

데이터베이스 설계 (Database Design)

Lecture 05: 관계 데이터 연산

담당교수: 전강욱(컴퓨터공학부)
kw.chon@koreatech.ac.kr

지난시간 복습

■ 관계 데이터 모델에 관해 학습

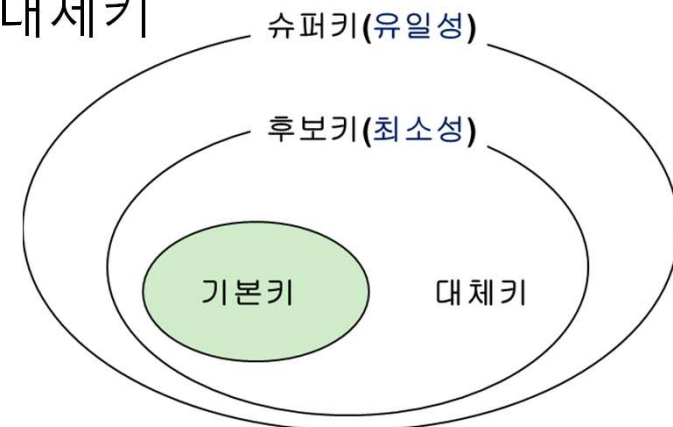
- 개념적 구조를 **행과 열로 구성된 표(릴레이션) 형태**로 다루는 모델
- 관련 용어: 릴레이션, 튜플, 속성, 차수, 카디널리티 등

■ 릴레이션(relation)의 특성

- 튜플의 유일성, 무순서
- 속성의 무순서, 원자성

■ 릴레이션의 키

- 튜플들을 유일하게 구별하는 속성 또는 속성들의 집합
 - 유일성, 최소성
- 키의 종류: 기본키, 수퍼키, 외래키, 대체키



지난시간 복습 (계속)

■ 관계 데이터 모델의 제약 (무결성 제약조건)

- 개체 무결성 제약조건(entity integrity constraint): 기본키를 구성하는 모든 속성은 null 값이 올 수 없음
- 참조 무결성 제약조건(referential integrity constraint): 외래키는 참조할 수 없는 값을 가질 수 없음(단, null 값은 허용)

| WORKS_ON | | | PROJECT | | | |
|-----------|-----|-------|-----------------|---------|-----------|------|
| ESSN | PNO | HOURS | PNAME | PNUMBER | PLOCATION | DNUM |
| 123456789 | 1 | 32.5 | ProductX | 1 | Bellaire | 5 |
| 123456789 | 2 | 7.5 | ProductY | 2 | Sugarland | 5 |
| 666884444 | 3 | 40 | ProductZ | 3 | Houston | 5 |
| 453453453 | 1 | 20 | Computerization | 10 | Stafford | 4 |
| 453453453 | 2 | 20 | Reorganization | 20 | Houston | 1 |
| 333445555 | 2 | 10 | Newbenefits | 30 | Stafford | 4 |
| 333445555 | 3 | 10 | | | | |

외래키(FK) → 기본키(PK)

복습 문제

- 릴레이션의 속성을 튜플이라고 한다. ()
- 속성 하나가 가질 수 있는 모든 값의 집합을 해당 속성의 () 라고 한다.
- 관계 데이터 모델에서는 속성의 값으로 더는 분해할 수 없는 원자 값만 사용할 수 있다. ()
- Null 값은 특정 속성에 해당되는 값이 없음을 나타낸다. ()
- 최소성과 유일성을 모두 만족하는 키를 수퍼키라고한다. ()

복습 문제 (계속)

- 기본키의 속성은 null이 아니어야 하고, 중복이 불가능해야 한다. ()
- 위의 제약 조건은 ()라고 한다.
- 동일한 구조의 관점에서 모든 데이터를 논리적으로 구성하며 선언적인 질의어를 통한 데이터 접근을 제공하는 데이터 모델을 (A)(이)라고 하고, 하나의 릴레이션에서 속성 전체 개수를 릴레이션의 (B)(이)라하고, 하나의 릴레이션에서 튜플의 전체 개수를 릴레이션의 (C)(이)라 한다.

NCS 정보

- 능력 단위명 : 데이터베이스 구현
- 능력 단위요소 : 데이터베이스 오브젝트 생성하기
- 학습목표(수행 준거) :
 - 3.2 물리 데이터베이스 설계에 따라 데이터베이스 오브젝트를 생성하기 위한 DDL(Data Definition Language)을 작성할 수 있고 생성된 오브젝트에 대한 유효성 여부를 검사할 수 있다.
 - 3.3 생성된 데이터베이스 오브젝트의 적정성과 무결성을 판단하고 데이터 베이스 오브젝트 명세를 포함한 완료보고서를 작성할 수 있다.

NCS 정보(계속)

■ 지식

- DDL과 DML 문법 이해 및 사용법

■ 기술

- DDL로 테이블을 생성/삭제하는 능력
- DDL로 인덱스를 생성하는 능력

■ 태도

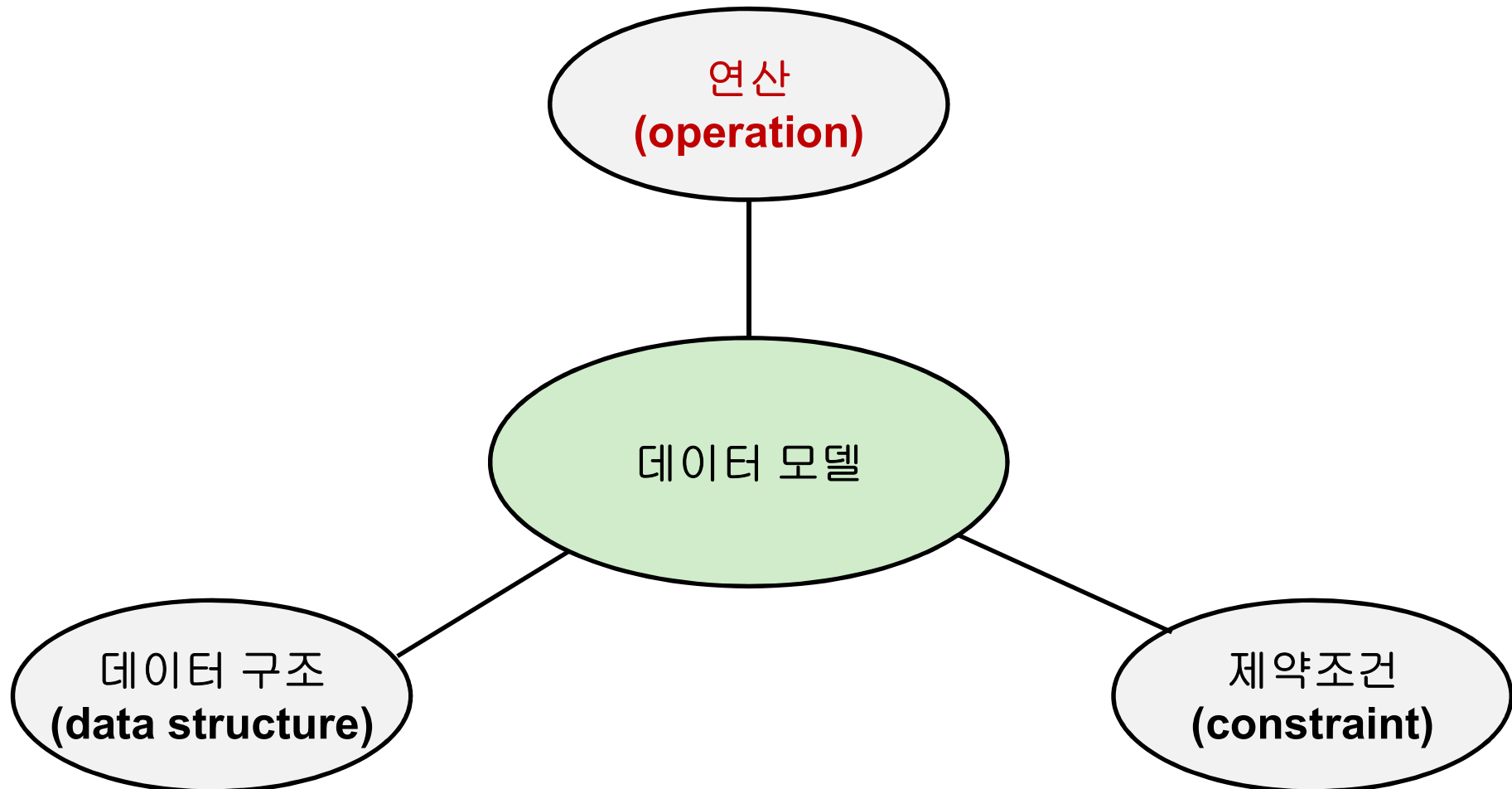
- SQL문을 이해하고 활용하려는 노력
- 테이블 정의서를 준수하려는 태도
- 생성된 테이블이 활용도를 다각도로 테스트하려는 적극적인 태도

개요

- 관계 데이터 연산
- 관계 대수
 - 일반 집합 연산자
 - 순수 관계 연산자
- 관계 해석

관계 데이터 연산의 개념

- 데이터 모델은 데이터 구조 + 연산 + 제약조건



관계 데이터 연산의 개념 (계속)

- 관계 데이터 모델의 연산
- 원하는 데이터를 얻기 위해 릴레이션에 필요한 처리 요구를 수행하는 것
- 관계 대수와 관계 해석: 기능과 표현력은 동등함
 - 관계 대수: 원하는 결과를 얻기 위해 데이터의 처리 과정을 순서대로 기술하는 **절차적 언어**
 - 관계 해석: 원하는 결과를 얻기 위해 처리를 원하는 데이터가 무엇인지만 기술하는 **비절차적 언어**

관계 데이터 연산의 개념 (계속)

- 관계 대수나 관계 해석은 개념적인 언어(DBMS에서 사용되진 않음)
- 관계 대수나 관계 해석은 데이터 언어의 유용성 검증의 기준으로 사용
 - 관계 대수나 관계 해석으로 기술할 수 있는 모든 질의(query)를 기술할 수 있는 데이터 언어를 **관계적으로 완전(relationally complete)**하다고 판단
 - 질의(query): 데이터에 대한 처리 요구

관계 대수

- 원하는 결과를 얻기 위해 릴레이션의 처리 과정을 순서대로 기술하는 언어
 - 절차 언어(procedural language)
- 릴레이션을 다루는 연산들의 모임
 - 수학적 집합 연산
 - 합집합, 교집합, 차집합, 카티션 프로덕트
 - 관계 데이터베이스를 위한 특별 연산
 - 선택(select), 프로젝트(project), 조인(join)
- 릴레이션에 대한 연산의 결과도 릴레이션
 - 중첩(nested)된 수식의 표현 가능
 - a.k.a., 폐쇄 특성(closure property)

관계 대수 (계속)

- 일반 집합 연산자 (set operation)
 - 릴레이션이 튜플의 집합이라는 개념을 이용하는 연산자
 - 피연산자(i.e., 릴레이션)가 두 개 필요
 - i.e., 합집합(\cup), 교집합(\cap), 차집합($-$), 카티션 프로덕트(\times)
- 합집합, 교집합, 차집합의 합병 가능(union-compatible) 조건
 - 두 릴레이션의 차수가 같아야 함
 - 두 릴레이션에서 서로 대응되는 속성의 도메인이 같아야 함

| <u>CID</u> | <u>NAME</u> | <u>AGE</u> |
|------------|-------------|------------|
| INT | CHAR(20) | INT |
| 100 | Kim | 32 |
| 200 | Lee | 7 |
| 300 | Park | 40 |
| 400 | Kang | 20 |
| 500 | Yeom | 20 |
| 600 | Hwang | 10 |
| 700 | Chang | 10 |

| <u>EID</u> | <u>NAME</u> | <u>AGE</u> |
|------------|-------------|------------|
| INT | CHAR(20) | INT |
| 100 | Kim | 32 |
| 200 | Lee | 37 |
| 300 | Park | 44 |
| 400 | Kang | 20 |
| 500 | Yeom | 63 |
| 600 | Hwang | 38 |
| 700 | Chang | 20 |

합병 가능한 릴레이션들.

일반 집합 연산자

- **합집합 (union, \cup)**

- $R \cup S = \{ t \mid t \in R \vee t \in S \}$
- $|R \cup S| \leq |R| + |S|$

- **교집합 (intersect, \cap)**

- $R \cap S = \{ t \mid t \in R \wedge t \in S \}$
- $|R \cap S| \leq \min\{ |R|, |S| \}$

- **차집합 (difference, $-$)**

- $R - S = \{ t \mid t \in R \wedge t \notin S \}$
- $|R - S| \leq |R|$

- **카티션 프로덕트 (cartesian product, \times)**

- $R \times S = \{ r \cdot s \mid r \in R \wedge s \in S \}$
- $|R \times S| = |R| \times |S|$ 차수(degree) = R의 차수 + S의 차수
- \cdot : 접속(concatenation)

일반집합연산자: 합집합(union)

- **합병 가능한 두 릴레이션 R과 S의 합집합: $R \cup S$**
 - 릴레이션 R에 속하거나 릴레이션 S에 속하는 모든 튜플로 결과 릴레이션 구성
 - i.e., $R \cup S = \{ t \mid t \in R \vee t \in S \}$
- **결과 릴레이션 특성**
 - 차수는 R과 S의 차수와 동일
 - 카디널리티는 릴레이션 R과 S의 카디널리티를 더한 것과 같거나 작음
 - $|R \cup S| \leq |R| + |S|$
- **교환적 특징 및 결합적 특징**
 - $R \cup S = S \cup R$
 - $(R \cup S) \cup T = R \cup (S \cup T)$

일반집합연산자: 교집합(intersection)

- **합병 가능한 두 릴레이션 R과 S의 합집합: $R \cup S$**
 - 릴레이션 R에 속하거나 릴레이션 S에 **둘 다 속하는** 모든 튜플로 결과 릴레이션 구성
 - i.e., $R \cup S = \{ t \mid t \in R \vee t \in S \}$
- **결과 릴레이션 특성**
 - 차수는 R과 S의 차수와 동일
 - 카디널리티는 릴레이션 R과 S의 어떤 카디널리티보다 크지 않음
 - $|R \cup S| \leq \max\{ |R|, |S| \}$
- **교환적 특징 및 결합적 특징**
 - $R \cap S = S \cap R$
 - $(R \cap S) \cap T = R \cap (S \cap T)$

일반집합연산자: 차집합(difference)

- **합병 가능한 두 릴레이션 R과 S의 합집합: R-S**
 - 릴레이션 R에 존재하고 릴레이션 S에 존재하지 않는 튜플로 결과 릴레이션 구성
 - i.e., $R \cap S = \{ t \mid t \in R \wedge t \in S \}$
- **결과 릴레이션 특성**
 - 차수는 R과 S의 차수와 동일
 - R-S의 카디널리티는 |R|과 같거나 작음
 - S-R의 카디널리티는 |S|과 같거나 작음
- **교환적 및 결합적 특징 없음**

일반집합연산자: 카디션프로덕트 (cartesian product)

- 두 릴레이션 R과 S의 카디션프로덕트: $R \times S$
 - 릴레이션 R에 속한 각 튜플과 릴레이션 S에 속한 각 튜플을 모두 연결
- 결과 릴레이션 특성
 - 차수는 릴레이션 R과 S의 차수의 합과 동일
 - 카디널리티는 릴레이션 R과 S의 카디널리티를 곱한 것과 동일
- 교환적 특징 및 결합적 특징
 - $R \times S = S \times R$
 - $R \times (S \times T) = (R \times S) \times T$

합집합, 교집합, 차집합 예제

(a)

| STUDENT | FN | LN |
|---------|---------|---------|
| | Susan | Yao |
| | Ramesh | Shah |
| | Johnny | Kohler |
| | Barbara | Jones |
| | Amy | Ford |
| | Jimmy | Wang |
| | Ernest | Gilbert |

| INSTRUCTOR | FNAME | LNAME |
|------------|---------|---------|
| | John | Smith |
| | Ricardo | Browne |
| | Susan | Yao |
| | Francis | Johnson |
| | Ramesh | Shah |

STUDENT \cup INSTRUCTOR

(b)

| FN | LN |
|---------|---------|
| Susan | Yao |
| Ramesh | Shah |
| Johnny | Kohler |
| Barbara | Jones |
| Amy | Ford |
| Jimmy | Wang |
| Ernest | Gilbert |
| John | Smith |
| Ricardo | Browne |
| Francis | Johnson |

STUDENT \cap INSTRUCTOR

(c)

| FN | LN |
|--------|------|
| Susan | Yao |
| Ramesh | Shah |

STUDENT - INSTRUCTOR

(d)

| FN | LN |
|---------|---------|
| Johnny | Kohler |
| Barbara | Jones |
| Amy | Ford |
| Jimmy | Wang |
| Ernest | Gilbert |

INSTRUCTOR - STUDENT

(e)

| FNAME | LNAME |
|---------|---------|
| John | Smith |
| Ricardo | Browne |
| Francis | Johnson |

Illustrating the set operations UNION, INTERSECTION, and DIFFERENCE.
 (a) Two union compatible relations. (b) STUDENT \cup INSTRUCTOR. (c) STUDENT \cap INSTRUCTOR. (d) STUDENT $-$ INSTRUCTOR. (e) INSTRUCTOR $-$ STUDENT.

카디션프로덕트 예제

| FEMALE_EMPS | FNAME | MINIT | LNAME | SSN | BDATE | ADDRESS | SEX | SALARY | SUPERSSN | DNO |
|-------------|----------|-------|---------|-----------|------------|-------------------------|-----|--------|-----------|-----|
| | Alicia | J | Zelaya | 999887777 | 1968-07-19 | 3321 Castle, Spring, TX | F | 25000 | 987654321 | 4 |
| | Jennifer | S | Wallace | 987654321 | 1941-06-20 | 291 Berry, Bellaire, TX | F | 43000 | 888665555 | 4 |
| | Joyce | A | English | 453453453 | 1972-07-31 | 5631 Rice, Houston, TX | F | 25000 | 333445555 | 5 |

| EMPNAMES | FNAME | LNAME | SSN |
|----------|----------|---------|-----------|
| | Alicia | Zelaya | 999887777 |
| | Jennifer | Wallace | 987654321 |
| | Joyce | English | 453453453 |

| EMP_DEPENDENTS | FNAME | LNAME | SSN | ESSN | DEPENDENT_NAME | SEX | BDATE | • • • |
|----------------|----------|---------|-----------|-----------|----------------|-----|------------|-------|
| | Alicia | Zelaya | 999887777 | 333445555 | Alice | F | 1986-04-05 | • • • |
| | Alicia | Zelaya | 999887777 | 333445555 | Theodore | M | 1983-10-25 | • • • |
| | Alicia | Zelaya | 999887777 | 333445555 | Joy | F | 1958-05-03 | • • • |
| | Alicia | Zelaya | 999887777 | 987654321 | Abner | M | 1942-02-28 | • • • |
| | Alicia | Zelaya | 999887777 | 123456789 | Michael | M | 1988-01-04 | • • • |
| | Alicia | Zelaya | 999887777 | 123456789 | Alice | F | 1988-12-30 | • • • |
| | Alicia | Zelaya | 999887777 | 123456789 | Elizabeth | F | 1967-05-05 | • • • |
| | Jennifer | Wallace | 987654321 | 333445555 | Alice | F | 1986-04-05 | • • • |
| | Jennifer | Wallace | 987654321 | 333445555 | Theodore | M | 1983-10-25 | • • • |
| | Jennifer | Wallace | 987654321 | 333445555 | Joy | F | 1958-05-03 | • • • |
| | Jennifer | Wallace | 987654321 | 987654321 | Abner | M | 1942-02-28 | • • • |
| | Jennifer | Wallace | 987654321 | 123456789 | Michael | M | 1988-01-04 | • • • |
| | Jennifer | Wallace | 987654321 | 123456789 | Alice | F | 1988-12-30 | • • • |
| | Jennifer | Wallace | 987654321 | 123456789 | Elizabeth | F | 1967-05-05 | • • • |
| | Joyce | English | 453453453 | 333445555 | Alice | F | 1986-04-05 | • • • |
| | Joyce | English | 453453453 | 333445555 | Theodore | M | 1983-10-25 | • • • |
| | Joyce | English | 453453453 | 333445555 | Joy | F | 1958-05-03 | • • • |
| | Joyce | English | 453453453 | 987654321 | Abner | M | 1942-02-28 | • • • |
| | Joyce | English | 453453453 | 123456789 | Michael | M | 1988-01-04 | • • • |
| | Joyce | English | 453453453 | 123456789 | Alice | F | 1988-12-30 | • • • |
| | Joyce | English | 453453453 | 123456789 | Elizabeth | F | 1967-05-05 | • • • |

$$3 \times 7 = 21$$

| ACTUAL_DEPENDENTS | FNAME | LNAME | SSN | ESSN | DEPENDENT_NAME | SEX | BDATE | • • • |
|-------------------|----------|---------|-----------|-----------|----------------|-----|------------|-------|
| | Jennifer | Wallace | 987654321 | 987654321 | Abner | M | 1942-02-28 | • • • |

| RESULT | FNAME | LNAME | DEPENDENT_NAME |
|--------|----------|---------|----------------|
| | Jennifer | Wallace | Abner |

An illustration of the CARTESIAN PRODUCT operation.

순수 관계 연산자

- 릴레이션의 구조와 특성을 이용하는 연산자
 - 선택 (select)
 - 프로젝트 (project)
 - 조인 (join)
 - 디비전 (division)

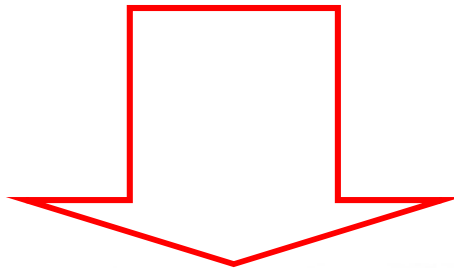
순수 관계 연산자: 선택(select)

- 릴레이션에서 조건을 만족하는 튜플들만을 선택하여 결과 릴레이션 구성
- 하나의 릴레이션(i.e., R)을 대상으로 연산을 수행
 - 연산 형식: $\sigma_{\langle \text{조건식} \rangle}(R)$
 - 데이터 언어적 표현법: R where <조건식>
 - 조건식: R의 속성들에 대한 임의의 부울식
 - a.k.a., 비교식, 프레디캣(predicate)
 - <속성 이름> <비교연산자> <상수값>
 - <속성 이름> <비교연산자> <속성 이름>
 - 비교연산자: =, <, ≤, >, ≥, ≠, SUBSTRING_OF (문자열 연산시 이용)

순수 관계 연산자: 셀렉트(계속)

| EMPLOYEE | FNAME | MINIT | LNAME | SSN | BDATE | ADDRESS | SEX | SALARY | SUPERSSN | DNO |
|----------|----------|-------|---------|-----------|------------|--------------------------|-----|--------|-----------|-----|
| | John | B | Smith | 123456789 | 1965-01-09 | 731 Fondren, Houston, TX | M | 30000 | 333445555 | 5 |
| | Franklin | T | Wong | 333445555 | 1955-12-08 | 638 Voss, Houston, TX | M | 40000 | 888665555 | 5 |
| | Alicia | J | Zelaya | 999887777 | 1968-07-19 | 3321 Castle, Spring, TX | F | 25000 | 987654321 | 4 |
| | Jennifer | S | Wallace | 987654321 | 1941-06-20 | 291 Berry, Bellaire, TX | F | 43000 | 888665555 | 4 |
| | Ramesh | K | Narayan | 666884444 | 1962-09-15 | 975 Fire Oak, Humble, TX | M | 38000 | 333445555 | 5 |
| | Joyce | A | English | 453453453 | 1972-07-31 | 5631 Rice, Houston, TX | F | 25000 | 333445555 | 5 |
| | Ahmad | V | Jabbar | 987987987 | 1969-03-29 | 980 Dallas, Houston, TX | M | 25000 | 987654321 | 4 |
| | James | E | Borg | 888665555 | 1937-11-10 | 450 Stone, Houston, TX | M | 55000 | null | 1 |

$\sigma_{(DNO=4 \text{ AND } SALARY>25000) \text{ OR } (DNO=5 \text{ AND } SALARY>30000)} (EMPLOYEE)$



(a)

| FNAME | MINIT | LNAME | SSN | BDATE | ADDRESS | SEX | SALARY | SUPERSSN | DNO |
|----------|-------|---------|-----------|------------|-------------------------|-----|--------|-----------|-----|
| Franklin | T | Wong | 333445555 | 1955-12-08 | 638 Voss, Houston, TX | M | 40000 | 888665555 | 5 |
| Jennifer | S | Wallace | 987654321 | 1941-06-20 | 291 Berry, Bellaire, TX | F | 43000 | 888665555 | 4 |
| Ramesh | K | Narayan | 666884444 | 1962-09-15 | 975 FireOak, Humble, TX | M | 38000 | 333445555 | 5 |

순수 관계 연산자: 셀렉트(계속)

- 논리연산자 이용

- AND, OR, NOT 연산자 이용 가능

- 교환법칙 성립

- $\sigma_{\langle \text{조건1} \rangle}(\sigma_{\langle \text{조건2} \rangle}(R)) = \sigma_{\langle \text{조건2} \rangle}(\sigma_{\langle \text{조건1} \rangle}(R))$

- 결합법칙 성립

- $\sigma_{\langle \text{조건1} \rangle}(\sigma_{\langle \text{조건2} \rangle}(\dots(\sigma_{\langle \text{조건n} \rangle}(R))\dots)) = \sigma_{\langle \text{조건1} \rangle \text{AND} \langle \text{조건2} \rangle \dots \text{AND} \langle \text{조건n} \rangle}(R)$

순수 관계 연산자: 프로젝트(project)

- 릴레이션에서 선택한 속성의 값으로 결과 릴레이션을 구성
- 하나의 릴레이션(i.e., R)을 대상으로 연산을 수행
 - 연산 형식: $\Pi_{\langle \text{속성 목록} \rangle}(R)$
 - 데이터 언어적 표현법: $R[\text{속성 목록}]$
- 결과 릴레이션의 튜플 수는 원래 릴레이션의 튜플 수보다 작거나 같음
 - 릴레이션은 중복 튜플을 인정하지 않음

순수 관계 연산자: 프로젝트(계속)

| EMPLOYEE | FNAME | MINIT | LNAME | SSN | BDATE | ADDRESS | SEX | SALARY | SUPERSSN | DNO |
|----------|----------|-------|---------|-----------|------------|--------------------------|-----|--------|-----------|-----|
| | John | B | Smith | 123456789 | 1965-01-09 | 731 Fondren, Houston, TX | M | 30000 | 333445555 | 5 |
| | Franklin | T | Wong | 333445555 | 1955-12-08 | 638 Voss, Houston, TX | M | 40000 | 888665555 | 5 |
| | Alicia | J | Zelaya | 999887777 | 1968-07-19 | 3321 Castle, Spring, TX | F | 25000 | 987654321 | 4 |
| | Jennifer | S | Wallace | 987654321 | 1941-06-20 | 291 Berry, Bellaire, TX | F | 43000 | 888665555 | 4 |
| | Ramesh | K | Narayan | 666884444 | 1962-09-15 | 975 Fire Oak, Humble, TX | M | 38000 | 333445555 | 5 |
| | Joyce | A | English | 453453453 | 1972-07-31 | 5631 Rice, Houston, TX | F | 25000 | 333445555 | 5 |
| | Ahmad | V | Jabbar | 987987987 | 1969-03-29 | 980 Dallas, Houston, TX | M | 25000 | 987654321 | 4 |
| | James | E | Borg | 888665555 | 1937-11-10 | 450 Stone, Houston, TX | M | 55000 | null | 1 |

$\Pi_{FNAME, LNAME, SALARY}(EMPLOYEE)$



(b)

| LNAME | FNAME | SALARY |
|---------|----------|--------|
| Smith | John | 30000 |
| Wong | Franklin | 40000 |
| Zelaya | Alicia | 25000 |
| Wallace | Jennifer | 43000 |
| Narayan | Ramesh | 38000 |
| English | Joyce | 25000 |
| Jabbar | Ahmad | 25000 |
| Borg | James | 55000 |

$\Pi_{SEX, SALARY}(EMPLOYEE)$



(c)

| SEX | SALARY |
|-----|--------|
| M | 30000 |
| M | 40000 |
| F | 25000 |
| F | 43000 |
| M | 38000 |
| M | 25000 |
| M | 55000 |

순수 관계 연산자: 조인(join)

- 조인 속성을 이용해 두 릴레이션(i.e., R과 S)을 조합하여 결과 릴레이션을 구성
 - R과 S로부터 값이 같은 튜플만을 결합하여 생성된 튜플들을 결과 릴레이션에 포함
 - 조인 속성: 두 릴레이션이 공통으로 가지고 있는 속성
 - 표현법: $R \bowtie S$
 - a.k.a., 자연조인(natural join)

순수 관계 연산자: 조인(계속)

■ 세타조인(theta join, Θ -join)

- 자연 조인에 비하여 더 일반화된 조인
- 주어진 **조인 조건을 만족하는 두 릴레이션의 모든 튜플을 연결**하여 생성된 새로운 튜플로 결과 릴레이션을 구성
- 결과 릴레이션의 차수는 두 릴레이션의 차수를 더한 것과 같음
- 표현법: $R \bowtie \langle \text{조인조건} \rangle S$ (두 릴레이션들 R과 S에 대해)
 - $\langle \text{조건} \rangle$ AND $\langle \text{조건} \rangle$ AND ... AND $\langle \text{조건} \rangle$
 - 각 조건의 형태는 $A_i \Theta B_j$ 이며, A_i 는 R의 속성, B_j 는 S의 속성
 - $\Theta = \{=, <, \leq, >, \geq, \neq\}$

■ 동일 조인(equi-join)

- Θ 연산자가 "="인 세타조인

순수 관계 연산자: 조인(계속)

| DEPARTMENT | DNAME | DNUMBER | MGRSSN | MGRSTARTDATE |
|----------------|-------|---------|-----------|--------------|
| Research | | 5 | 333445555 | 1988-05-22 |
| Administration | | 4 | 987654321 | 1995-01-01 |
| Headquarters | | 1 | 888665555 | 1981-06-19 |

| EMPLOYEE | FNAME | MINIT | LNAME | SSN | BDATE | ADDRESS | SEX | SALARY | SUPERSSN | DNO |
|----------|----------|-------|---------|-----------|------------|--------------------------|-----|--------|-----------|-----|
| | John | B | Smith | 123456789 | 1965-01-09 | 731 Fondren, Houston, TX | M | 30000 | 333445555 | 5 |
| | Franklin | T | Wong | 333445555 | 1955-12-08 | 638 Voss, Houston, TX | M | 40000 | 888665555 | 5 |
| | Alicia | J | Zelaya | 999887777 | 1968-07-19 | 3321 Castle, Spring, TX | F | 25000 | 987654321 | 4 |
| | Jennifer | S | Wallace | 987654321 | 1941-06-20 | 291 Berry, Bellaire, TX | F | 43000 | 888665555 | 4 |
| | Ramesh | K | Narayan | 666884444 | 1962-09-15 | 975 Fire Oak, Humble, TX | M | 38000 | 333445555 | 5 |
| | Joyce | A | English | 453453453 | 1972-07-31 | 5631 Rice, Houston, TX | F | 25000 | 333445555 | 5 |
| | Ahmad | V | Jabbar | 987987987 | 1969-03-29 | 980 Dallas, Houston, TX | M | 25000 | 987654321 | 4 |
| | James | E | Borg | 888665555 | 1937-11-10 | 450 Stone, Houston, TX | M | 55000 | null | 1 |

DEPT_MGR ← DEPARTMENT ⋈_{MGRSSN=SSN} EMPLOYEE



| DEPT_MGR | DNAME | DNUMBER | MGRSSN | ... | FNAME | MINIT | LNAME | SSN | ... |
|----------|----------------|---------|-----------|-----|----------|-------|---------|-----------|-----|
| | Research | 5 | 333445555 | ... | Franklin | T | Wong | 333445555 | ... |
| | Administration | 4 | 987654321 | ... | Jennifer | S | Wallace | 987654321 | ... |
| | Headquarters | 1 | 888665555 | ... | James | E | Borg | 888665555 | ... |

Figure 7.13 Illustrating the JOIN operation.

순수 관계 연산자: 조인(계속)

■ 세미 조인(semi-join)

- 조인 속성으로 **프로젝트 연산을 수행한 릴레이션을 이용**하는 조인
- 표현법: $R \bowtie S$ (두 릴레이션들 R과 S에 대해)
 - S를 조인 속성으로 프로젝트 연산 후에 R에 자연조인하여 결과 릴레이션을 구성
- 불필요한 속성을 미리 제거하여 연산 비용을 줄이는 효과
- 교환적 특징이 없음
 - $R \bowtie S \neq S \bowtie R$

순수 관계 연산자: 조인(계속)

릴레이션 R

(b)

| LNAME | FNAME | SALARY |
|---------|----------|--------|
| Smith | John | 30000 |
| Wong | Franklin | 40000 |
| Zelaya | Alicia | 25000 |
| Wallace | Jennifer | 43000 |
| Narayan | Ramesh | 38000 |
| English | Joyce | 25000 |
| Jabbar | Ahmad | 25000 |
| Borg | James | 55000 |

릴레이션 S

(c)

| SEX | SALARY |
|-----|--------|
| M | 30000 |
| M | 40000 |
| F | 25000 |
| F | 43000 |
| M | 38000 |
| M | 25000 |
| M | 55000 |

$\Pi_{\text{SALARY}}(S)$

| SALARY |
|--------|
| 30000 |
| 40000 |
| 25000 |
| 43000 |
| 38000 |
| 25000 |
| 55000 |

세미조인연산

$R \bowtie S$

| LNAME | FNAME | SALARY |
|---------|----------|--------|
| Smith | John | 30000 |
| Wong | Franklin | 40000 |
| Zelaya | Alicia | 25000 |
| Wallace | Jennifer | 43000 |
| Narayan | Ramesh | 38000 |
| English | Joyce | 25000 |
| Jabbar | Ahmad | 25000 |
| Borg | James | 55000 |

순수 관계 연산자: 조인(계속)

■ 외부 조인(outer-join)

- 자연 조인 연산에서 제외되는 모든 튜플을 결과 릴레이션에 포함시키는 조인
 - 동일 조인이나 자연 조인 연산에서 조인 조건을 만족하지 않는 튜플들은 결과 릴레이션에도 나타나지 않음
 - 조인에 참여하는 릴레이션의 모든 튜플들이 조인 여부와 관계 없이 결과 릴레이션에서 나타나고 싶은 경우 사용됨
 - 외부 조인에서 상대방 릴레이션에 대응되는 튜플이 없으면 빈 속성에 NULL 값 할당
- 표현법: $R \bowtie S$ (두 릴레이션들 R과 S에 대해)

순수 관계 연산자: 조인(계속)

■ 외부 조인 종류

- 왼쪽 외부 조인 (left outer join)
 - $R \bowtie L S$ 는 R의 모든 튜플들이 결과 릴레이션에 존재
- 오른쪽 외부 조인 (right outer join)
 - $R \bowtie R S$ 는 S의 모든 튜플들이 결과 릴레이션에 존재
- 완전 외부 조인(full outer join)
 - $R \bowtie R S$ 는 R과 S의 모든 튜플들이 결과 릴레이션에 존재

순수 관계 연산자: 조인(계속)

| S | A | B1 |
|---|---|----|
| | 1 | 가 |
| | 2 | 나 |
| | 3 | 다 |

| T | B2 | C |
|---|----|----|
| | 나 | # |
| | 다 | % |
| | 라 | \$ |
| | 마 | @ |

조인

$V \leftarrow S \bowtie_{B1=B2} T$

| V | A | B1 | B2 | C |
|---|---|----|----|---|
| | 2 | 나 | 나 | # |
| | 3 | 다 | 다 | % |

왼쪽 외부조인

$V \leftarrow S \ltimes_{B1=B2} T$

| V | A | B1 | B2 | C |
|---|---|----|------|------|
| | 1 | 가 | null | null |
| | 2 | 나 | 나 | # |
| | 3 | 다 | 다 | % |

오른쪽 외부조인

$V \leftarrow S \bowtie_{B1=B2} T$

| V | A | B1 | B2 | C |
|---|------|------|----|----|
| | 2 | 나 | 나 | # |
| | 3 | 다 | 다 | % |
| | null | null | 라 | \$ |
| | null | null | 마 | @ |

양쪽 외부조인

$V \leftarrow S \ltimes_{B1=B2} T$

| V | A | B1 | B2 | C |
|---|------|------|------|------|
| | 1 | 가 | null | null |
| | 2 | 나 | 나 | # |
| | 3 | 다 | 다 | % |
| | null | null | 라 | \$ |
| | null | null | 마 | @ |

순수 관계 연산자: 조인(계속)

| DEPARTMENT | DNAME | DNUMBER | MGRSSN | MGRSTARTDATE |
|------------|----------------|---------|-----------|--------------|
| | Research | 5 | 333445555 | 1988-05-22 |
| | Administration | 4 | 987654321 | 1995-01-01 |
| | Headquarters | 1 | 888665555 | 1981-06-19 |

| EMPLOYEE | FNAME | MINIT | LNAME | SSN | BDATE | ADDRESS | SEX | SALARY | SUPERSSN | DNO |
|----------|----------|-------|---------|-----------|------------|--------------------------|-----|--------|-----------|-----|
| | John | B | Smith | 123456789 | 1965-01-09 | 731 Fondren, Houston, TX | M | 30000 | 333445555 | 5 |
| | Franklin | T | Wong | 333445555 | 1955-12-08 | 638 Voss, Houston, TX | M | 40000 | 888665555 | 5 |
| | Alicia | J | Zelaya | 999887777 | 1968-07-19 | 3321 Castle, Spring, TX | F | 25000 | 987654321 | 4 |
| | Jennifer | S | Wallace | 987654321 | 1941-06-20 | 291 Berry, Bellaire, TX | F | 43000 | 888665555 | 4 |
| | Ramesh | K | Narayan | 666884444 | 1962-09-15 | 975 Fire Oak, Humble, TX | M | 38000 | 333445555 | 5 |
| | Joyce | A | English | 453453453 | 1972-07-31 | 5631 Rice, Houston, TX | F | 25000 | 333445555 | 5 |
| | Ahmad | V | Jabbar | 987987987 | 1969-03-29 | 980 Dallas, Houston, TX | M | 25000 | 987654321 | 4 |
| | James | E | Borg | 888665555 | 1937-11-10 | 450 Stone, Houston, TX | M | 55000 | null | 1 |

TEMP ← EMPLOYEE ⋈_{MGRSSN=SSN} DEPARTMENT
 RESULT ← $\Pi_{FNAME, MINIT, LNAME, DNAME}(TEMP)$

| RESULT | FNAME | MINIT | LNAME | DNAME |
|--------|----------|-------|---------|----------------|
| | John | B | Smith | null |
| | Franklin | T | Wong | Research |
| | Alicia | J | Zelaya | null |
| | Jennifer | S | Wallace | Administration |
| | Ramesh | K | Narayan | null |
| | Joyce | A | English | null |
| | Ahmad | V | Jabbar | null |
| | James | E | Borg | Headquarters |

순수 관계 연산자: 디비전(division)

■ 표현법: $T(Y) = R(Z) \div S(X)$ ($X \subseteq Z$ 이고, $Y = Z - X$)

- S의 모든 튜플과 관련이 있는 R의 튜플로 결과 릴레이션을 구성
 - $T1 = \Pi_Y(R)$
 - $T2 = \Pi_Y((S \times T1) - R)$
 - $T = T1 - T2$
- 이경우 **R이 S의 모든 속성을 포함하고 있어야 연산 가능**

| SSN_PNOS | ESSN | PNO |
|----------|-----------|-----|
| | 123456789 | 1 |
| | 123456789 | 2 |
| | 666884444 | 3 |
| | 453453453 | 1 |
| | 453453453 | 2 |
| | 333445555 | 2 |
| | 333445555 | 3 |
| | 333445555 | 10 |
| | 333445555 | 20 |
| | 999887777 | 30 |
| | 999887777 | 10 |
| | 987987987 | 10 |
| | 987987987 | 30 |
| | 987654321 | 30 |
| | 987654321 | 20 |
| | 888665555 | 20 |

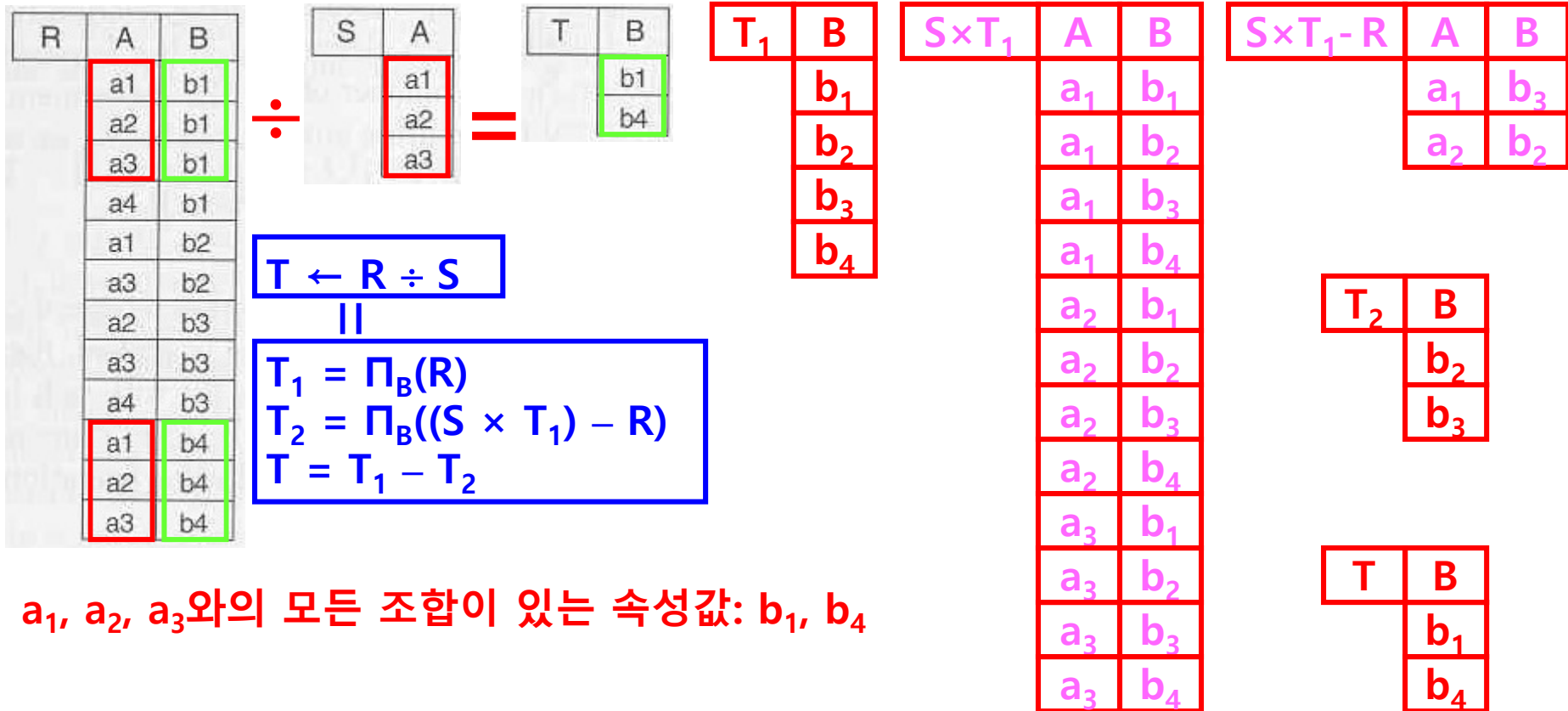
÷

| SMITH_PNOS | PNO |
|------------|-----|
| | 1 |
| | 2 |

=

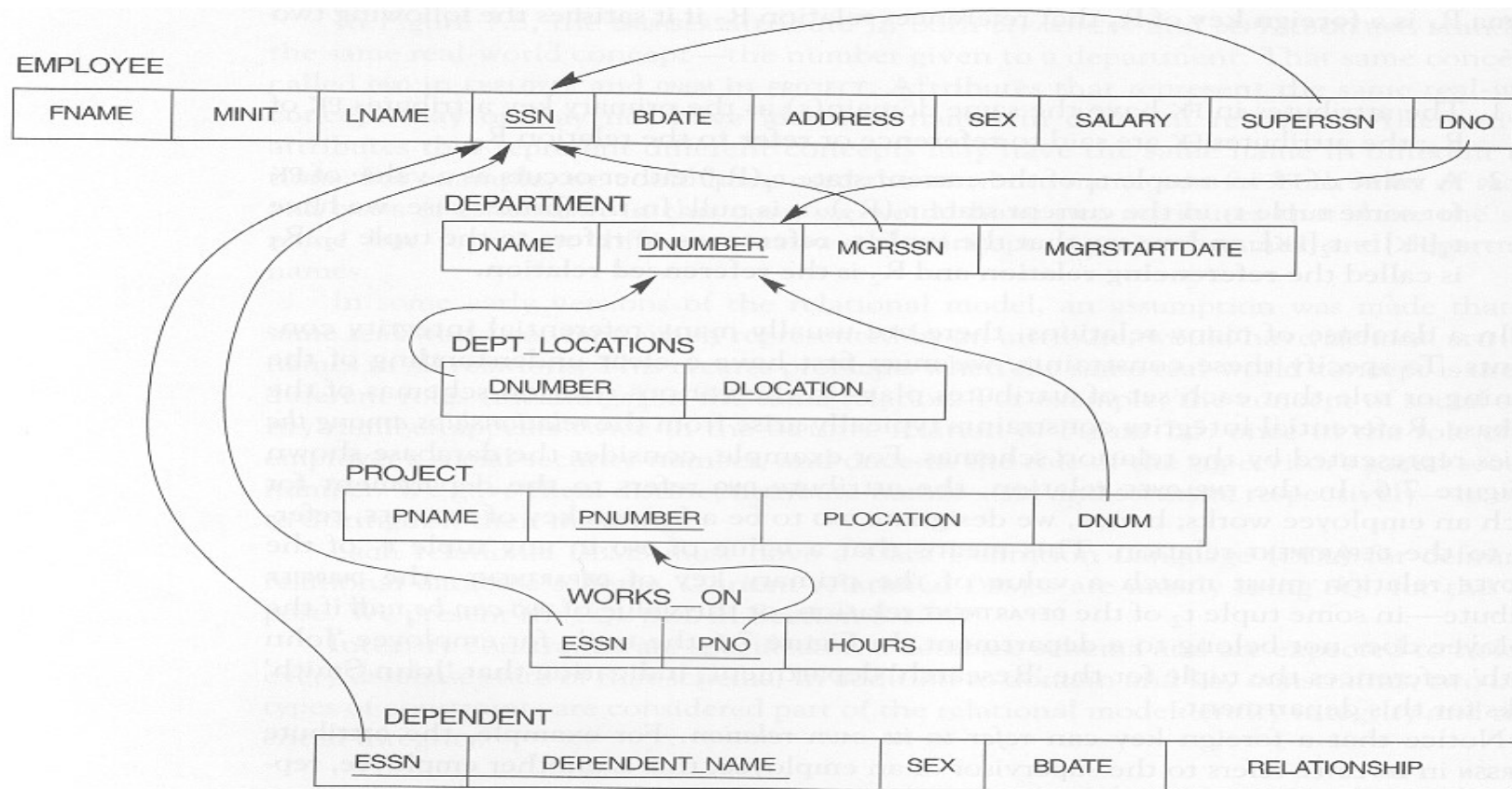
| SSNS | SSN |
|------|-----------|
| | 123456789 |
| | 453453453 |

순수 관계 연산자: 디비전(계속)



관계 대수를 이용한 질의 표현

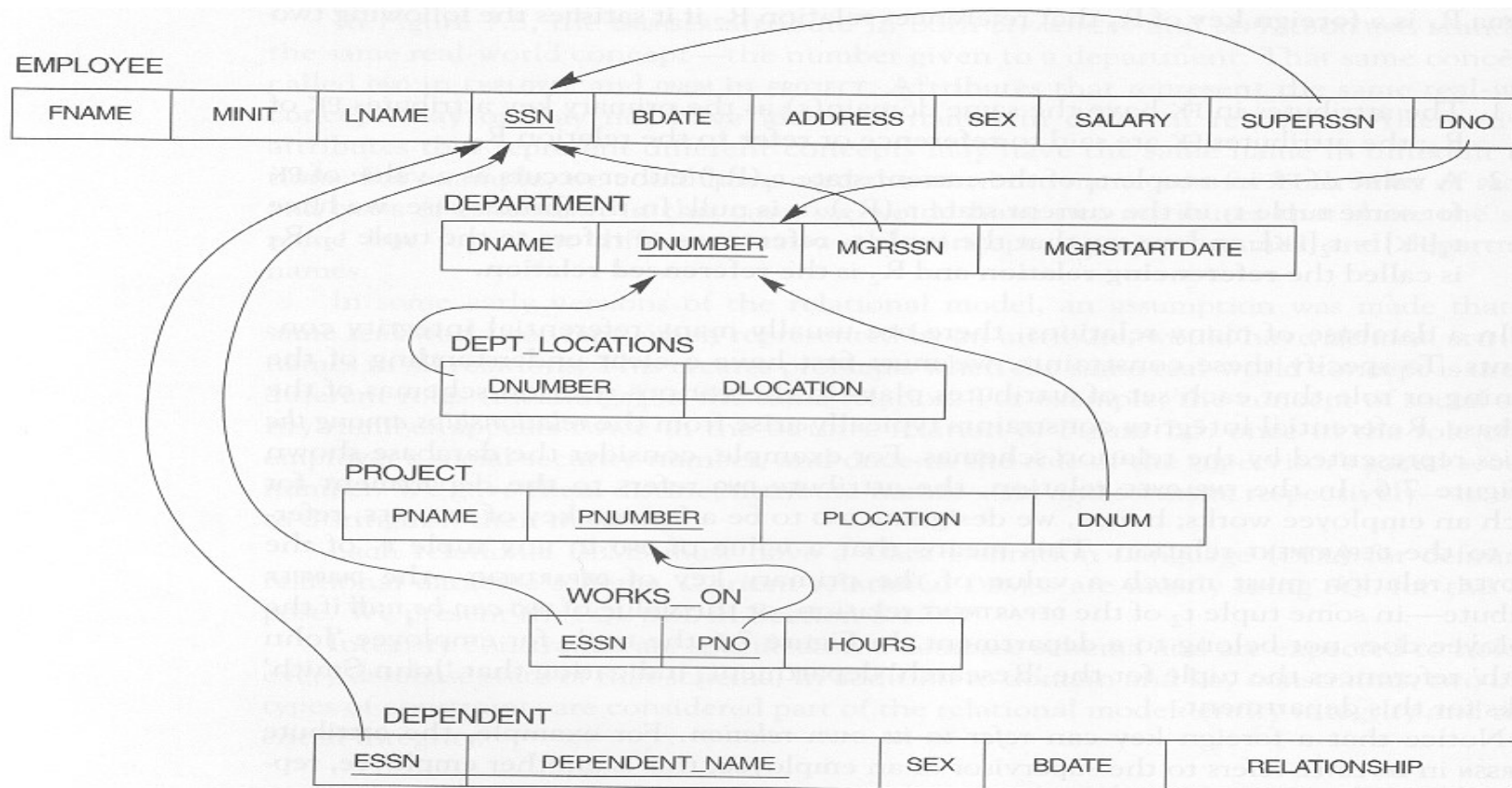
- Retrieve the name and address of all employees who work for the 'Research' department



Referential integrity constraints displayed on the COMPANY relational database schema diagram.

관계 대수를 이용한 질의 표현(계속)

- For every project located in 'Stafford', list the project number, the controlling department number, and the department manager's last name, address, and birthdate



Referential integrity constraints displayed on the COMPANY relational database schema diagram.

관계 해석

- 처리를 원하는 데이터가 무엇인지만 기술하는 언어
 - 비절차적 언어(nonprocedural language)
- 수학의 프레디킷 해석(predicate calculus)에 기반
- 분류
 - 튜플 관계 해석(tuple relational calculus)
 - 도메인 관계해석(domain relational calculus)

감사합니다!

담당교수: 전강욱(컴퓨터공학부)
kw.chon@koreatech.ac.kr