 일관성을 향상시킬 수 있다.
- 데이터베이스의 확장성을 보장할 수 있다.
- 데이터베이스의 논리적 구조를 견고하게 만들 수 있다

어노말리(이상현상)

- 정규화를 진행하지 않은 경우 발생
- 릴레이션을 조작할 때 이상 (Anomaly)현상이 발생하게 된다.
 1. 삽입이상 (insertion Anomaly)
 2. 삭제이상 (Deletion Anomaly)
 3. 갱신이상 (Update Anomaly)

삽입 이상 (Insertion Anomaly)

- 데이터를 삽입할 때 불필요한 데이터가 함께 삽입되는 현상으로, 제 1정규형에 문제가 있는 경우 발생하는 현상
- 다른 데이터가 존재하지 않아 원하는 데이터를 입력할 수 없는 것이 삽입 이상 현상이다.

삭제 이상 (Deletion Anomaly)

- 릴레이션의 한 튜플을 삭제함으로써 연쇄 삭제로 인해 정보의 손실이 발생하는 현상

갱신 이상 (Update Anomaly)

- 튜플 중에서 일부 튜플의 값만을 갱신함으로써 정보의 모순성이 발생하는 현상
- 이유는 데이터의 중복

이상현상예

- 이상 현상의 발생 예

#선수번호	선수이름	포지션	#팀번호	팀이름
10	홍길동	유격수	25	라이거스
14	강길동	투수	25	라이거스
15	박길동	중견수	25	라이거스
13	이길동	포수	20	라이언스
14	강길동	투수	10	타이언스
12	김길동	1루수	05	타이거스
10	홍길동	유격수	10	타이언스
11	최길동	2루수	10	타이언스
12	김길동	1루수	15	엘리펀츠
10	홍길동	유격수	30	캣츠

이상현상예

- 삽입 이상: 장길동이라는 새로운 선수를 등록할 때

#선수번호	선수이름	포지션	#팀번호	팀이름
10	홍길동	2루수	25	라이거스
14	강길동	투수	25	라이거스
15	박길동	중견수	25	라이거스
13	이길동	포수	20	라이언스
14	강길동	투수	10	타이언스
12	김길동	1루수	05	타이거스
10	홍길동	유격수	10	타이언스
11	최길동	2루수	10	타이언스
12	김길동	1루수	15	엘리펀츠
10	홍길동	유격수	30	캣츠
17	장길동		???	???

입력 불가 오류

이상현상예

- 수정 이상 : 홍길동을 2루수로 바꿀때

#선수번호	선수이름	포지션	#팀번호	팀이름
10	홍길동	2루수	25	라이거스
14	강길동	투수	25	라이거스
15	박길동	중견수	25	라이거스
13	이길동	포수	20	라이언스
14	강길동	투수	10	타이언스
12	김길동	1루수	05	타이거스
10	홍길동	유격수	10	타이언스
11	최길동	2루수	10	타이언스
12	김길동	1루수	15	엘리펀츠
10	홍길동	유격수	30	캣츠

입력 불가 오류

데이터 오류

이상현상예

- 삭제 이상 : 김길동 삭제시 팀 정보 삭제

#선수번호	선수이름	포지션	#팀번호	팀이름
10	홍길동	유격수	25	라이거스
14	강길동	투수	25	라이거스
15	박길동	중견수	25	라이거스
13	이길동	포수	20	라이언스
14	강길동	투수	10	타이언스
12	김길동	1루수	05	타이거스
10	홍길동	유격수	10	타이언스
11	최길동	2루수	10	타이언스
12	김길동	1루수	15	엘리펀츠
10	홍길동	유격수	30	캣츠

삭제 오류

어노말리(Anomaly)를 해결하는 방법 => 정규화

[선수]

#선수번호	선수이름	포지션
10	홍길동	유격수
14	강길동	투수
15	박길동	중견수
13	이길동	포수
12	김길동	1루수
11	최길동	2루수

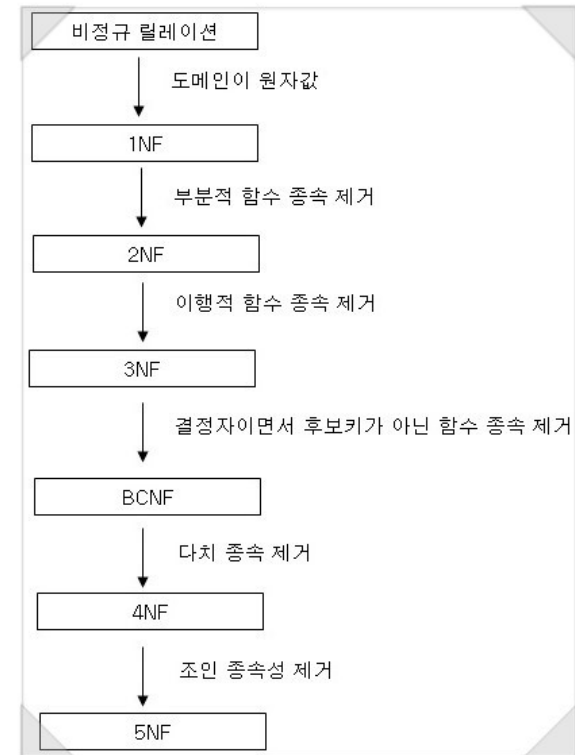
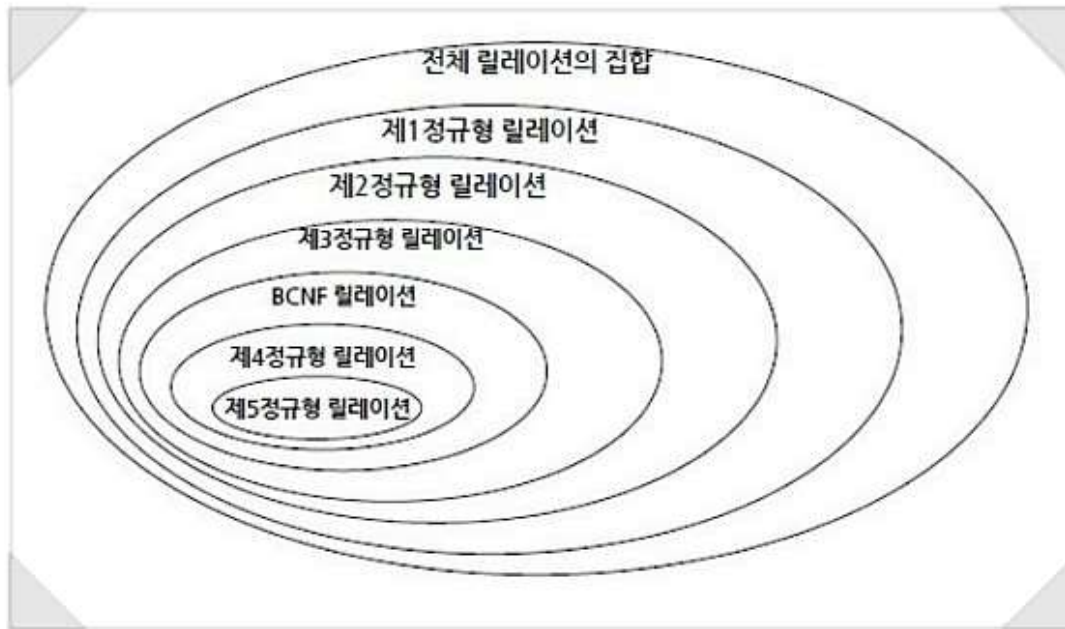
[소속]

#팀번호	#팀번호
10	25
14	25
15	25
13	20
14	10
12	05
10	10
11	10
12	15
10	30

[팀]

#팀번호	팀이름
25	라이거스
20	라이언스
10	타이언스
5	타이거스
15	엘리펀츠
30	캣츠

정규형의 구조와 단계



제1 정규형

- 모든 속성은 반드시 하나의 값을 가져야 한다
- 값이라는 것은 원자성(ATOM)을 가져야 한다. 즉 더 이상 쪼갤 수 없는 하나의 값만을 가져야 한다
 - 다가 속성
 - 복합 속성

제1 정규형

• 다가 속성

#고객번호	고객명	주민번호	전화번호
100	홍길동	123456-7890123	123-4568, 234-5678, 345-6789
101	황진이	234567-8901234	456-7890, 567-8901

- 다가 속성은 같은 종류의 값을 여러 개 가지는 속성을 의미
- 고객 번호를 알더라도 유일하게 식별할 수 있는 전화번호가 없음
- 따라서 전화번호는 주 식별자인 고객번호에 함수적으로 종속되지 않았다.
- 따라서 한 속성에는 반드시 하나의 값을 가져야 한다는 1정규형에 어긋나는 릴레이션이다.
- FD: 고객번호->고객명
- FD1: 고객번호-> 주민번호
- FD2: 고객번호-> 전화번호 ????? <여러개 존재>

제1 정규형

- 다가 속성

- 한 속성이 하나의값을 가지도록 풀어보면

#고객번호	고객명	주민번호	전화번호
100	홍길동	123456-7890123	123-4568
100	홍길동	123456-7890123	234-5678
100	홍길동	123456-7890123	345-6789
101	황진이	234567-8901234	456-7890
101	황진이	234567-8901234	567-8901

- 고객번호, 고객명, 주민번호 중복을 제거하면

#고객번호	고객명	주민번호
100	홍길동	123456-7890123
101	황진이	234567-8901234

#고객번호	전화번호
100	123-4568
100	234-5678
100	345-6789
101	456-7890
101	567-8901

제1 정규형

- 복합속성

#고객번호	고객명	주민번호
100	홍길동	123456-7890123

- 고객명은 성과 이름으로 구성된 복합 속성
- 만일 고객의 성 또는 이름을 따로 조회해야 하는 경우가 업무에 효율적이라고 판단된다면

#고객번호	고객성	고객명	주민번호
100	홍	길동	123456-7890123

- 복합 속성 예시 : 주민번호 123456-7890123
- 123456: 생년, 월, 일
- 7890123: 성별/ 지역/ 고유번호

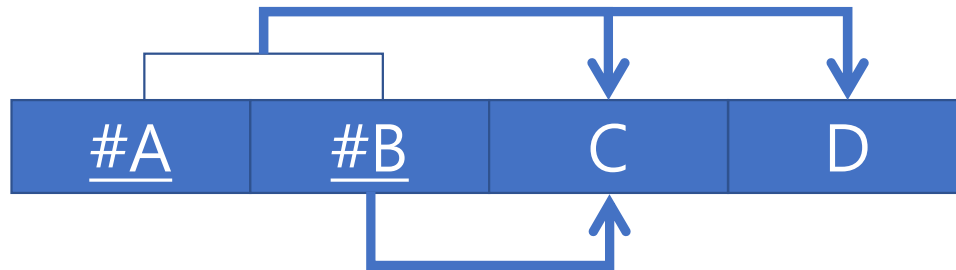
제2 정규형 – 부분적 함수 종속

- 두 개 이상으로 구성된 PK에서 발생
- R의 모든 속성이 후보 식별자 전체에 종속
- 일반 속성이 후보 식별자 전체에 종속되지 않고 일부에 종속된다면 2정규형 아님
- 따라서 2정규형이기 위해서는 모든 비 식별자 속성은 후보 식별자 속성에 완전 함수 종속돼야 한다

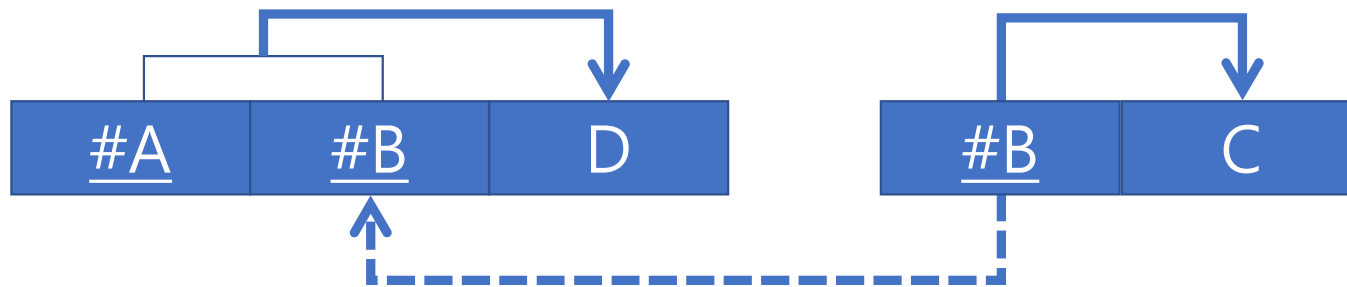
제2 정규형 – 부분적 함수 종속

- 제 2 정규형 방법
 - 일반 속성 중에서
 - 후보 식별자 전체에 종속적이지 않은 속성을 찾아
 - 기본 엔터티에서 제거하고,
 - 그 속성의 결정자를 주 식별자로 하는 새로운 상위 엔터티를 생성

제2 정규형 - 부분적 함수 종속



[부분 함수 종속이 존재하는 릴레이션]



[정규화한 릴레이션]

제2 정규형 – 부분적 함수 종속

- 예제

주문_상품	
PK	주문번호
PK	상품번호
	상품명
	단가
	주문수량

- 함수적 종속관계를 분석하고 2정규화 수행

제2 정규형 – 부분적 함수 종속

- 예제

주문_상품	
PK	주문번호
PK	상품번호
	상품명
	단가
	주문수량

fd:(상품번호)-> 단가

fd1:(주문번호,상품번호)->주문수량

fd2:상품번호->상품명

제2 정규형 – 부분적 함수 종속

- 예제
- FD1: (주문번호, 상품번호) -> 주문수량
- FD2: 상품번호 -> (상품명, 단가)

상품	
PK	상품번호
	상품명
	단가

주문_상품	
PK	주문번호
PK/FK	상품번호
	주문수량

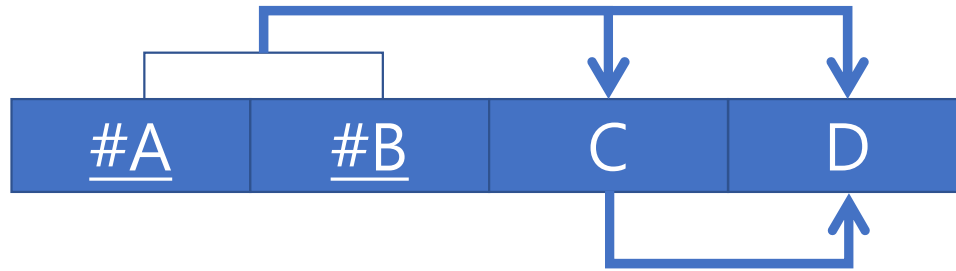
제3 정규형 – 이행적 함수 종속

- 식별자가 아닌 일반 속성 간에는 종속성이 존재하지 않는다

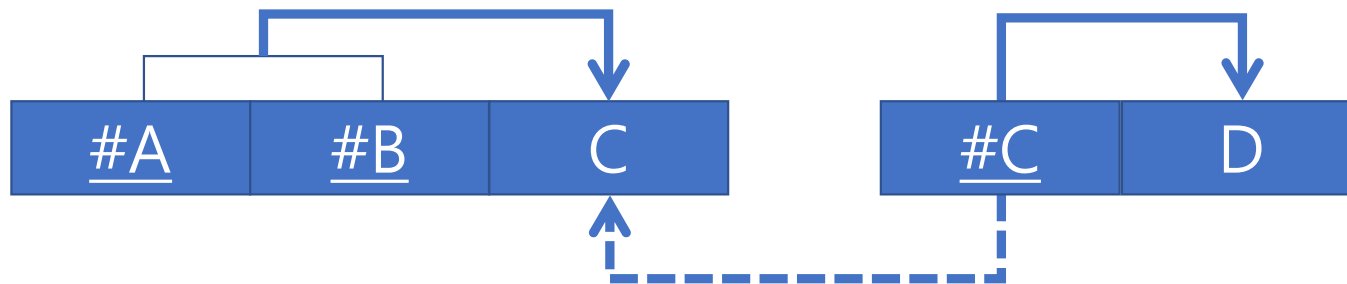
제3 정규화 방법

- 일반 속성간의 종속 관계를 분해하는 것

제3 정규형 - 이행적 함수 종속



[3정규화 대상 릴레이션]



[정규화한 릴레이션]

제3 정규형 - 이행적 함수 종속

- 예제

주문	
PK	주문번호
	고객번호
	고객명
	주문일자
	배송요청일자



고객	
PK	고객번호
	고객명



주문	
PK	주문번호
FK	고객번호
	주문일자
	배송요청일자