

C H A P T E R 0 9

정규화

Normalization

이상 현상을 제거하는 5단계 — 함수 종속부터 BCNF까지

박상돈 조교수 대전대학교 · 컴퓨터공학과 · AxGS Lab

목차

01 정규화와 이상 현상

왜 잘못된 설계가 문제인가 — 삽입·갱신·삭제 이상

02 함수 종속

속성 간 관련성을 판단하는 기준 — 결정자와 종속자

03 기본 정규형과 정규화 과정

1NF → 2NF → 3NF → BCNF 단계별 분해

S E C T I O N

01

정규화와 이상 현상

Normalization & Anomalies

잘못 설계된 릴레이션은 어떤 문제를 일으키는가

정규화란 무엇인가?

DEFINITION

이상(anomaly) 현상을 제거하면서 데이터베이스를 올바르게 설계해 나가는 과정

불필요한 데이터 중복이 만들어내는 세 가지 이상 현상



삽입 이상

Insertion anomaly

원하지 않는 값까지 함께 넣어야만 삽입이 되는 문제



갱신 이상

Update anomaly

중복된 값 중 일부만 바뀌어 데이터가 모순 되는 문제



삭제 이상

Deletion anomaly

하나를 지우면 꼭 필요한 다른 정보까지 사라지는 문제

문제의 출발 — 이벤트참여 릴레이션

관련 없는 속성들을 하나의 릴레이션에 모아두면, 데이터 중복과 이상 현상이 발생합니다.

고객아이디	이벤트번호	당첨여부	고객이름	등급
apple	E001	Y	정소화	gold
apple	E005	N	정소화	gold
apple	E010	Y	정소화	gold
banana	E002	N	김선우	vip
banana	E005	Y	김선우	vip
carrot	E003	Y	고명석	gold
carrot	E007	Y	고명석	gold
orange	E004	N	김용욱	silver
melon	NULL	NULL	성원용	gold

← 삽입 불가!

그림 9-3 이벤트참여 릴레이션의 삽입 이상



삽입 이상 기본키인 이벤트번호가 Null이 될 수 없어, 이벤트에 참여하지 않은 고객(melon)은 추가할 수 없습니다.

갱신 이상 — 일부만 바뀌면 모순

고객아이디	이벤트번호	당첨여부	고객이름	등급
apple	E001	Y	정소화	vip
apple	E005	N	정소화	vip
apple	E010	Y	정소화	gold
banana	E002	N	김선우	vip
banana	E005	Y	김선우	vip
carrot	E003	Y	고명석	gold
carrot	E007	Y	고명석	gold
orange	E004	N	김용욱	silver

← 데이터 불일치 발생!

그림 9-4 이벤트참여 릴레이션의 갱신 이상



갱신 이상 정소화 고객의 등급이 gold→vip로 바뀔 때, 중복 저장된 모든 튜플을 동시에 바꾸지 않으면 데이터가 서로 모순됩니다.

삭제 이상 — 같이 사라지는 정보

고객아이디	이벤트번호	당첨여부	고객이름	등급
apple	E001	Y	정소화	gold
apple	E005	N	정소화	gold
apple	E010	Y	정소화	gold
banana	E002	N	김선우	vip
banana	E005	Y	김선우	vip
carrot	E003	Y	고명석	gold
carrot	E007	Y	고명석	gold
orange	E004	N	김용욱	silver

← 데이터 손실 발생!

그림 9-5 이벤트참여 릴레이션의 삭제 이상



삭제 이상 orange 고객의 이벤트 기록만 지우려 해도, 그 고객의 유일한 정보였던 고객이름·등급까지 함께 삭제됩니다.

정규화의 필요성 — 해결의 방향

문제의 원인

관련 없는 속성들을 하나의 릴레이션에 함께 모아둔 것 → 데이터 중복 발생

해결 — 분해(Decomposition)

관련 있는 속성들로만 릴레이션을 구성하도록 쪼개는 것



판단 기준

속성들 간의 관련성을 함수적 종속성(Functional Dependency)으로 판단한다



분해 원칙

일반적으로 릴레이션에 함수적 종속성이 하나만 존재하도록 분해한다

S E C T I O N

02 함수 종속

Functional Dependency

결정자와 종속자 — 속성 간 관련성을 읽는 법

함수 종속이란?

DEFINITION

X 값이 Y 값을 유일하게 결정하면 "Y는 X에 함수적으로 종속된다" ($X \rightarrow Y$)

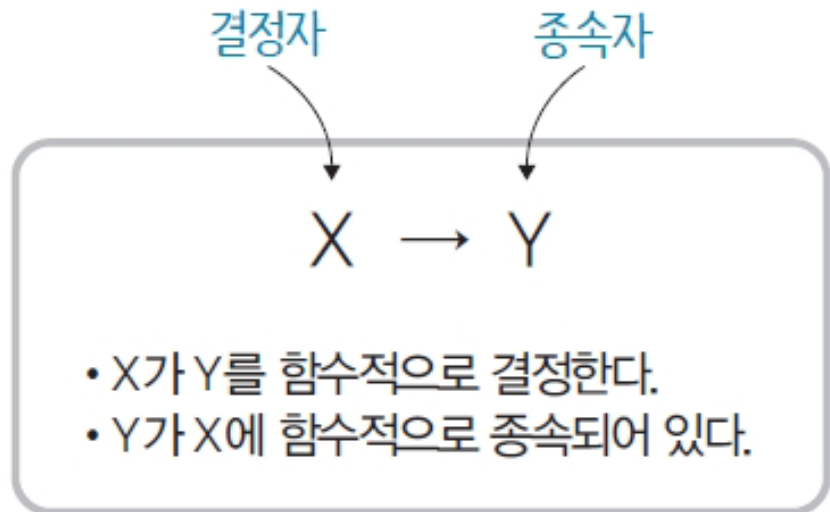


그림 9-6 함수 종속의 표현

결정자 (Determinant) X

종속자 (Dependent) Y



핵심 화살표 $X \rightarrow Y$ 는 'X를 알면 Y가 하나로 정해진다'는 뜻입니다. 예) 고객아이디 $\rightarrow$ 고객이름, 등급

함수 종속의 표현 — 고객 릴레이션

고객아이디 하나가 정해지면 고객이름·등급이 유일하게 결정됩니다.

고객아이디	고객이름	등급
apple	정소화	gold
banana	김선우	vip
carrot	고명석	gold
orange	김용욱	silver

그림 9-7 함수 종속 관계 설명을 위한 릴레이션의 예: 고객 릴레이션

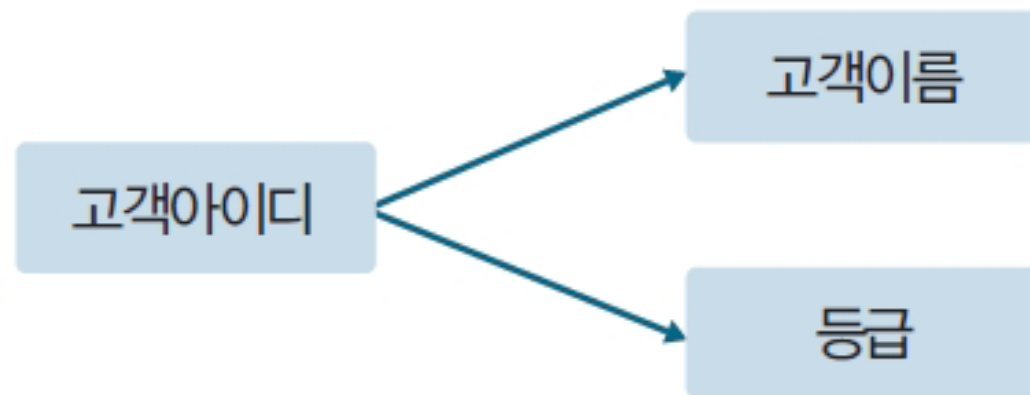


그림 9-9 고객 릴레이션의 함수 종속 다이어그램

고객아이디 → 고객이름
고객아이디 → 등급

또는

고객아이디 → (고객이름, 등급)

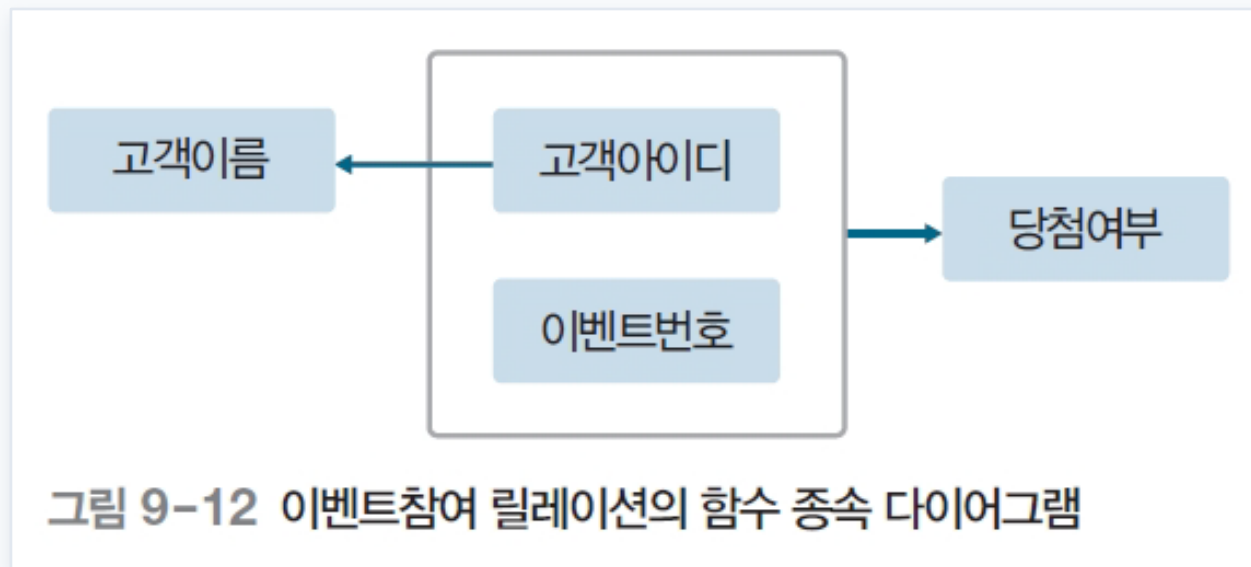
그림 9-8 고객 릴레이션에 존재하는 함수 종속 관계

완전 함수 종속 vs 부분 함수 종속

이벤트참여 릴레이션의 기본키 {고객아이디, 이벤트번호} 를 기준으로 살펴봅니다.

고객아이디	이벤트번호	당첨여부	고객이름
apple	E001	Y	정소화
apple	E005	N	정소화
apple	E010	Y	정소화
banana	E002	N	김선우
banana	E005	Y	김선우
carrot	E003	Y	고명석
carrot	E007	Y	고명석
orange	E004	N	김용욱

그림 9-10 함수 종속 관계를 설명하기 위한 릴레이션의 예 : 이벤트참여 릴레이션



완전 함수 종속 (FFD)
당첨여부 — 기본키 전체에 종속

부분 함수 종속 (PFD)
고객이름 — 기본키 일부(고객아이디)에만 종속

함수 종속 관계의 예외

다음 두 경우는 당연히 성립하므로, 함수 종속으로 따로 고려하지 않습니다.



결정자와 종속자가 같은 경우

$X \rightarrow X$ 처럼 자기 자신을 결정하는 것은 자명하다



결정자가 종속자를 포함하는 경우

$\{\text{고객아이디, 이벤트번호}\} \rightarrow \text{고객아이디}$ 도 자명하다

고객아이디 $\rightarrow$ 고객아이디

$\{\text{고객아이디, 이벤트번호}\} \rightarrow \text{이벤트번호}$

그림 9-13 고려할 필요가 없는 함수 종속 관계의 예

S E C T I O N

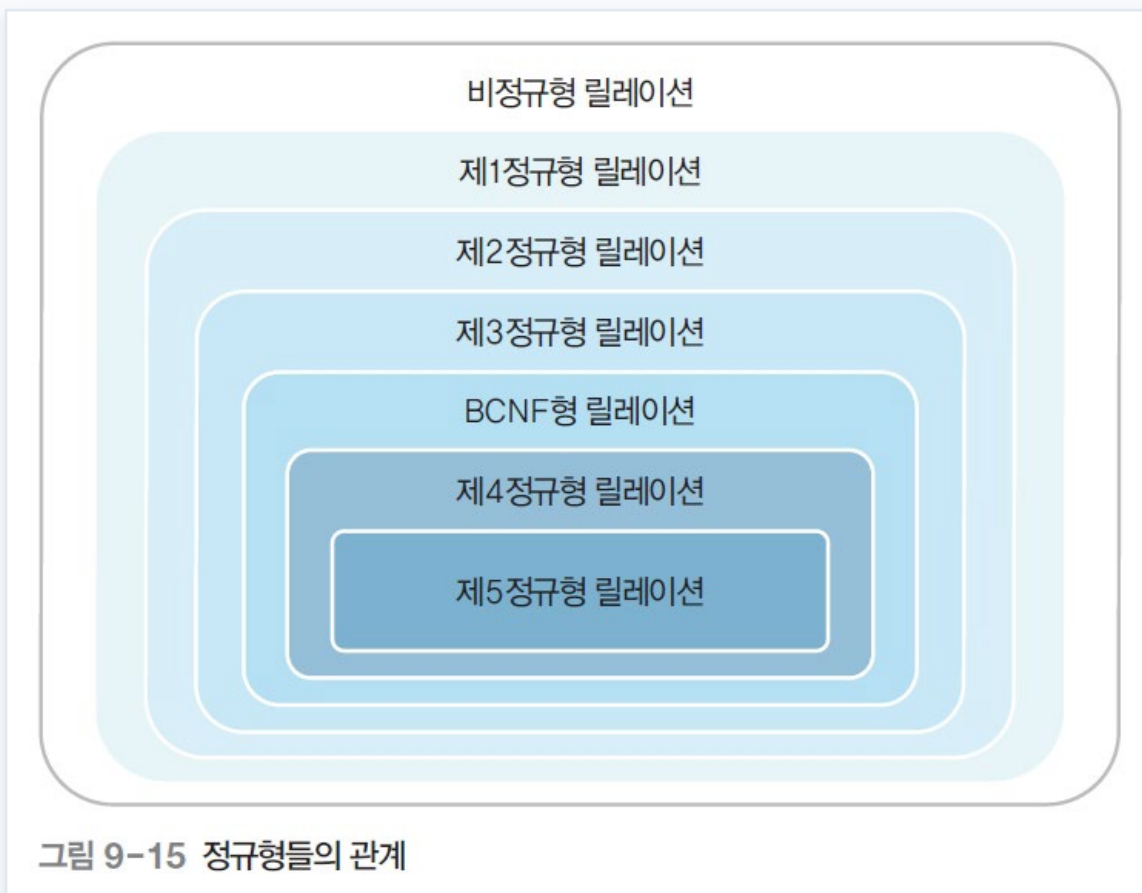
03 기본 정규형과 정규화 과정

Normal Forms

1NF → 2NF → 3NF → BCNF, 단계별로 분해하기

정규형의 종류와 포함 관계

정규형은 바깥쪽(약함)에서 안쪽(강함)으로 갈수록 더 엄격한 조건을 요구합니다.



제1정규형 (1NF)

제1정규형 1NF · First Normal Form

RULE 1NF

릴레이션의 모든 속성이 더 이상 분해되지 않는 원자 값(atomic value)만 가져야 한다

다중 값 → 원자 값으로

고객아이디	이벤트번호	당첨여부	등급	할인율
apple	E001, E005, E010	Y, N, Y	gold	10%
banana	E002, E005	N, Y	vip	20%
carrot	E003, E007	Y, Y	gold	10%
orange	E004	N	silver	5%



고객아이디	이벤트번호	당첨여부	등급	할인율
apple	E001	Y	gold	10%
apple	E005	N	gold	10%
apple	E010	Y	gold	10%
banana	E002	N	vip	20%
banana	E005	Y	vip	20%
carrot	E003	Y	gold	10%
carrot	E007	Y	gold	10%
orange	E004	N	silver	5%

그림 9-16 다중 값 속성을 포함하는 이벤트참여 릴레이션

그림 9-17 제1정규형에 속하는 이벤트참여 릴레이션

1NF에도 남아 있는 이상 현상

원자 값이 되었어도, 등급·할인율이 중복되어 세 가지 이상이 여전히 발생합니다.

삽입 이상

갱신 이상

삭제 이상

고객아이디	이벤트번호	당첨여부	등급	할인율
apple	E001	Y	gold	10%
apple	E005	N	gold	10%
apple	E010	Y	gold	10%
banana	E002	N	vip	20%
banana	E005	Y	vip	20%
carrot	E003	Y	gold	10%
carrot	E007	Y	gold	10%
orange	E004	N	silver	5%
grape	NULL	NULL	silver	5%

← 삽입 불가!

그림 9-20 이벤트참여 릴레이션의 삽입 이상

고객아이디	이벤트번호	당첨여부	등급	할인율
apple	E001	Y	vip	10%
apple	E005	N	vip	10%
apple	E010	Y	gold	10%
banana	E002	N	vip	20%
banana	E005	Y	vip	20%
carrot	E003	Y	gold	10%
carrot	E007	Y	gold	10%
orange	E004	N	silver	5%

← 데이터 불일치 발생!

그림 9-21 이벤트참여 릴레이션의 갱신 이상

고객아이디	이벤트번호	당첨여부	등급	할인율
apple	E001	Y	gold	10%
apple	E005	N	gold	10%
apple	E010	Y	gold	10%
banana	E002	N	vip	20%
banana	E005	Y	vip	20%
carrot	E003	Y	gold	10%
carrot	E007	Y	gold	10%
orange	E004	N	silver	5%

← 데이터 손실 발생!

그림 9-22 이벤트참여 릴레이션의 삭제 이상

1NF 이상 현상 한눈에 보기

삽입 · 갱신 · 삭제 이상을 하나의 표에서 함께 확인할 수 있습니다.

고객아이디	이벤트번호	당첨여부	등급	할인율
apple	E001	Y	vip	10%
apple	E005	N	vip	10%
apple	E010	Y	gold	10%
banana	E002	N	vip	20%
banana	E005	Y	vip	20%
carrot	E003	Y	gold	10%
carrot	E007	Y	gold	10%
orange	E004	N	silver	5%
grape	NULL	NULL	silver	5%

← 데이터 불일치로 인한 갱신 이상

← 데이터 손실로 인한 삭제 이상

← 삽입 불가로 인한 삽입 이상

그림 9-23 이벤트참여 릴레이션의 삽입 · 갱신 · 삭제 이상

● 삽입 이상

● 갱신 이상

● 삭제 이상

1NF 해결 — 부분 함수 종속 제거



해결 방법 기본키의 일부에만 종속되는 상황(부분 함수 종속)을 제거하면 됩니다.

고객아이디 → 등급

고객아이디 → 할인율

등급 → 할인율

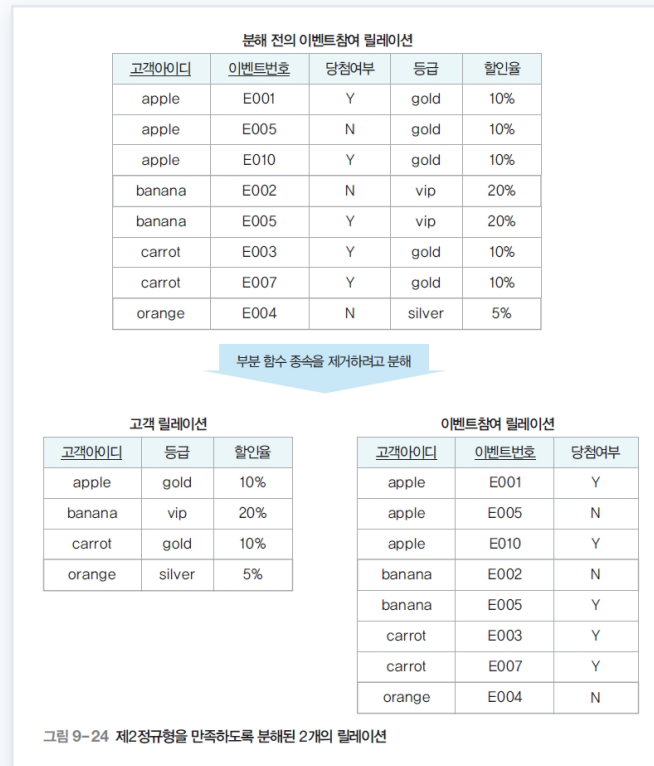
{고객아이디, 이벤트번호} → 당첨여부

그림 9-18 이벤트참여 릴레이션에 존재하는 함수 종속 관계

제2정규형 (2NF)

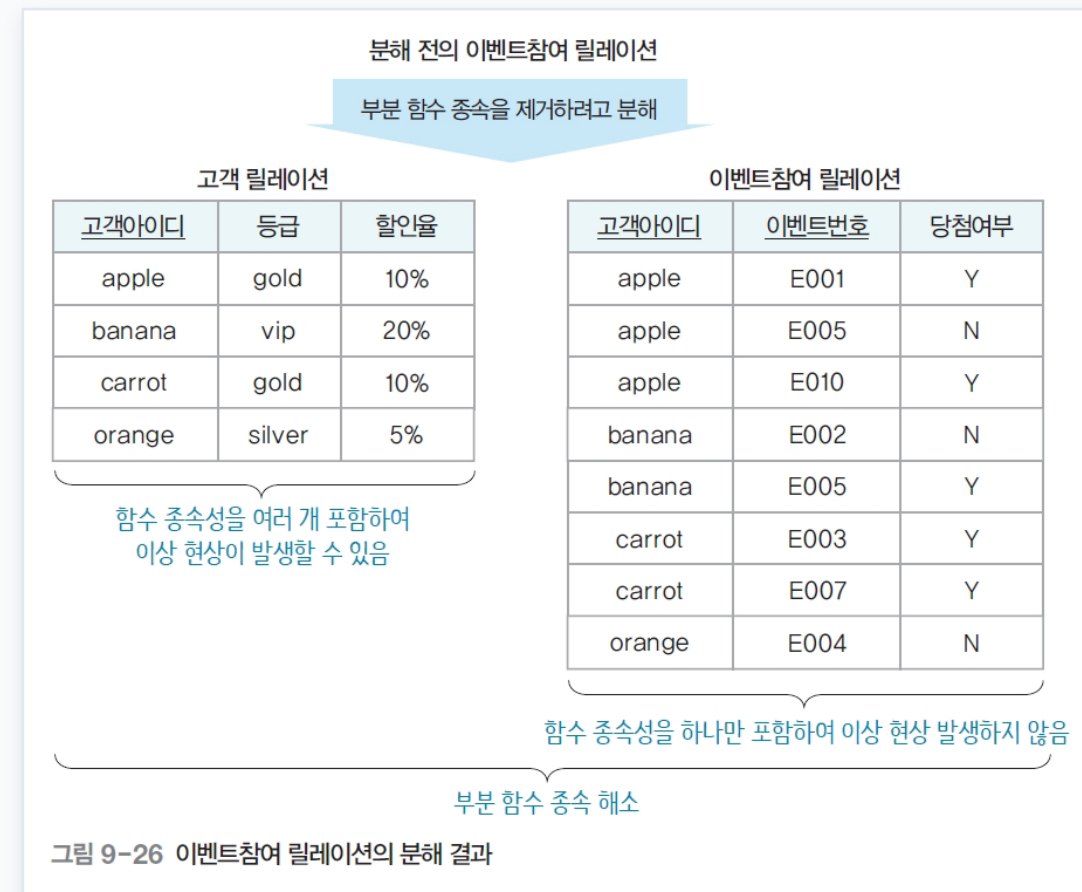
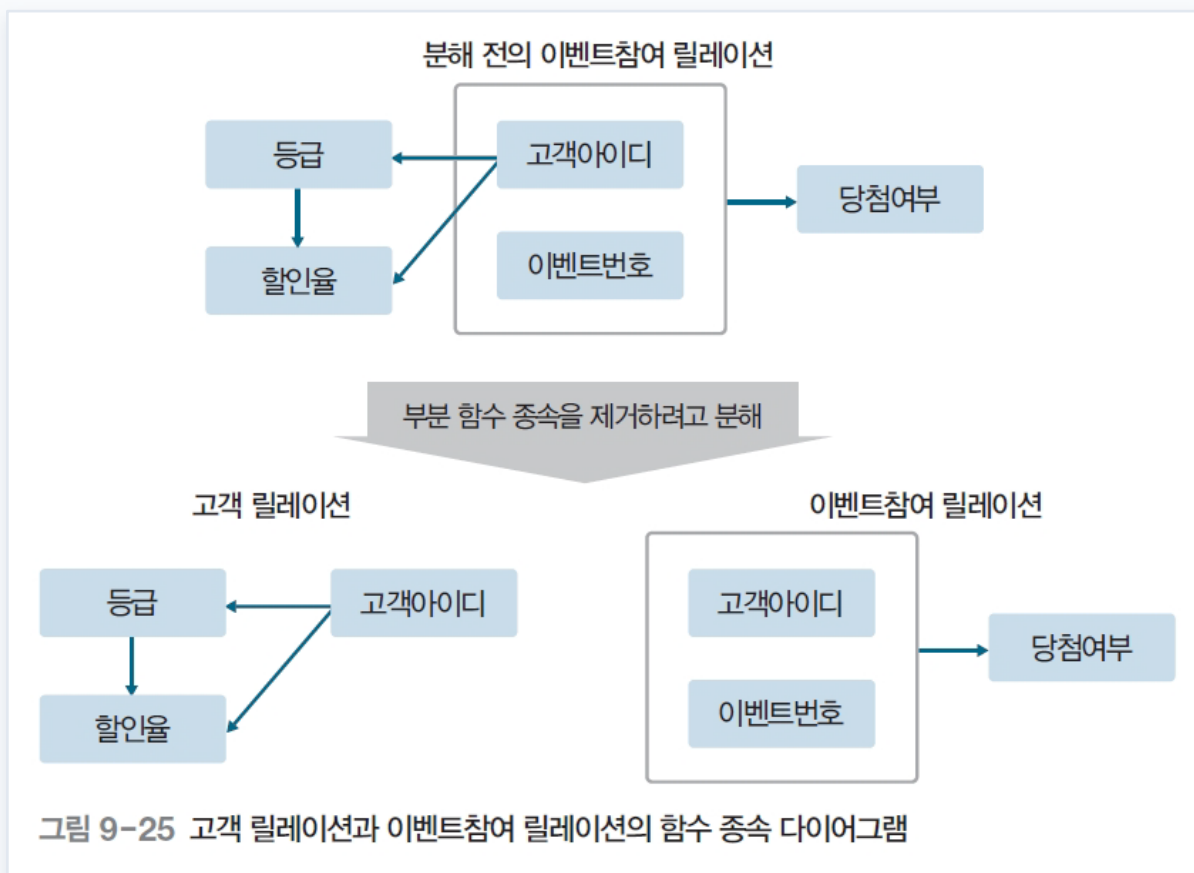
제2정규형 2NF · Second Normal Form

부분 함수 종속을 제거하고, 모든 속성이 기본키에 완전 함수 종속되도록 릴레이션을 분해합니다.



무손실 분해 (Nonloss Decomposition)

분해한 릴레이션을 자연 조인하면 원래의 이벤트참여 릴레이션을 손실 없이 복원할 수 있습니다.



2NF에도 남는 이상 — 고객 릴레이션

등급 → 할인율이라는 또 다른 종속 때문에, 고객 릴레이션에서 이상 현상이 다시 나타납니다.

삽입 이상

고객아이디	등급	할인율
apple	gold	10%
banana	vip	20%
carrot	gold	10%
orange	silver	5%
NULL	bronze	1%

← 삽입 불가!

그림 9-27 고객 릴레이션의 삽입 이상

고객이 없으면 등급·할인율 정책을 넣을 수 없음

갱신 이상

고객아이디	등급	할인율
apple	gold	15%
banana	vip	20%
carrot	gold	10%
orange	silver	5%

← 데이터 불일치 발생!

그림 9-28 고객 릴레이션의 갱신 이상

등급의 할인율 변경 시 모든 튜플 일괄 처리 필요

삭제 이상

고객아이디	등급	할인율
apple	gold	10%
banana	vip	20%
carrot	gold	10%
orange	silver	5%

← 데이터 손실 발생!

그림 9-29 고객 릴레이션의 삭제 이상

banana 탈퇴 시 vip·20% 정책 정보까지 삭제

2NF 해결 — 이행적 함수 종속 제거

TRANSITIVE FD

$X \rightarrow Y, Y \rightarrow Z$ 가 성립하면 논리적으로 $X \rightarrow Z$ 도 성립한다. 이때 Z 는 X 에 이행적으로 함수 종속(TFD)된다.

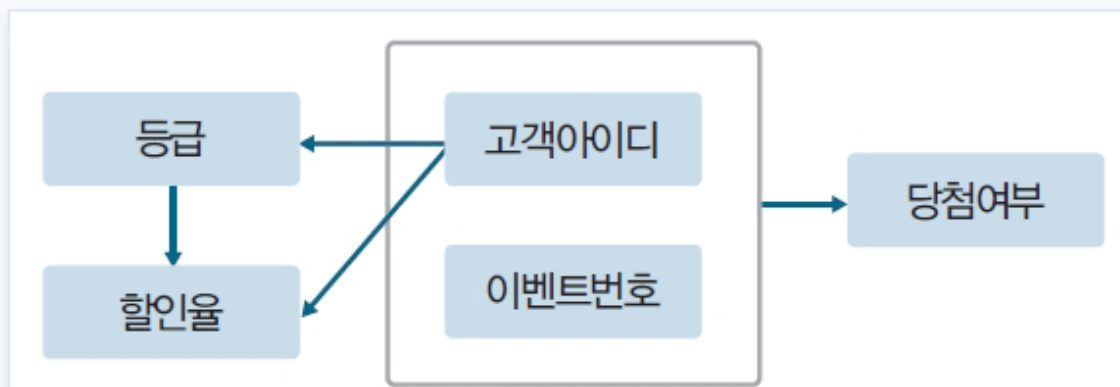


그림 9-19 이벤트참여 릴레이션의 함수 종속 다이어그램



적용 고객아이디 → 등급 → 할인율 이므로, 할인율은 고객아이디에 이행적으로 종속됩니다. 이 TFD를 제거해야 합니다.

제3정규형 (3NF)

제3정규형 3NF · Third Normal Form

RULE 3NF

이행적 함수 종속을 제거하여, 어떤 속성도 기본키에 이행적으로 종속되지 않도록 만드는 정규화 과정



분해 방법 $X \rightarrow Y \rightarrow Z$ 가 있으면 $X \rightarrow Y$ 와 $Y \rightarrow Z$ 로 분리한다. 즉 고객아이디 $\rightarrow$ 등급, 등급 $\rightarrow$ 할인을 로 분해합니다.

→ 고객 릴레이션과 고객등급 릴레이션으로 분리

3NF 분해 결과 — 고객 / 고객등급

분해 전의 고객 릴레이션

고객아이디	등급	할인율
apple	gold	10%
banana	vip	20%
carrot	gold	10%
orange	silver	5%

이행적 함수 종속을 제거하려고 분해

고객 릴레이션

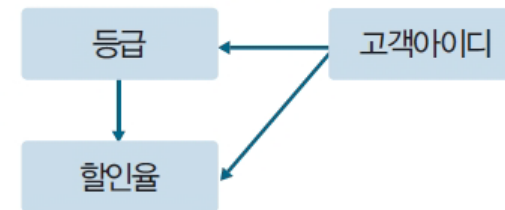
고객아이디	등급
apple	gold
banana	vip
carrot	gold
orange	silver

고객등급 릴레이션

등급	할인율
gold	10%
vip	20%
silver	5%

그림 9-32 제3정규형을 만족하도록 분해된 2개의 릴레이션

고객 릴레이션



이행적 함수 종속을 제거하려고 분해

고객 릴레이션



고객등급 릴레이션



그림 9-33 고객 릴레이션과 고객등급 릴레이션의 함수 종속 다이어그램

보이스/코드 정규형 (BCNF)

BCNF · Boyce/Codd Normal Form

RULE BCNF

모든 결정자가 후보키일 때 만족한다. '강한 제 3정규형(strong 3NF)' 이라고도 부른다.



후보키 강좌신청 릴레이션에는 후보키가 2개입니다: {고객아이디, 인터넷강좌}, {고객아이디, 담당강사번호}.

고객아이디	인터넷강좌	담당강사번호
apple	영어회화	P001
banana	기초토익	P002
carrot	영어회화	P001
carrot	기초토익	P004
orange	영어회화	P003
orange	기초토익	P004

그림 9-34 보이스/코드 정규형을 설명하기 위한 릴레이션의 예 : 강좌신청 릴레이션

BCNF — 결정자 분석

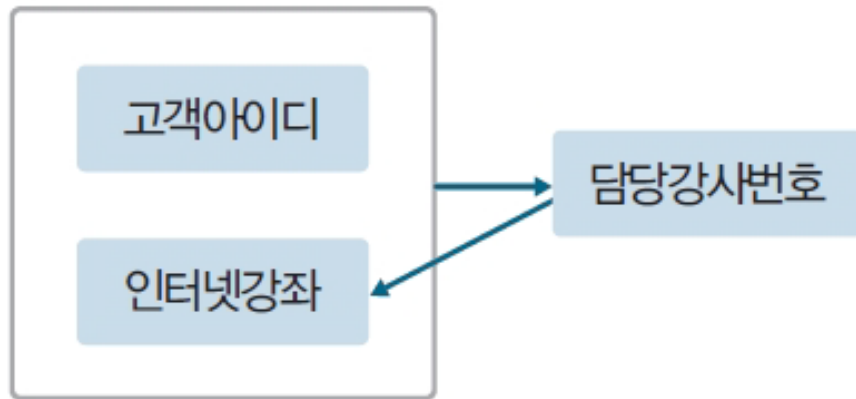


그림 9-35 강좌신청 릴레이션의 함수 종속 다이어그램

- 강사 한 명은 인터넷강좌를 하나만 담당한다
- 담당강사번호 → 인터넷강좌 종속이 존재한다
- 그러나 담당강사번호는 후보키가 아니다
- 이행적 함수 종속 케이스도 아니다

→ 후보키가 아닌 결정자가 있으므로 BCNF 위반

BCNF 위반이 부르는 이상 현상

삽입 이상

갱신 이상

삭제 이상

고객아이디	인터넷강좌	담당강사번호
apple	영어회화	P001
banana	기초토익	P002
carrot	영어회화	P001
carrot	기초토익	P004
orange	영어회화	P003
orange	기초토익	P004
NULL	중급토익	P005

← 삽입 불가!

그림 9-36 강좌신청 릴레이션의 삽입 이상

고객아이디	인터넷강좌	담당강사번호
apple	영어회화	P001
banana	기초토익	P002
carrot	영어회화	P001
carrot	중급토익	P004
orange	영어회화	P003
orange	기초토익	P004

← 데이터 불일치 발생!

그림 9-37 강좌신청 릴레이션의 갱신 이상

고객아이디	인터넷강좌	담당강사번호
apple	영어회화	P001
banana	기초토익	P002
carrot	영어회화	P001
carrot	기초토익	P004
orange	영어회화	P003
orange	기초토익	P004

← 데이터 손실 발생!

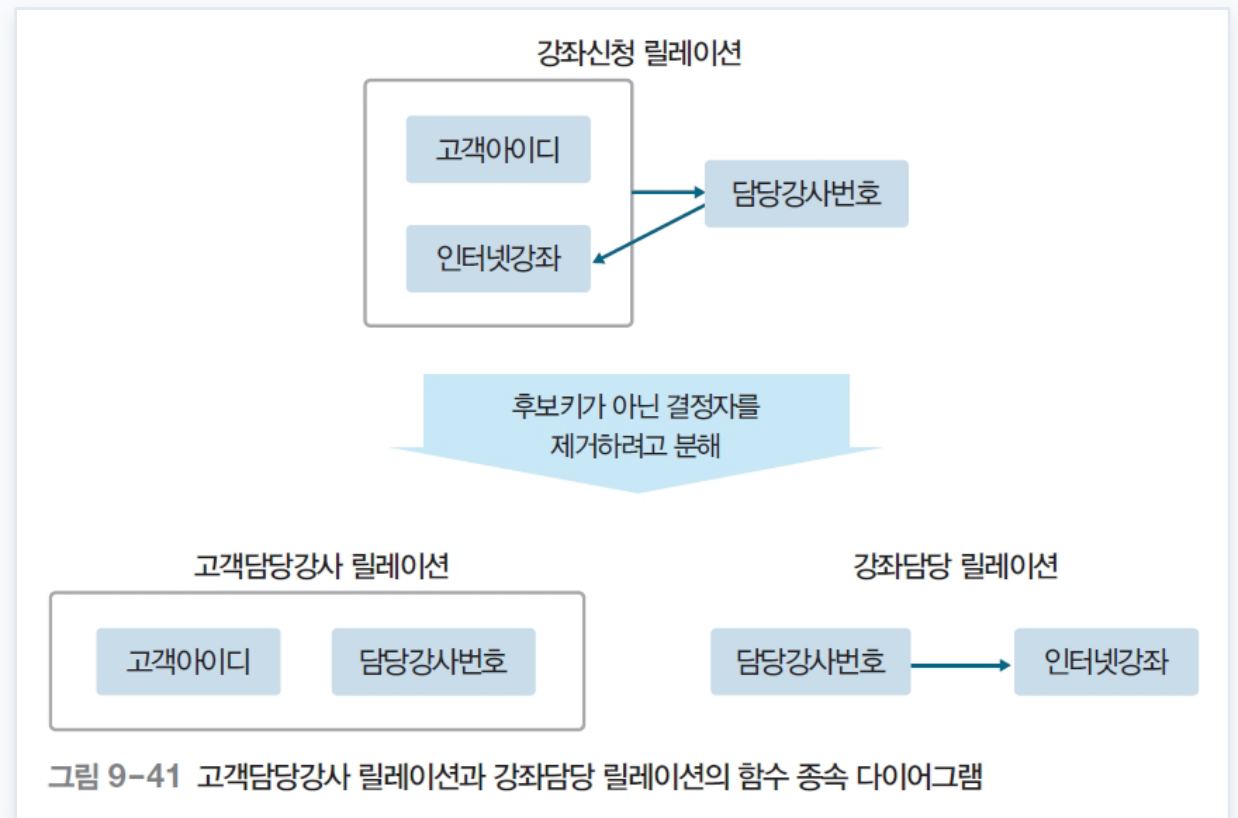
그림 9-38 강좌신청 릴레이션의 삭제 이상



갱신 이상 주의 강사 한 명이 강좌 하나만 담당한다는 전제와 모순되는 데이터가 생길 수 있습니다.

BCNF 해결 — 모든 결정자를 후보키로

모든 결정자가 후보키가 되도록, 고객담당강사 · 강좌담당 릴레이션으로 분리합니다.



기타 정규형 — 제4 · 제5정규형

 **실무 관점** 보통 제4·5정규형까지 진행하지 않고 3NF나 BCNF 수준에서 중복을 줄입니다. 과도한 분해는 오히려 비효율적일 수 있습니다.



제4정규형 (4NF)

BCNF를 만족하면서, 다치 종속(MVD: Multi-Valued Dependency)을 제거한 상태



제5정규형 (5NF)

BCNF를 만족하면서, 후보키를 통하지 않는 조인 종속(JD: Join Dependency)을 제거한 상태

정규화 과정 한눈에 보기 ★

비정규형 릴레이션 → 보이스/코드 정규형까지의 단계별 흐름



1NF	원자 값만 갖도록
2NF	부분 함수 종속 제거
3NF	이행적 함수 종속 제거
BCNF	모든 결정자를 후보키로

오늘 배운 내용 정리



이상 현상 3가지

삽입 · 갱신 · 삭제 이상 — 데이터 중복이 원인



함수 종속

$X \rightarrow Y$, 결정자가 종속자를 결정 · 완전/부분/이행적 종속



1NF · 2NF

원자 값 / 부분 함수 종속 제거 → 완전 함수 종속



3NF · BCNF

이행적 함수 종속 제거 / 모든 결정자를 후보키로



분해의 원칙

무손실 분해 — 자연 조인으로 원본 복원이 가능해야 함

T H A N K Y O U

수고하셨습니다

9장 정규화 · 끝

NEXT CLASS

Chapter 10 · 회복과 병행 제어

트랜잭션 · 장애 회복 · 동시 접근 제어 — 데이터를 안전하게 지키는 법