

오픈플랫폼활용

Chapter 02. 관계 데이터 모델

Section 01

관계 데이터 모델의 개념

릴레이션

■ 릴레이션(relation): 행과 열로 구성된 테이블

용어	한글 용어
relation	릴레이션, 테이블
relational data model	관계 데이터 모델
relational database	관계 데이터베이스
relational algebra	관계대수
relationship	관계

릴레이션

■ 데이터와 테이블(릴레이션)

도서 1, 축구의 역사, 굿스포츠, 7000
도서 2, 축구아는 여자, 나무수, 13000
도서 3, 축구의 이해, 대한미디어, 22000
도서 4, 골프 바이블, 대한미디어, 35000
도서 5, 피겨 교본, 굿스포츠, 8000



도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

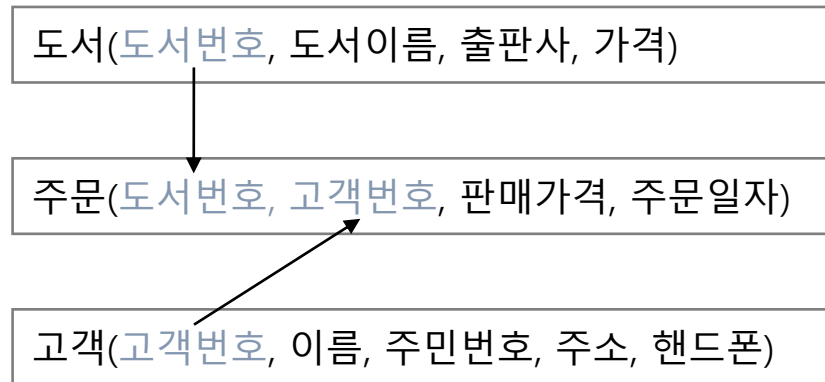
- 도서번호 = {1, 2, 3, 4, 5}
- 도서이름 = {축구의 역사, 축구아는 여자, 축구의 이해, 골프 바이블, 피겨 교본}
- 출판사 = {굿스포츠, 나무수, 대한미디어}
- 가격 = {7000, 13000, 22000, 35000, 8000}

→ 첫 번째 행(1, 축구의 역사, 굿스포츠, 7000)의 경우 네 개의 집합에서 각각 원소 한 개씩 선택하여 만들어진 것으로 이 원소들이 관계(relationship)를 맺고 있다.

릴레이션

■ 관계(relationship)

- 릴레이션 내에서 생성되는 관계: 릴레이션 내 데이터들의 관계
- 릴레이션 간에 생성되는 관계: 릴레이션 간의 관계



릴레이션 스키마와 인스턴스

- 관계(relationship)
 - 도서 릴레이션

속성(애트리뷰트),
열(column) 이라고도 함
(차수=4)

도서

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구하는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

스키마(내포)
Schema

인스턴스(외연)
Instance

튜플(tuple),
행(row) 이라고도 함
(카디널리티=5)

릴레이션 스키마와 인스턴스

■ 릴레이션 스키마

- 스키마의 요소

- 속성(attribute): 릴레이션 스키마의 열
- 도메인(domain): 속성이 가질 수 있는 값의 집합
- 차수(degree): 속성의 개수

- 스키마의 표현

- 릴레이션 이름(속성1: 도메인1, 속성2: 도메인2, 속성3: 도메인3 ...)
- ex) 도서(도서번호, 도서이름, 출판사, 가격)

■ 인스턴스 요소

- 튜플(tuple): 릴레이션의 행
- 카디널리티(cardinality): 튜플의 수

릴레이션의 특징

■ 속성은 단일 값을 가진다

- 각 속성의 값은 도메인에 정의된 값만을 가지며 그 값은 모두 단일 값이어야 함

■ 속성은 서로 다른 이름을 가진다

- 속성은 한 릴레이션에서 서로 다른 이름을 가져야만 함

■ 한 속성의 값은 모두 같은 도메인 값을 가진다

- 한 속성에 속한 열은 모두 그 속성에서 정의한 도메인 값만 가질 수 있음

■ 속성의 순서는 상관없다

- 속성의 순서가 달라도 릴레이션 스키마는 같음
ex) 릴레이션 스키마에서 (이름, 주소) 순으로 속성을 표시하거나 (주소, 이름) 순으로 표시하여도 상관없음

■ 릴레이션 내의 중복된 튜플은 허용하지 않는다

- 하나의 릴레이션 인스턴스 내에서는 서로 중복된 값을 가질 수 없음. 즉, 모든 튜플은 서로 값이 달라야 함

■ 튜플의 순서는 상관없다

- 튜플의 순서가 달라도 같은 릴레이션임. 관계 데이터 모델의 튜플은 실제적인 값을 가지고 있으며 이 값은 시간이 지남에 따라 데이터의 삭제, 수정, 삽입에 따라 순서가 바뀔 수 있음

릴레이션의 특징

■ 릴레이션의 특징에 위배된 경우

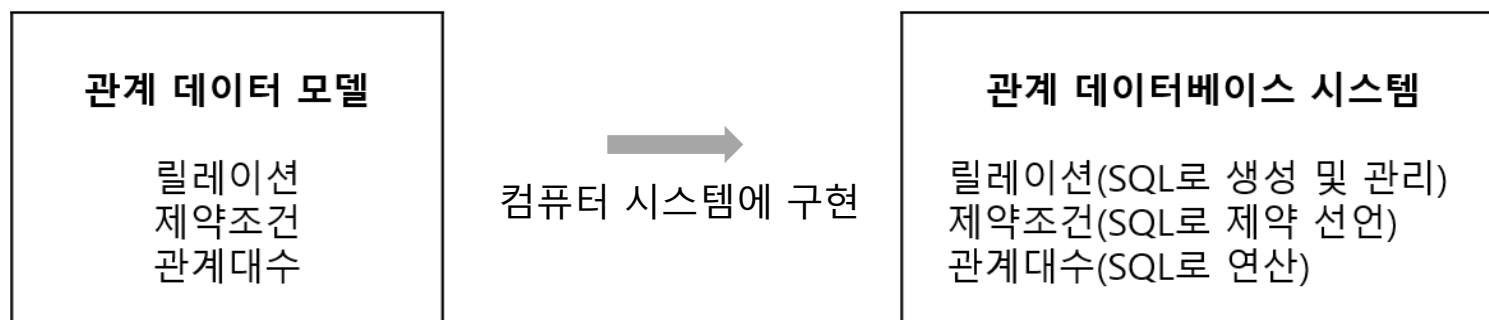
도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구 아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000
5	피겨 교본	굿스포츠	8000
6	피겨 교본, 피겨 기초	굿스포츠	8000

동일한 튜플이 중복되면 안 됨

속성의 값은 단일 값이어야 함

관계 데이터 모델

- 관계 데이터 모델은 데이터를 2차원 테이블 형태인 릴레이션으로 표현함
- 릴레이션에 대한 제약조건과 관계 연산을 위한 관계대수를 정의함



Section 02

무결성 제약조건

키

- 특정 튜플을 식별할 때 사용하는 속성 혹은 속성의 집합
- 릴레이션은 중복된 튜플을 허용하지 않음
 - 각각의 튜플에 포함된 속성들 중 어느 하나(혹은 하나 이상)는 값이 달라야 함.
 - 즉 키가 되는 속성(혹은 속성의 집합)은 반드시 값이 달라서 튜플들을 서로 구별할 수 있어야 함
- 키는 릴레이션 간의 관계를 맺는 데도 사용됨

키

■ 마당서점 데이터베이스

고객

고객번호	이름	주민번호	주소	핸드폰
1	박지성	810101-1111111	영국 맨체스터	000-5000-0001
2	김연아	900101-2222222	대한민국 서울	000-6000-0001
3	장미란	830101-2333333	대한민국 강원도	000-7000-0001
4	추신수	820101-1444444	미국 클리블랜드	000-8000-0001

도서

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

주문

고객번호	도서번호	판매가격	주문일자
1	1	7000	2014-07-01
1	2	13000	2014-07-03
2	5	8000	2014-07-03
3	2	13000	2014-07-04
4	4	35000	2014-07-05
1	3	22000	2014-07-07
4	3	22000	2014-07-07

키

■ 슈퍼키

- 튜플을 유일하게 식별할 수 있는 하나의 속성 혹은 속성의 집합
→ 튜플을 유일하게 식별할 수 있는 값이면 모두 슈퍼키가 될 수 있음
- (고객 릴레이션 예)
 - 고객번호 : 고객별로 유일한 값이 부여되어 있기 때문에 튜플을 식별할 수 있음
 - 이름 : 동명이인이 있을 경우 튜플을 유일하게 식별할 수 없음
 - 주민번호 : 개인별로 유일한 값이 부여되어 있기 때문에 튜플을 식별할 수 있음
 - 주소 : 가족끼리는 같은 정보를 사용하므로 튜플을 식별할 수 없음
 - 핸드폰 : 한 사람이 여러 개의 핸드폰을 사용할 수 있고
반대로 핸드폰을 사용하지 않는 사람이 있을 수 있기 때문에 튜플을 식별할 수 없음
- 고객 릴레이션은 고객번호와 주민번호를 포함한 모든 속성의 집합이 슈퍼키가 됨
 - EX) (주민번호), (주민번호, 이름), (주민번호, 이름, 주소), (주민번호, 이름, 핸드폰), (고객번호), (고객번호, 이름, 주소), (고객번호, 이름, 주민번호, 주소, 핸드폰) 등

키

■ 후보키

- 튜플을 유일하게 식별할 수 있는 속성의 최소 집합
- 주문 릴레이션의 후보키는 2개의 속성을 합한 (고객번호, 도서번호)가 됨
- 2개 이상의 속성으로 이루어진 키를 복합키(composite key)라고 함

■ 기본키

- 여러 후보키 중 하나를 선정하여 대표로 삼는 키
- 후보키가 하나뿐이라면 그 후보키를 기본키로 사용하면 되고, 여러 개라면 릴레이션의 특성을 반영하여 하나를 선택하면 됨
- 기본키 선정 시 고려사항
 - 릴레이션 내 튜플을 식별할 수 있는 고유한 값을 가져야 함
 - NULL 값은 허용하지 않음
 - 키 값의 변동이 일어나지 않아야 함
 - 최대한 적은 수의 속성을 가진 것이어야 함
 - 향후 키를 사용하는 데 있어서 문제 발생 소지가 없어야 함
- 릴레이션 스키마를 표현할 때 기본키는 밑줄을 그어 표시함
 - 릴레이션 이름(속성1, 속성2, ..., 속성N)
 - EX) 고객(고객번호, 이름, 주민번호, 주소, 핸드폰)
 - 도서(도서번호, 도서이름, 출판사, 가격)

키

■ 대리키

- 기본키가 보안을 요하거나, 여러 개의 속성으로 구성되어 복잡하거나, 마땅한 기본키가 없을 때는 일련번호 같은 가상의 속성을 만들어 기본키로 삼는 경우가 있음 이러한 키를 대리키(surrogate key) 혹은 인조키(artificial key)라고 함
 - 대리키는 DBMS나 관련 소프트웨어에서 임의로 생성하는 값으로 사용자가 직관적으로 그 값의 의미를 알 수 없음
-
- 대리키를 사용하도록 변경된 주문 릴레이션

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	1	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	3	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	1	3	22000	2014-07-07
7	4	3	22000	2014-07-07

키

■ 대체키

- 기본키로 선정되지 않은 후보키
- 고객 릴레이션의 경우 고객번호와 주민번호 중 고객번호를 기본키로 정하면 주민번호는 대체키가 됨

■ 외래키

- 다른 릴레이션의 기본키를 참조하는 속성을 말함
- 다른 릴레이션의 기본키를 참조하여 관계 데이터 모델의 특징인 릴레이션 간의 관계(relationship)를 표현함
- 외래키의 특징
 - 관계 데이터 모델의 릴레이션 간의 관계를 표현함
 - 다른 릴레이션의 기본키를 참조하는 속성임
 - 참조하고(외래키) 참조되는(기본키) 양쪽 릴레이션의 도메인은 서로 같아야 함
 - 참조되는(기본키) 값이 변경되면 참조하는(외래키) 값도 변경됨
 - NULL 값과 중복 값 등이 허용됨
 - 자기 자신의 기본키를 참조하는 외래키도 가능함
 - 외래키가 기본키의 일부가 될 수 있음

키

■ 외래키

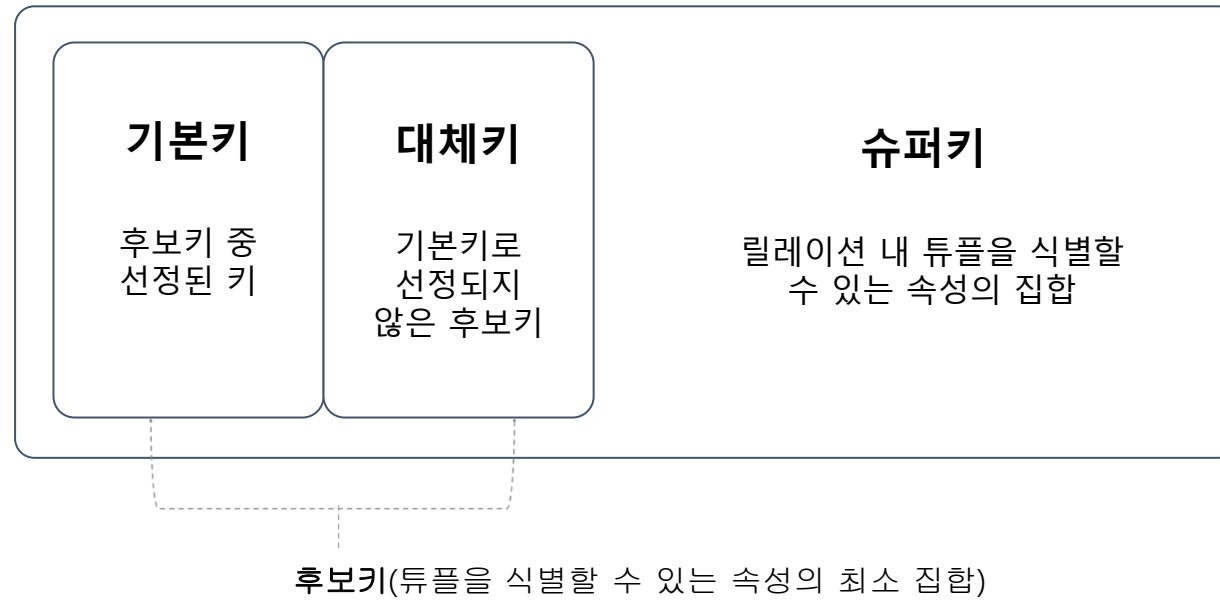
- 외래키 사용 시 참조하는 릴레이션과 참조되는 릴레이션이 꼭 다른 릴레이션일 필요는 없음.
- 즉, 자기 자신의 기본키를 참조할 수도 있음

참조

기본키				외래키
선수번호	이름	주소	멘토번호	
1	박지성	영국 맨체스타	NULL	
2	김연아	대한민국 서울	3	
3	장미란	대한민국 강원도	4	
4	추신수	미국 클리블랜드	NULL	

키

■ 요약: 키의 포함 관계



키

■ 릴레이션 간의 참조 관계

고객

고객번호	이름	주민번호	주소	핸드폰
1	박지성	810101-1111111	영국 맨체스터	000-5000-0001
2	김연아	900101-2222222	대한민국 서울	000-6000-0001
3	장미란	830101-2333333	대한민국 강원도	000-7000-0001
4	추신수	820101-1444444	미국 클리블랜드	000-8000-0001

기본키

도서

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

기본키

참조

외래키

참조

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	1	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	3	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	1	3	22000	2014-07-07
7	4	3	22000	2014-07-07

기본키

무결성 제약조건

■ 데이터 무결성(integrity, 無缺性)

- 데이터베이스에 저장된 데이터의 일관성과 정확성을 지키는 것

■ 도메인 무결성 제약조건

- 도메인 제약(domain constraint)이라고도 하며, 릴레이션 내의 튜플들이 각 속성의 도메인에 지정된 값만을 가져야 한다는 조건
- SQL 문에서 데이터 형식(type), 널(null/not null), 기본 값(default), 체크(check) 등을 사용하여 지정할 수 있음

■ 개체 무결성 제약조건

- 기본키 제약(primary key constraint)이라고도 함
- 릴레이션은 기본키를 지정하고 그에 따른 무결성 원칙, 즉 기본키는 NULL 값을 가져서는 안 되며 릴레이션 내에 오직 하나의 값만 존재해야 한다는 조건임

■ 참조 무결성 제약조건

- 외래키 제약(foreign key constraint)이라고도 하며, 릴레이션 간의 참조 관계를 선언하는 제약조건임
- 자식 릴레이션의 외래키는 부모 릴레이션의 기본키와 도메인이 동일해야 하며, 자식 릴레이션의 값이 변경될 때 부모 릴레이션의 제약을 받는다는 것임

무결성 제약조건

■ 제약조건 정리

구분	도메인	키	
	도메인 무결성 제약조건	개체 무결성 제약조건	참조 무결성 제약조건
제약 대상	속성	튜플	속성과 튜플
같은 용어	도메인 제약 (Domain Constraint)	기본키 제약 (Primary Key Constraint)	외래키 제약 (Foreign Key Constraint)
해당되는 키	-	기본키	외래키
NULL 값 허용 여부	허용	불가	허용
릴레이션 내 제약조건의 개수	속성의 개수와 동일	1개	0~여러 개
기타	<ul style="list-style-type: none"> • 튜플 삽입, 수정 시 제약 사항 우선 확인 	<ul style="list-style-type: none"> • 튜플 삽입/수정 시 제약 사항 우선 확인 	<ul style="list-style-type: none"> • 튜플 삽입/수정 시 제약사항 우선 확인 • 부모 릴레이션의 튜플 수정/삭제 시 제약사항 우선 확인

무결성 제약조건의 수행

■ 개체 무결성 제약조건

- 삽입: 기본키 값이 같으면 삽입이 금지됨
- 수정: 기본키 값이 같거나 NULL로도 수정이 금지됨
- 삭제: 특별한 확인이 필요하지 않으며 즉시 수행함

■ 개체 무결성 제약조건 수행 예

(501, 남슬찬, 1001)

↓ 삽입 거부

학번	이름	학과코드
501	박지성	1001
401	김연아	2001
402	장미란	2001
502	추신수	1001

(NULL, 남슬찬, 1001)

↓ 삽입 거부

학번	이름	학과코드
501	박지성	1001
401	김연아	2001
402	장미란	2001
502	추신수	1001

무결성 제약조건의 수행

■ 참조 무결성 제약조건

• 삽입

- 학생(자식 릴레이션): 참조받는 테이블에 외래키 값이 없으므로 삽입이 금지됨
- 학과(부모 릴레이션): 튜플 삽입한 후 수행하면 정상적으로 진행

학생(자식 릴레이션)

학번	이름	학과코드
501	박지성	1001
401	김연아	2001
402	장미란	2001
502	추신수	1001

학과(부모 릴레이션)

학과코드	학과명
1001	컴퓨터학과
2001	체육학과

참조

무결성 제약조건의 수행

■ 참조 무결성 제약조건

• 삭제

- 학과(부모 릴레이션) : 참조하는 테이블을 같이 삭제할 수 있어서 금지하거나 다른 추가 작업이 필요함
- 학생(자식 릴레이션) : 바로 삭제 가능함

→ ※ 부모 릴레이션에서 튜플을 삭제할 경우 참조 무결성 조건을 수행하기 위한 고려사항

- ✓ 즉시 작업을 중지
- ✓ 자식 릴레이션의 관련 튜플을 삭제
- ✓ 초기에 설정된 다른 어떤 값으로 변경
- ✓ NULL 값으로 설정

• 수정

- 삭제와 삽입 명령이 연속해서 수행됨
- 부모 릴레이션의 수정이 일어날 경우 삭제 옵션에 따라 처리된 후 문제가 없으면 다시 삽입 제약조건에 따라 처리됨

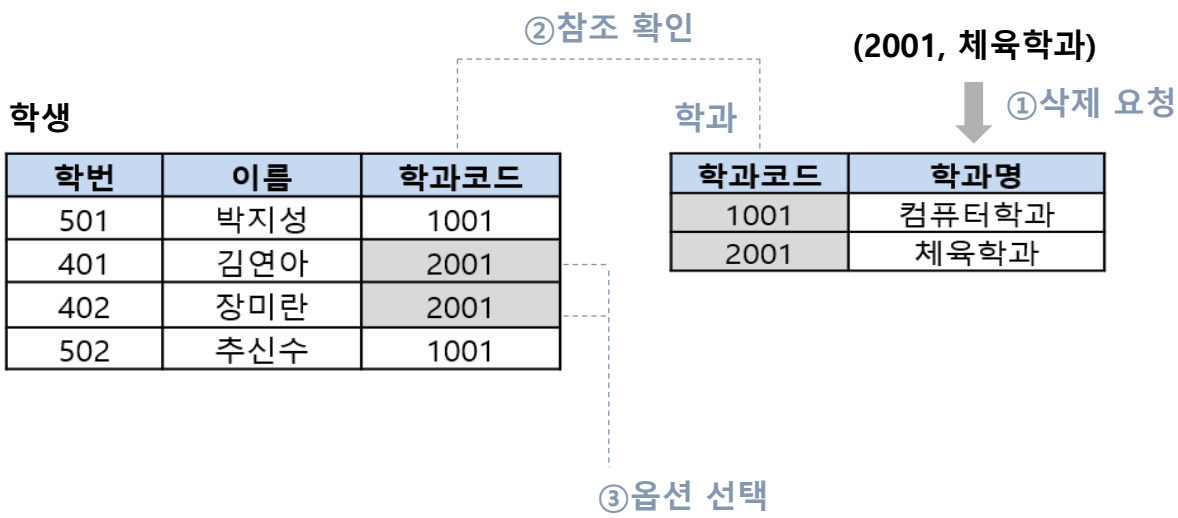
무결성 제약조건의 수행

■ 참조 무결성 제약조건의 옵션(부모 릴레이션에서 튜플을 삭제할 경우)

명령어	의미	예
RESTRICTED	자식 릴레이션에서 참조하고 있을 경우 부모 릴레이션의 삭제 작업을 거부함	학과 릴레이션의 튜플 삭제 거부
CASCADE	자식 릴레이션의 관련 튜플을 같이 삭제 처리함	학생 릴레이션의 관련 튜플을 삭제
DEFAULT	자식 릴레이션의 관련 튜플을 미리 설정해둔 값으로 변경함	학생 릴레이션의 학과가 다른 학과로 자동 배정
NULL	자식 릴레이션의 관련 튜플을 NULL 값으로 설정함(NULL 값을 허가한 경우)	학과 릴레이션의 학과가 NULL 값으로 변경

무결성 제약조건의 수행

■ 참조 무결성 제약조건



Section 03

관계대수

관계대수

■ 관계대수(relational algebra, 關係代數)

- 릴레이션에서 원하는 결과를 얻기 위해 수학의 대수와 같은 연산을 이용하여 질의하는 방법을 기술하는 언어

■ 관계대수와 관계해석

- 관계대수 : 어떤 데이터를 어떻게 찾는지에 대한 처리 절차를 명시하는 절차적인 언어이며, DBMS 내부의 처리 언어로 사용됨
- 관계해석 : 어떤 데이터를 찾는지만 명시하는 선언적인 언어로 관계대수와 함께 관계 DBMS의 표준 언어인 SQL의 이론적인 기반을 제공함

■ ※ 관계대수와 관계해석은 모두 관계 데이터 모델의 중요한 언어이며 실제 동일한 표현 능력을 가지고 있음

관계대수

■ 릴레이션(relation)의 수학적 개념

예) $A = \{2, 4\}$, $B = \{1, 3, 5\}$ 일 때

$$A \times B = \{(2, 1), (2, 3), (2, 5), (4, 1), (4, 3), (4, 5)\}$$

릴레이션 R은 카티전 프로덕트의 부분집합으로 정의

예) $R1 = \{(2, 1), (4, 1)\}$, $R2 = \{(2, 1), (2, 3), (2, 5)\}$, $R3 = \{(2, 3), (2, 5), (4, 3), (4, 5)\} \dots$

원소 개수가 n인 집합 S의 부분집합의 개수는 2^n 이므로, 카티전 프로덕트 $A \times B$ 의 부분집합의 개수는 $2^{|A| \times |B|}$ 임

카티전 프로덕트의 기초 집합 A, B 각각이 가질 수 있는 값의 범위를 도메인(domain)이라고 함
즉 집합 A의 도메인은 $\{2, 4\}$ 임

릴레이션 역시 집합이므로 집합에서 집합에서 가능한 연산은 합집합(\cup), 교집합(\cap), 카티전 프로덕트(\times) 등이 있음

$$R1 \cup R2 = \{(2, 1), (4, 1), (2, 3), (2, 5)\}$$

$$R1 \cap R2 = \{(2, 1)\}$$

관계대수

■ 관계의 수학적 의미

• 릴레이션의 현실 세계 적용

→ ex) 학번={2, 4}, 과목={데이터베이스, 자료구조, 프로그래밍}일 때

두 집합의 카티전 프로덕트 학번×과목은 학번 원소와 과목 원소의 순서쌍의 집합임

즉, 학번×과목={(2, 데이터베이스), (2, 자료구조), (2, 프로그래밍), (4, 데이터베이스), (4, 자료구조), (4, 프로그래밍)}을 말함

학번×과목의 각 원소는 학생이 과목을 수강할 수 있는 모든 경우를 나열한 것임

수강={(2, 데이터베이스), (2, 자료구조), (4, 프로그래밍)}은 카티전 프로덕트 학번×과목의 부분집합으로 하나의 릴레이션 인스턴스임

수강 릴레이션의 튜플은 위에서 나열한 여섯 개 원소 중 하나로, 아래 수강 테이블을 데이터베이스에서는 릴레이션 (relation)이라고 함

수강

학번	과목
2	데이터베이스
2	자료구조
4	프로그래밍

관계대수

■ 관계대수 연산자

연산자 종류	대상	연산자 이름	기호	설명
기본	단항	선택	σ	릴레이션에서 조건에 만족하는 튜플을 선택
기본	단항	프로젝션	π	릴레이션의 속성을 선택
추가	단항	개명	ρ	릴레이션이나 속성의 이름을 변경
유도	이항	디비전	\div	부모 릴레이션에 포함된 튜플의 값을 모두 갖고 있는 튜플을 분자 릴레이션에서 추출
기본	이항	합집합	\cup	두 릴레이션의 합집합
기본	이항	차집합	$-$	두 릴레이션의 차집합
유도	이항	교집합	\cap	두 릴레이션의 교집합
기본	이항	카디전 프로덕트	\times	두 릴레이션에 속한 모든 튜플의 집합

관계대수

■ 관계대수 연산자

연산자 종류	대상	연산자 이름		기호	설명
유도	이항	조인	세타	\bowtie_{θ}	두 릴레이션 간의 비교 조건에 만족하는 집합
			동등	\bowtie	두 릴레이션 간의 같은 값을 가진 집합
			자연	\bowtie_N	동등 조인에서 중복 속성을 제거
			세미	left	자연 조인 후 오른쪽 속성을 제거
				right	자연 조인 후 왼쪽 속성을 제거
		외부	left	\bowtie_{\leftarrow}	<ul style="list-style-type: none"> 자연 조인 후 각각 왼쪽(left), 오른쪽(right), 양쪽(full)의 모든 값을 결과로 추출 조인이 실패(또는 값이 없을 경우)한 쪽의 값을 NULL로 채움
			right	\bowtie_{\rightarrow}	
			full	\bowtie_{\times}	

선택과 프로젝션

■ 선택(selection)

- 릴레이션의 튜플을 추출하기 위한 연산임
- 하나의 릴레이션을 대상으로 하는 단항 연산자며, 찾고자 하는 튜플의 조건(predicate)을 명시하고 그 조건에 만족하는 튜플을 반환함
- 형식 : $\sigma_{\langle \text{조건} \rangle}(R)$

질의 2-1 마당서점에서 판매하는 도서 중 8,000원 이하인 도서를 검색하시오.

$\sigma_{\text{가격} \leq 8000}(\text{도서})$

도서

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

$\sigma_{\text{가격} \leq 8000}$
(도서)



도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
5	피겨 교본	굿스포츠	8000

선택과 프로젝션

■ 프로젝션(projection)

- 릴레이션의 속성을 추출하기 위한 연산으로 단항 연산자임
- 형식 : $\pi_{\langle \text{속성리스트} \rangle} (R)$

질의 2-2 신간도서 안내를 위해
고객의 (이름, 주소, 핸드폰)이 적힌
카탈로그 주소록을 만드시오.

$\pi_{\text{이름, 주소, 핸드폰}}(\text{고객})$

고객

고객번호	이름	주민번호	주소	핸드폰
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001
2	김연아	900905-2222222	대한민국 서울	000-6000-0001
3	장미란	831009-2333333	대한민국 강원도	000-7000-0001
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001

$\pi_{\text{이름, 주소, 핸드폰}}(\text{고객})$

이름	주소	핸드폰
박지성	영국 맨체스타	000-5000-0001
김연아	대한민국 서울	000-6000-0001
장미란	대한민국 강원도	000-7000-0001
추신수	미국 클리블랜드	000-8000-0001

셀렉션과 프로젝션

■ 합집합

- 두 개의 릴레이션을 합하여 하나의 릴레이션을 반환함
- 이 때 두 개의 릴레이션은 서로 같은 속성 순서와 도메인을 가져야 함
- 형식 : $R \cup S$

질의 2-3 마당서점은 지점A와 지점B가 있다. 두 지점의 도서는 각 지점에서 관리하며 릴레이션 이름은 각각 도서A, 도서B다. 마당서점의 도서를 하나의 릴레이션으로 보이시오.

도서A \cup 도서B

도서A

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000

U

도서B

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000



도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

셀렉션과 프로젝션

■ 교집합

- 합병가능한 두 릴레이션을 대상으로 하며, 두 릴레이션이 공통으로 가지고 있는 튜플을 반환함
- 형식 : $R \cap S$

질의 2-4 마당서점의 두 지점에서 동일하게 보유하고 있는 도서 목록을 보이시오.

도서A \cap 도서B

도서A

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000

\cap

도서B

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000



도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000

집합 연산

■ 차집합

- 첫 번째 릴레이션에는 속하고 두 번째 릴레이션에는 속하지 않는 튜플을 반환함
- 형식 : $R - S$

질의 2-5 마당서점 두 지점 중 지점 A에서만 보유하고 있는 도서 목록을 보이시오.

도서A - 도서B

도서A

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000

-

도서B

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000



도서번호	도서이름	출판사	가격
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000

집합 연산

■ 카티전 프로덕트(cartesian product)

- 두 릴레이션을 연결시켜 하나로 합칠 때 사용함
- 결과 릴레이션은 첫 번째 릴레이션의 오른쪽에 두 번째 릴레이션의 모든 튜플을 순서대로 배열하여 반환함
- 결과 릴레이션의 차수는 두 릴레이션의 차수의 합이며, 카디널리티는 두 릴레이션의 카디널리티의 곱임
- 형식 : $R \times S$

질의 2-6 고객 릴레이션과 주문 릴레이션의 카티전 프로덕트를 구하시오
(결과가 많으므로 튜플을 일부 삭제한 릴레이션을 사용함).

고객 \times 주문

집합 연산

고객

고객번호	이름	주소	핸드폰
1	박지성	영국 맨체스타	000-5000-0001
2	김연아	대한민국 서울	000-6000-0001
3	장미란	대한민국 강원도	000-7000-0001

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	2	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	1	2	13000	2014-07-04

×



고객번호	이름	주소	핸드폰	주문번호	고객번호	도서번호	판매가격	주문일자
1	박지성	영국 맨체스타	000-5000-0001	1	2	1	7000	2014-07-01
1	박지성	영국 맨체스타	000-5000-0001	2	1	2	13000	2014-07-03
1	박지성	영국 맨체스타	000-5000-0001	3	2	5	8000	2014-07-03
1	박지성	영국 맨체스타	000-5000-0001	4	1	2	13000	2014-07-04
2	김연아	대한민국 서울	000-6000-0001	1	2	1	7000	2014-07-01
2	김연아	대한민국 서울	000-6000-0001	2	1	2	13000	2014-07-03
2	김연아	대한민국 서울	000-6000-0001	3	2	5	8000	2014-07-03
2	김연아	대한민국 서울	000-6000-0001	4	1	2	13000	2014-07-04
3	장미란	대한민국 강원도	000-7000-0001	1	2	1	7000	2014-07-01
3	장미란	대한민국 강원도	000-7000-0001	2	1	2	13000	2014-07-03
3	장미란	대한민국 강원도	000-7000-0001	3	2	5	8000	2014-07-03
3	장미란	대한민국 강원도	000-7000-0001	4	1	2	13000	2014-07-04

조인

- 두 릴레이션의 공통 속성을 기준으로 속성 값이 같은 튜플을 수평으로 결합하는 연산임
- 조인을 수행하기 위해서는 두 릴레이션의 조인에 참여하는 속성이 서로 동일한 도메인으로 구성되어야 함
- 조인 연산의 결과는 공통 속성의 속성 값이 동일한 튜플만을 반환함

■ 형식 : $R \bowtie_C S = \sigma_c (R \times S)$ (c 는 조인조건)

■ 조인 연산의 구분

- 기본연산 : 세타조인(\bowtie_θ), 동등조인(\bowtie), 자연조인(\bowtie_N)
- 확장된 조인 연산 : 세미조인($\bowtie_{\leftarrow}, \bowtie_{\rightarrow}$), 외부조인($\bowtie_{\leftarrow\leftarrow}, \bowtie_{\rightarrow\rightarrow}, \bowtie_{\leftrightarrow}$)

조인

■ 세타조인과 동등조인

- 세타조인(theta join, θ)

- 조인에 참여하는 두 릴레이션의 속성 값을 비교하여 조건을 만족하는 튜플만 반환함

- 세타조인의 조건은 $\{=, \neq, \leq, \geq, <, >\}$ 중 하나가 됨

- 형식 : $R \bowtie_{(r \text{ 조건 } s)} S$

- 동등조인(equi join)

- 세타조인에서 $=$ 연산자를 사용한 조인을 말함. 보통 조인 연산이라고 하면 동등조인을 지칭함

- 형식 : $R \bowtie_{(r = s)} S$

질의 2-7 고객과 고객의 주문 사항을 모두 보이시오.

고객 $\bowtie_{\text{고객.고객번호=주문.고객번호}}$ 주문


조인

고객

고객번호	이름	주소	핸드폰
1	박지성	영국 맨체스타	000-5000-0001
2	김연아	대한민국 서울	000-6000-0001
3	장미란	대한민국 강원도	000-7000-0001
4	추신수	미국 클리블랜드	000-8000-0001

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	2	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	1	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	5	3	22000	2014-07-07
7	4	3	22000	2014-07-07

고객  고객.고객번호=주문.고객번호 주문

고객번호	이름	주소	핸드폰	주문번호	고객번호	도서번호	판매가격	주문일자
1	박지성	영국 맨체스타	000-5000-0001	2	1	2	13000	2014-07-03
1	박지성	영국 맨체스타	000-5000-0001	4	1	2	13000	2014-07-04
2	김연아	대한민국 서울	000-6000-0001	1	2	1	7000	2014-07-01
2	김연아	대한민국 서울	000-6000-0001	3	2	5	8000	2014-07-03
4	추신수	미국 클리블랜드	000-8000-0001	5	4	4	35000	2014-07-05
4	추신수	미국 클리블랜드	000-8000-0001	7	4	3	22000	2014-07-07

조인

■ 자연조인(natural join)

- 동등조인에서 조인에 참여한 속성이 두 번 나오지 않도록 두 번째 속성을 제거한 결과를 반환함
- 형식 : $R \bowtie_{N(r, s)} S$

질의 2-8 고객과 고객의 주문 사항을 모두 보여주되 같은 속성은 한 번만 표시하시오.

고객 $\bowtie_{N(\text{고객.고객번호}=\text{주문.고객번호})}$ 주문


조인

고객

고객번호	이름	주소	핸드폰
1	박지성	영국 맨체스터	000-5000-0001
2	김연아	대한민국 서울	000-6000-0001
3	장미란	대한민국 강원도	000-7000-0001
4	추신수	미국 클리블랜드	000-8000-0001

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	2	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	1	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	5	3	22000	2014-07-07
7	4	3	22000	2014-07-07

고객  N(고객.고객번호=주문.고객번호) 주문

고객번호	이름	주소	핸드폰	주문번호	도서번호	판매가격	주문일자
1	박지성	영국 맨체스터	000-5000-0001	2	2	13000	2014-07-03
1	박지성	영국 맨체스터	000-5000-0001	4	2	13000	2014-07-04
2	김연아	대한민국 서울	000-6000-0001	1	1	7000	2014-07-01
2	김연아	대한민국 서울	000-6000-0001	3	5	8000	2014-07-03
4	추신수	미국 클리블랜드	000-8000-0001	5	4	35000	2014-07-05
4	추신수	미국 클리블랜드	000-8000-0001	7	3	22000	2014-07-07

조인

■ 외부조인과 세미조인

• 외부조인(outer join)

- 자연조인 시 조인에 실패한 튜플을 모두 보여주되 값이 없는 대응 속성에는 NULL 값을 채워서 반환
- 모든 속성을 보여주는 기준 릴레이션 위치에 따라 왼쪽(left) 외부조인, 오른쪽(right) 외부조인, 완전(full, 양쪽) 외부조인으로 나뉨
- 형식 : 왼쪽(left) 외부조인 - $R \bowtie_{(r, s)} S$
완전(full) 외부조인 - $R \bowtie_{(r, s)} S$
오른쪽(right) 외부조인 - $R \bowtie_{(r, s)} S$

질의 2-9 마당서점의 고객과 고객의 주문 내역을 보이시오.

- ① 고객 기준으로 주문내역이 없는 고객도 모두 보이시오.
- ② 주문내역이 없는 고객과, 고객 릴레이션에 고객번호가 없는 주문을 모두 보이시오.
- ③ 주문내역 기준으로 고객 릴레이션에 고객번호가 없는 주문도 모두 보이시오.

- ① 고객 $\bowtie_{(고객.고객번호, 주문.고객번호)}$ 주문
- ② 고객 $\bowtie_{(고객.고객번호, 주문.고객번호)}$ 주문
- ③ 고객 $\bowtie_{(고객.고객번호, 주문.고객번호)}$ 주문

조인

고객

고객번호	이름
1	박지성
2	김연아
3	장미란
4	추신수

주문

주문번호	고객번호	판매가격
1	2	7000
2	1	13000
3	2	8000
4	1	13000
5	4	35000
6	5	22000
7	4	22000

고객 ⋈ 고객.고객번호=주문.고객번호 주문

고객번호	이름	주문번호	판매가격
1	박지성	2	13000
1	박지성	4	13000
2	김연아	1	7000
2	김연아	3	8000
3	장미란	NULL	NULL
4	추신수	5	35000
4	추신수	7	22000

① 왼쪽 외부조인

고객 ⋈ 고객.고객번호=주문.고객번호 주문

고객번호	이름	주문번호	판매가격
1	박지성	2	13000
1	박지성	4	13000
2	김연아	1	7000
2	김연아	3	8000
3	장미란	NULL	NULL
4	추신수	5	35000
4	추신수	7	22000
5	NULL	6	22000

② 완전 외부조인

고객 ⋈ 고객.고객번호=주문.고객번호 주문

이름	주문번호	고객번호	판매가격
김연아	1	2	7000
박지성	2	1	13000
김연아	3	2	8000
박지성	4	1	13000
추신수	5	4	35000
NULL	6	5	22000
추신수	7	4	22000

③ 오른쪽 외부조인

조인

■ 외부조인과 세미조인

- 세미조인(semi join)

→ 자연조인을 한 후 두 릴레이션 중 한쪽 릴레이션의 결과만 반환하며, 기호에서 닫힌 쪽 릴레이션의 튜플만 반환함

→ 형식 : $R \bowtie_{(r, s)} S$

조인

질의 2-10 마당서점의 고객 중 주문 내역이 있는 고객의 고객 정보를 보이시오.

고객 \bowtie 고객.고객번호, 주문.고객번호 주문

고객

고객번호	이름	주소	핸드폰
1	박지성	영국 맨체스타	000-5000-0001
2	김연아	대한민국 서울	000-6000-0001
3	장미란	대한민국 강원도	000-7000-0001
4	추신수	미국 클리블랜드	000-8000-0001

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	2	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	1	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	5	3	22000	2014-07-07
7	4	3	22000	2014-07-07

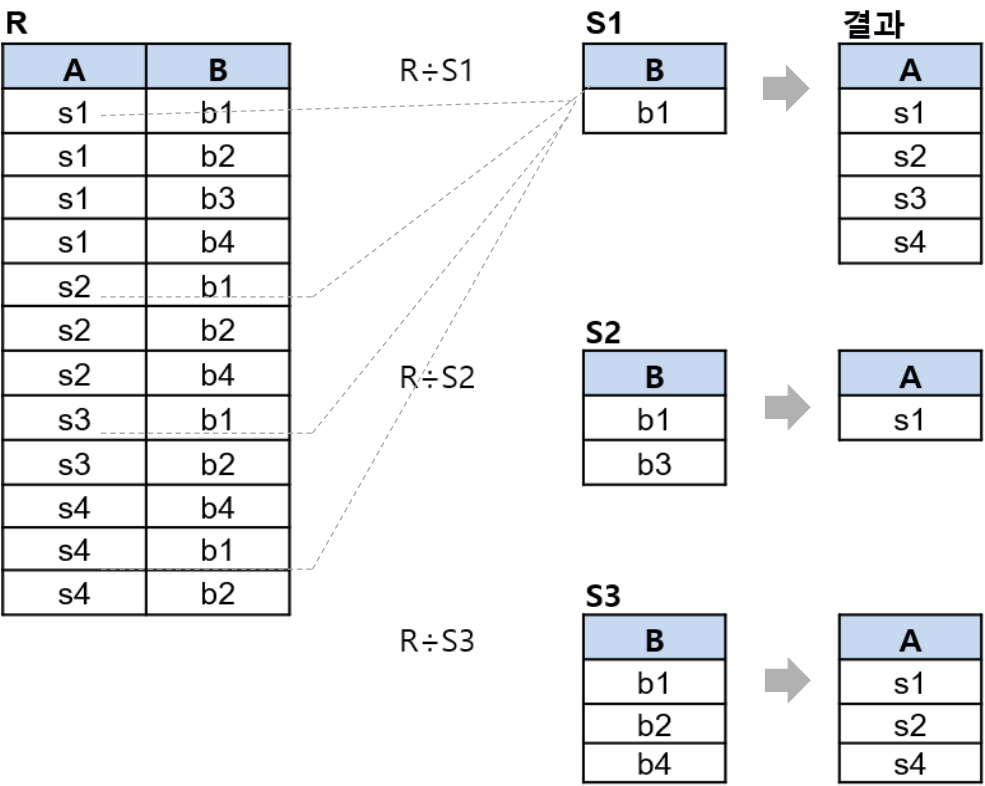


고객 \bowtie (고객.고객번호, 주문.고객번호) 주문

고객번호	이름	주소	핸드폰
1	박지성	영국 맨체스타	000-5000-0001
2	김연아	대한민국 서울	000-6000-0001
4	추신수	미국 클리블랜드	000-8000-0001

디비전

- 릴레이션의 속성 값의 집합으로 연산을 수행함
- 형식 : $R \div S$



관계대수 예제

■ 셀렉션, 프로젝션, 집합연산의 복합사용

질의 2-11 마당서점의 도서 중 가격이 8,000원 이하인 도서이름과 출판사를 보이시오.

■ 마당서점의 지점이 하나인 경우

- $\pi_{\text{도서이름, 출판사}}(\sigma_{\text{가격} \leq 8000} \text{도서})$

도서

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

↓ $\sigma_{\text{가격} \leq 8000} (\text{도서})$

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
5	피겨 교본	굿스포츠	8000



$\pi_{\text{도서이름, 출판사}}$

도서이름	출판사
축구의 역사	굿스포츠
피겨 교본	굿스포츠

관계대수 예제

■ 마당서점의 지점이 둘 이상인 경우

$\pi_{\text{도서이름, 출판사}}((\sigma_{\text{가격} \leq 8000} \text{도서A}) \cup (\sigma_{\text{가격} \leq 8000} \text{도서B}))$

도서A

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
2	축구아는 여자	나무수	13000
3	축구의 이해	대한미디어	22000

도서B

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
4	골프 바이블	대한미디어	35000
5	피겨 교본	굿스포츠	8000

$\sigma_{\text{가격} \leq 8000}$ (도서A)

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000

$\sigma_{\text{가격} \leq 8000}$ (도서B)

도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
5	피겨 교본	굿스포츠	8000

U



도서번호	도서이름	출판사	가격
1	축구의 역사	굿스포츠	7000
5	피겨 교본	굿스포츠	8000



$\pi_{\text{도서이름, 출판사}}$

도서이름	출판사
축구의 역사	굿스포츠
피겨 교본	굿스포츠

관계대수 예제

■ 카티전 프로덕트를 사용한 연산과 조인을 사용한 연산

질의 2-12 마당서점의 박지성 고객의 거래 내역 중 주문번호, 이름, 가격을 보이시오.

■ 카티전 프로덕트를 사용한 연산

- $\pi_{\text{주문.주문번호, 고객.이름, 주문.판매가격}}(\sigma_{\text{고객.고객번호=주문.고객번호 AND 고객.이름='박지성'}}(\text{고객} \times \text{주문}))$

관계대수 예제

고객

고객번호	이름	주민번호	주소	핸드폰
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001
2	김연아	900905-2222222	대한민국 서울	000-6000-0001
3	장미란	831009-2333333	대한민국 강원도	000-7000-0001
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	1	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	3	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	1	3	22000	2014-07-07
7	4	3	22000	2014-07-07

고객 × 주문

고객번호	이름	주민번호	주소	핸드폰	주문번호	고객번호	도서번호	판매가격	주문일자
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	1	1	1	7000	2014-07-01
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	2	1	2	13000	2014-07-03
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	3	2	5	8000	2014-07-03
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	4	3	2	13000	2014-07-04
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	5	4	4	35000	2014-07-05
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	6	1	3	22000	2014-07-07
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	7	4	3	22000	2014-07-07
2	김연아	900905-2222222	대한민국 서울	000-6000-0001	1	1	1	7000	2014-07-01
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001	5	4	4	35000	2014-07-05
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001	6	1	3	22000	2014-07-07
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001	7	4	3	22000	2014-07-07

$\sigma_{\text{고객.고객번호=주문.고객번호 AND 고객.이름='박지성'}}$

고객번호	이름	주민번호	주소	핸드폰	주문번호	고객번호	도서번호	판매가격	주문일자
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	1	1	1	7000	2014-07-01
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	2	1	2	13000	2014-07-03
1	박지성	810225-1111111	영국 맨체스터	000-5000-0001	6	1	3	22000	2014-07-07

$\pi_{\text{주문.주문번호, 고객.이름, 주문.판매가격}}$

주문번호	이름	판매가격
1	박지성	7000
2	박지성	13000
6	박지성	22000

관계대수 예제

■ 카티전 프로덕트를 사용한 연산과 조인을 사용한 연산

■ 조인을 사용한 연산

- $\pi_{\text{주문번호, 이름, 판매가격}}(\sigma_{\text{이름='박지성'}}(\text{고객} \bowtie \text{고객.고객번호=주문.고객번호 주문}))$

관계대수 예제

고객

고객번호	이름	주민번호	주소	핸드폰
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001
2	김연아	900905-2222222	대한민국 서울	000-6000-0001
3	장미란	831009-2333333	대한민국 강원도	000-7000-0001
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001

주문

주문번호	고객번호	도서번호	판매가격	주문일자
1	1	1	7000	2014-07-01
2	1	2	13000	2014-07-03
3	2	5	8000	2014-07-03
4	3	2	13000	2014-07-04
5	4	4	35000	2014-07-05
6	1	3	22000	2014-07-07
7	4	3	22000	2014-07-07

고객 ⋈ 고객.고객번호=주문.고객번호 주문

고객번호	이름	주민번호	주소	핸드폰	주문번호	고객번호	도서번호	판매가격	주문일자
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001	1	1	1	7000	2014-07-01
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001	2	1	2	13000	2014-07-03
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001	6	1	3	22000	2014-07-07
2	김연아	900905-2222222	대한민국 서울	000-6000-0001	3	2	5	8000	2014-07-03
3	장미란	831009-2333333	대한민국 강원도	000-7000-0001	4	3	2	13000	2014-07-04
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001	5	4	4	35000	2014-07-05
4	추신수	820713-1444444	미국 클리블랜드	000-8000-0001	7	4	3	22000	2014-07-07

$\sigma_{\text{이름} = \text{'박지성'}}$

고객번호	이름	주민번호	주소	핸드폰	주문번호	고객번호	도서번호	판매가격	주문일자
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001	1	1	1	7000	2014-07-01
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001	2	1	2	13000	2014-07-03
1	박지성	810225-1111111	영국 맨체스타	000-5000-0001	6	1	3	22000	2014-07-07

$\pi_{\text{주문번호, 이름, 판매가격}}$

주문번호	이름	판매가격
1	박지성	7000
2	박지성	13000
6	박지성	22000

오픈플랫폼활용

Copyright © Lee Seungwon Professor
All rights reserved.

<Q&A : lsw@kopo.ac.kr>