



정규화(Normalization)

컴퓨터정보공학과
교수 허 태 성



정규화 (normalization)

- DB에서 표현될 데이터들에 대하여 적절한
논리적 자료구조를 어떻게 구성할 것인가?
→논리적 DB설계시 릴레이션과 어트리뷰트의
구성 결정

→데이터의 중복성을 최소화함.

정규화 (normalization)

전체 릴레이션의 집합 (정규화, 혹은 비정규화된 모든 릴레이션)

제 1 정규형 (정규화된 모든 릴레이션)

제 2 정규형 릴레이션

제 3 정규형 릴레이션

BCNF 릴레이션

제 4 정규형 릴레이션

제 5 정규형 릴레이션
(PJ/NF)



정규화 (normalization)

학생1

*학 번	이 름	학 년	이수학점	*과목번호	성 적
2007001	홍길동	2	20	101	95
2007001	홍길동	2	20	102	80
2007001	홍길동	2	20	103	60
2007002	이일삼	4	16	101	80
2007002	이일삼	4	16	104	95
2006012	김이사	2	20	102	70
2006012	김이사	2	20	103	80
2007033	이육사	3	18	102	85
2004014	김일구	4	16	104	90
2004014	김일구	4	16	105	75



정규화 (normalization)

학사 데이터베이스

학생

학 번	이 름	학 년	학 과
2007001	홍길동	2	컴퓨터과
2006002	이일삼	4	기계과
2006012	김이사	2	전기과
2007033	이육사	3	전자과
2004014	김일구	4	컴퓨터과
2005032	정소일	4	컴퓨터과

수강

학번	과목번호	성적
2007001	101	95
2007001	102	80
2007001	103	60
2006002	101	80
2006002	104	95
2006012	102	70
2006012	103	80
2007033	102	85
2004014	104	90
2004014	105	75

과목

과목번호	교수명	과목명
101	이진필	데이터베이스
102	김기용	자료구조
103	김운석	컴퓨터시스템
104	박이중	멀티미디어
105	김철제	프로그래밍

정규화 (normalization)

●학생1 테이블

- 학과 열에 중복 데이터 발생

●문제점

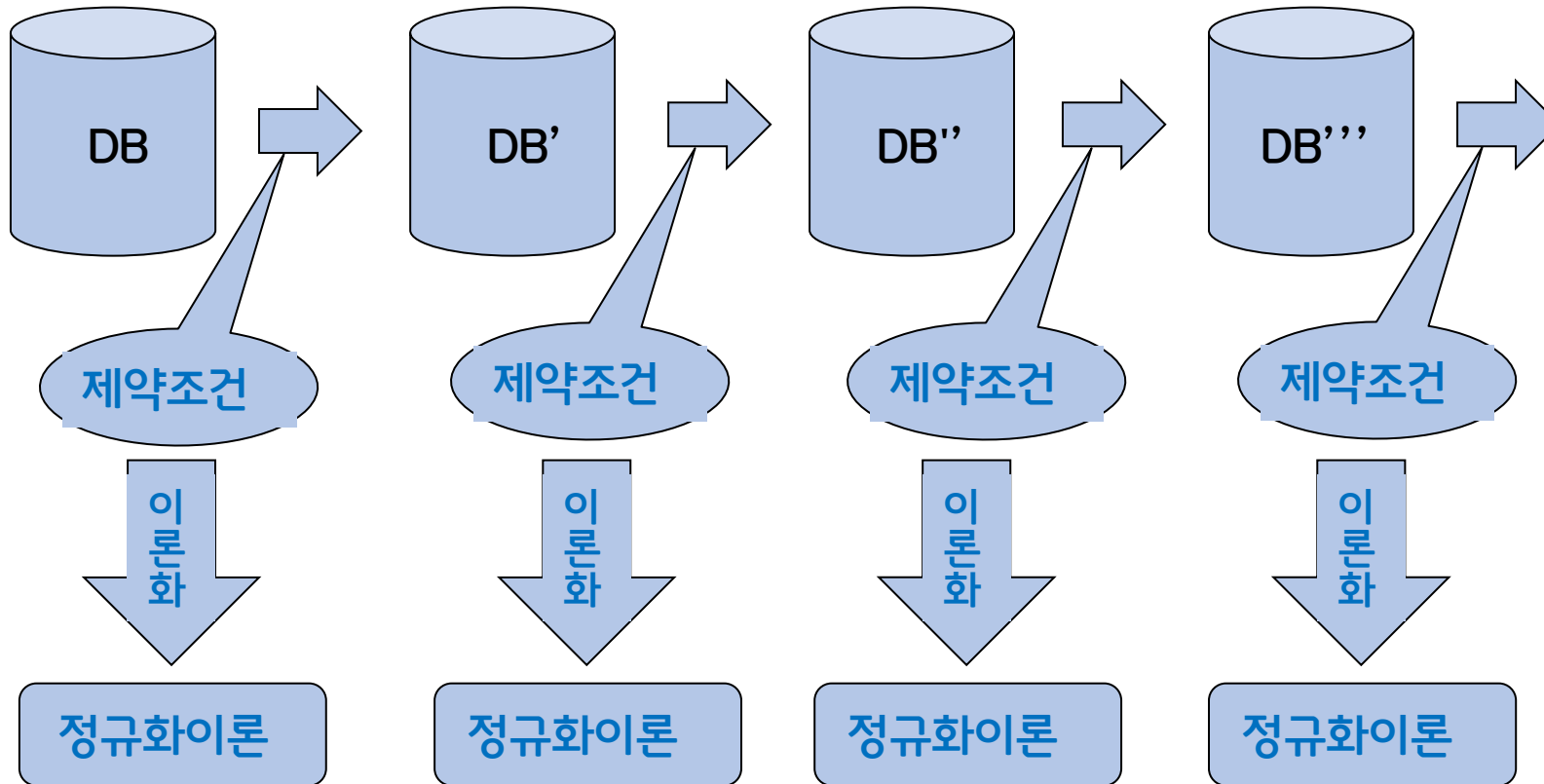
- 자료 변경시 **이상현상(anomaly)** 발생의 원인

●해결책

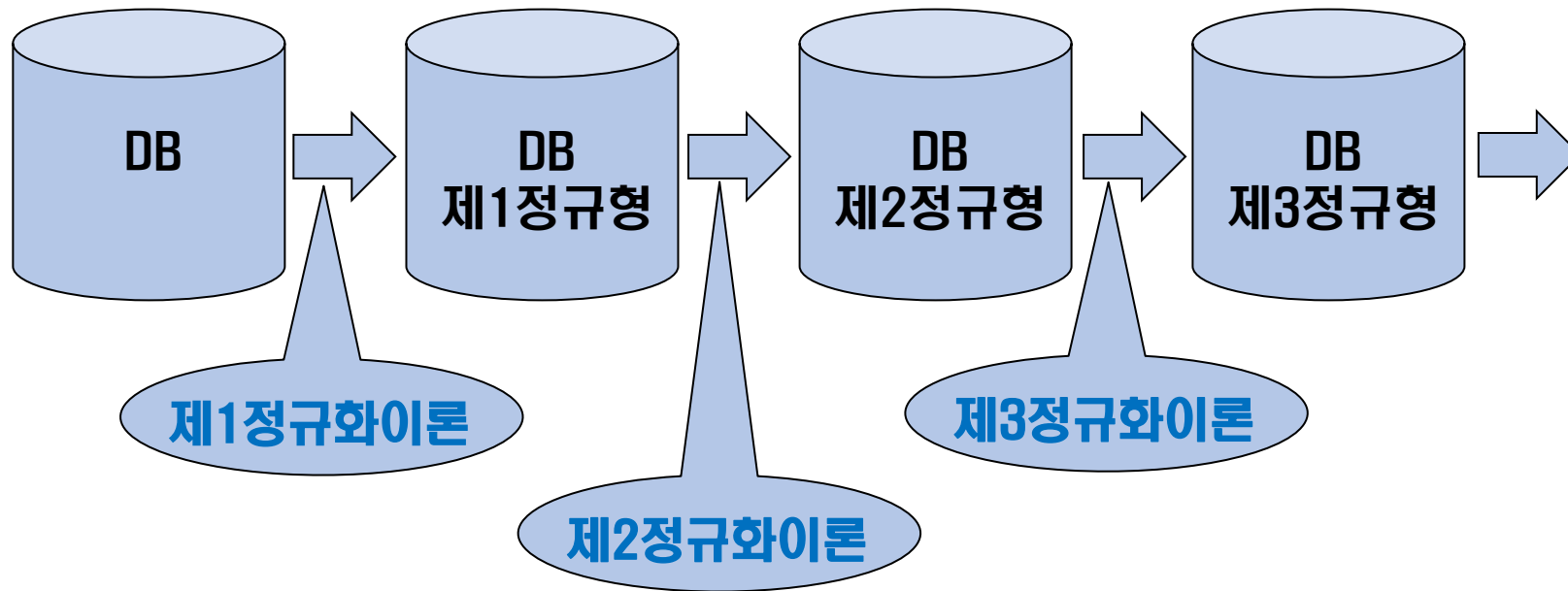
- 중복을 최소화하기 위한 제약조건을 이론화하여 적용함

→ 릴레이션을 분해함.

정규화 (normalization)



정규화 (normalization)



정규화 (normalization)

◆ 제 1 정규형

*학 번	이 름	학 년	이수학점	*번 호	성 적
2007001	홍길동	2	20	101	95
2007001	홍길동	2	20	102	80
2007001	홍길동	2	20	103	60
2007002	이일삼	4	16	101	80
2007002	이일삼	4	16	104	95
2006012	김이사	2	20	102	70
2006012	김이사	2	20	103	80
2007033	이육사	3	18	102	85
2004014	김일구	4	16	104	90
2004014	김일구	4	16	105	75

➤ 이수학점 : 1,2학년 20학점, 3학년 18학점, 4학년 16학점



정규화 (normalization)

◆ 제 1 정규형(계속)

➤ 정의

- ✓ 주어진 릴레이션 R의 모든 도메인들은 단지 **단일값 (atomic value)**만을 가질때, R은 제1정규형에 속한다.

➤ 이상현상 발생

- ✓ 삽입의 이상현상
- ✓ 삭제의 이상 현상
- ✓ 갱신의 이상 현상

→ 제 2 정규형



정규화 (normalization)

◆ 제 2 정규형

➤ 함수적 종속성

학 번	이 름	학 년	이수학점
2002001	홍길동	2	20
2001002	이일삼	4	16
2001012	김이사	2	20
2002033	이육사	3	18
1999014	김일구	4	16
2000032	정소일	4	16

✓ 학번, 이름, 학년, 이수학점의 어트리뷰트의 값들간의 관계



정규화 (normalization)

◆ 제 2 정규형(계속)

➤ 함수적 종속성

- ✓ 학생2테이블의 어트리뷰트인 이름, 학년, 이수학점 은 각각 학생 테이블의 어트리뷰트 학번에 **함수적 종속(Functional Dependency : FD)** 된다.
- ✓ 학번은 이름, 학년, 이수학점을 **함수적으로 결정(Functional Determine)** 한다.

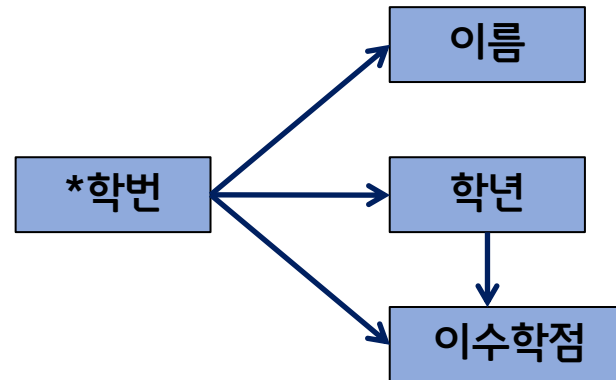
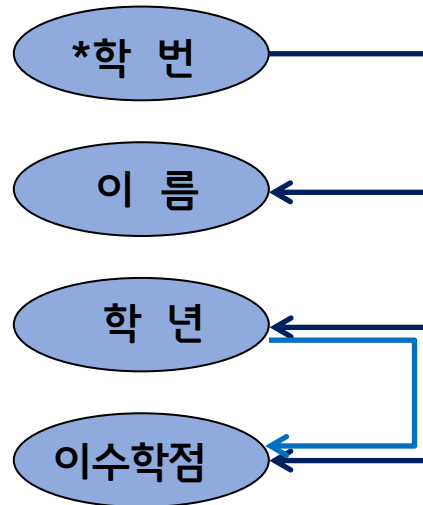
학생2.학번 → 학생2.이름
학생2.학번 → 학생2.학년
학생2.학번 → 학생2.이수학점



학생2.학번 → 학생2. (이름, 학년, 이수학점)

정규화 (normalization)

◆ 제 2 정규형(계속)



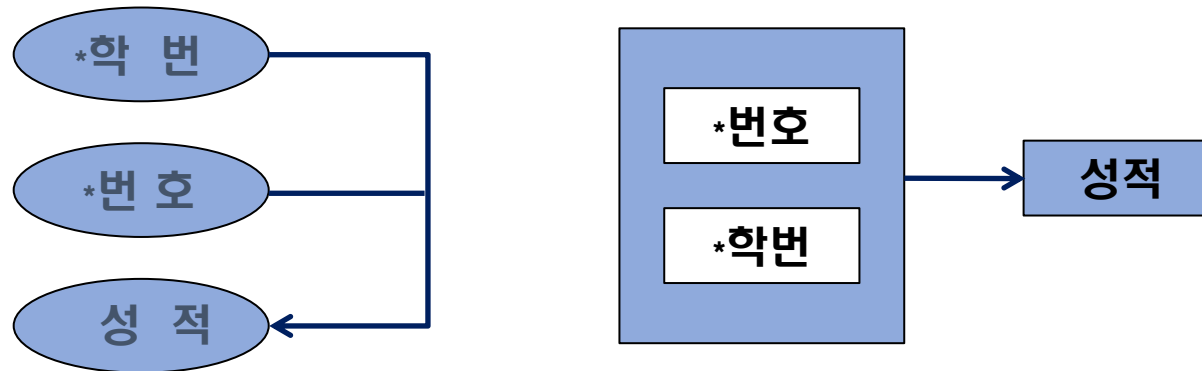
정규화 (normalization)

◆ 제 2 정규형(계속)

➤ 함수적 종속성 정의

주어진 릴레이션 R에서의 임의의 어트리뷰트 X, Y에 대하여, X의 데이터(값)들에서 Y의 데이터들로의 대응관계가 함수 성질을 만족하면 Y는 X에 함수적 종속성(FD)을 갖는다.

Ex) 수강(*학번, *번호, 성적)



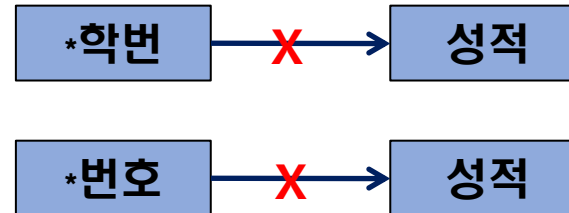
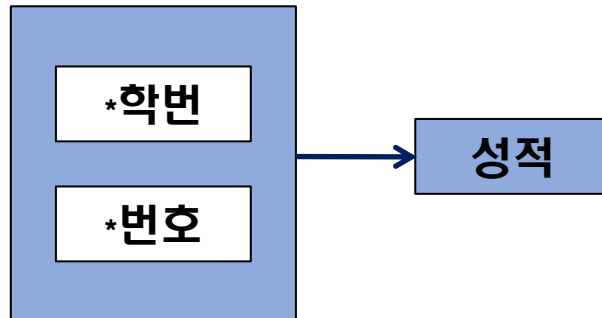
정규화 (normalization)

◆ 제 2 정규형(계속)

➤ 완전 함수적 종속성 정의

릴레이션 R에서의 임의의 어트리뷰트 X, Y에 대하여, Y가 X에 함수적 종속이고, X의 진부분 집합 어트리뷰트에 대해서는 함수적 종속성이 없을 경우 Y는 X에 완전 함수적 종속성(FFD)을 갖는다.

Ex) 수강(*학번, *번호, 성적)



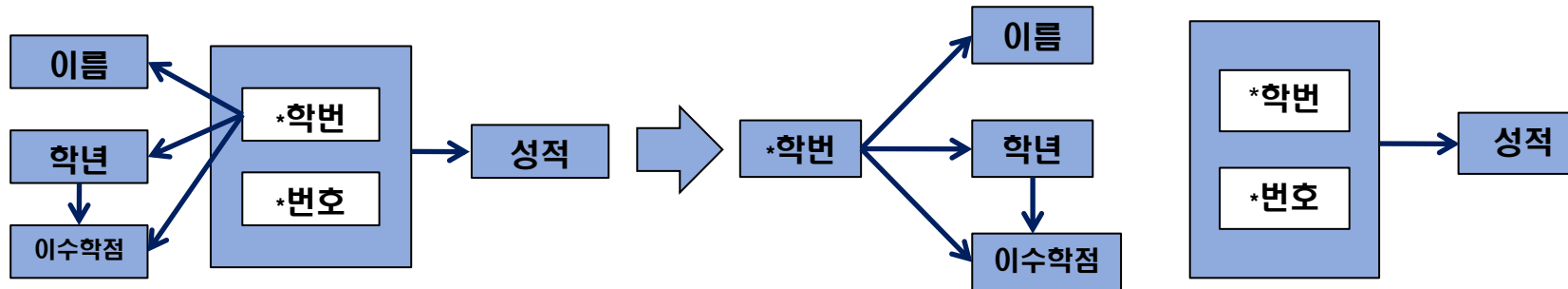
정규화 (normalization)

◆ 제 2 정규형(계속)

➤ 정의

릴레이션 R이 1NF이고, 키가 아닌 모든 어트리뷰트가 기본 키 어트리뷰트에 **완전 함수적 종속성[FFD]**을 만족하면 릴레이션 R은 제2정규형(2NF)이다.

➤Ex) 학생1(*학번, 이름, 학년, 이수학점, *번호, 성적)



➤Ex) 학생2(*학번, 이름, 학년, 이수학점)
수강(*학번, *번호, 성적)

정규화 (normalization)

◆ 제 3 정규형

➤ 이상현상 발생

- ✓삽입의 이상현상
- ✓삭제의 이상 현상
- ✓갱신의 이상 현상

→ 제 3 정규형

➤ 이행적 종속성 [P238]

어떤 릴레이션 R의 어트리뷰트 A,B,C에 대해

$R.A \rightarrow R.B$ 이고,

$R.B \rightarrow R.C$ 이면

$R.A \rightarrow R.C$ 이다.

이와같은 함수적 종속성을 가질 때 릴레이션 이 때 어트리뷰트 C는 A에
이행적 종속성(Transitive Dependency)을 갖는다.

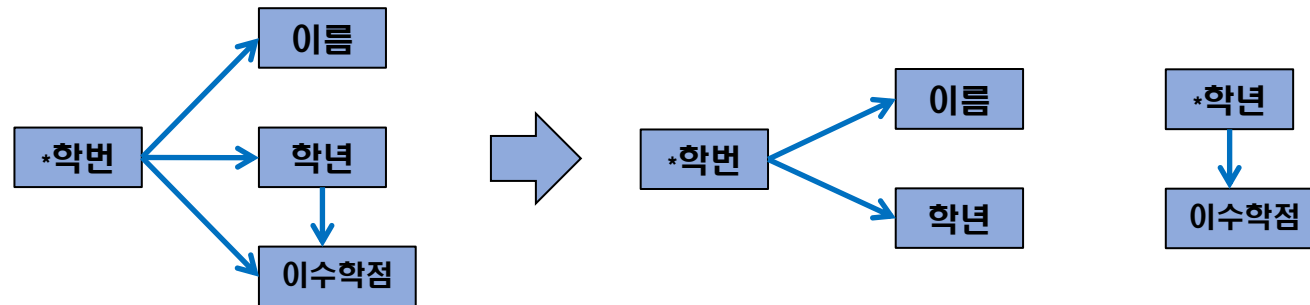
정규화 (normalization)

◆ 제 3 정규형(계속)

➤ 정의

릴레이션 R이 2NF이고, 키가 아닌 모든 어트리뷰트가 기본 키 어트리뷰트에 대하여 **이행적 함수 종속이 아닐 때**, 릴레이션 R은 제3정규형 (3NF)이다.

➤ Ex) 학생2(*학번, 이름, 학년, 이수학점)



➤ Ex) 학년(*학번, 이름, 학년)

이수학점(*학년, 이수학점)

정규화 (normalization)

◆ 제 1 정규형

➤ 정의

- ✓ 주어진 릴레이션 R의 모든 도메인들은 단지 **단일값 (atomic value)**만을 가질때, R은 제1정규형에 속한다.

◆ 제 2 정규형

➤ 정의

- ✓ 릴레이션 R이 1NF이고, 키가 아닌 모든 어트리뷰트가 기본 키 어트리뷰트에 **완전 함수적 종속성[FFD]**을 만족하면 릴레이션 R은 제2정규형(2NF)이다.

정규화 (normalization)

◆ 제 3 정규형

➤ 정의

- ✓ 릴레이션 R이 2NF이고, 키가 아닌 모든 어트리뷰트가 기본 키 어트리뷰트에 **이행적 함수 종속이 아닐 때**, 릴레이션 R은 제3정규형 (3NF)이다.

◆ 보이스/코드 정규형(BCNF)

➤ 정의

- ✓ 어떤 릴레이션 R에서 **모든 결정자가 후보 키**이면 이 릴레이션 R은 보이스/코드 정규형(BCNF)이다.

정규화 (normalization)

◆ 제 4 정규형

➤ 정의

- ✓ 어떤 릴레이션 R에서 $R.A \rightarrow R.B$ 인 **다중치 종속성(MVD)**가 존재할 경우 R의 모든 어트리뷰트들이 A에 함수적으로 종속되면 릴레이션 R은 제4정규형(4NF)이다.

◆ 제 5 정규형(5NF)

➤ 정의

- ✓ 어떤 릴레이션 R에서 의 모든 **조인 종속성**이 후보 키를 귀결(consequence) 지을 수 있다면 R은 제5정규형(5NF)이다.
→ **PJ/NF(Projection-Join Normal Form)**

정규화 (normalization)

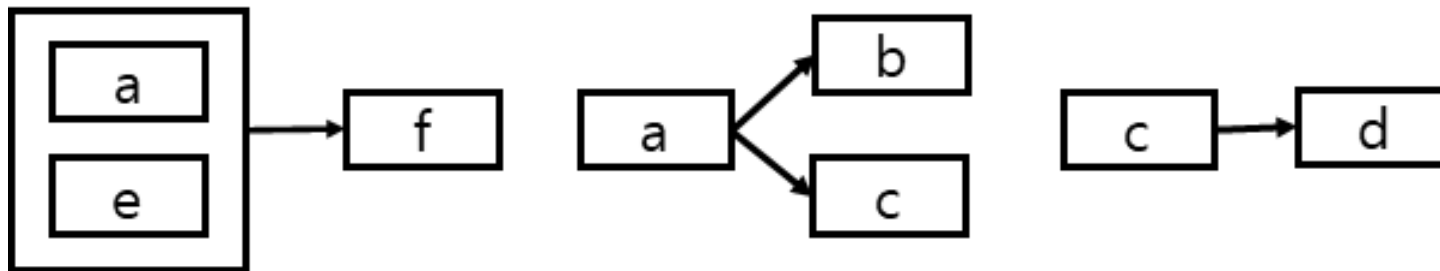
- ◆ 테이블 R에 속성(attribute)들 간의 FD(함수적종속성) 관계가 다음과 같을 때, 제3정규형의 FD 다이어그램으로 작성하시오.

$R(\underline{a}, b, c, d, \underline{e}, f)$ (단, $\underline{\quad}$ 은 기본키 임)

$R(a) \rightarrow R(b, c, d)$

$R(a, e) \rightarrow R(f)$

$R(c) \rightarrow R(d)$



질의 & 응답

