

Relational Algebra (1)

Week 6: Database Systems (4CCS1DBS)

20 Feb 2016

REVIEW: Summary of SQL Queries

- The **SELECT**-clause lists the attributes or functions to be retrieved
- The **FROM**-clause specifies all relations (or aliases) needed in the query but not those needed in nested queries
- The **WHERE**-clause specifies the conditions for selection and join of tuples from the relations specified in the FROM-clause
- **GROUP BY** specifies grouping attributes
- **HAVING** specifies a condition for selection of groups
- **ORDER BY** specifies an order for displaying the result of a query
- A query is evaluated by:
 1. applying the WHERE-clause
 2. then GROUP BY and HAVING
 3. and finally the SELECT-clause
 4. and then ORDER BY the resulting tuples

REVIEW: JOINS

EMPLOYEE JOIN DEPENDENT ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE JOIN DEPENDENT ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE **LEFT JOIN** DEPENDENT
ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE **LEFT JOIN** DEPENDENT
ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE **RIGHT JOIN** DEPENDENT
ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE **RIGHT JOIN** DEPENDENT
ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE **FULL OUTER JOIN** DEPENDENT
ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE **FULL OUTER JOIN** DEPENDENT
ON Ssn = Essn

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE NATURAL JOIN DEPENDENT

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: JOINS

EMPLOYEE NATURAL JOIN DEPENDENT

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

REVIEW: Constraints as Assertions

- General constraints: constraints that do not fit in the basic SQL categories
- Useful for ***Schema Assertions*** - Outside the scope of the built-in relational model constraints (primary and unique keys, entity integrity, referential integrity).
- Defines whether the State of the Database is VALID at any given point of time.
- **CREATE ASSERTION**, Components include:
 - A constraint name
 - Followed by a CHECK keyword
 - Followed by a condition clause
- Enforcing the Assertion is up to the Database Implementation – i.e. Rejecting a Query that will violate the CHECK ASSERTION.

REVIEW: CREATE ASSERTION Example

What is being asserted here?

assertion name

```
CREATE ASSERTION SALARY_CONSTRAINT
CHECK (NOT EXISTS (
    SELECT *
    FROM   EMPLOYEE E, EMPLOYEE M, DEPARTMENT D
    WHERE  E.SALARY > M.SALARY AND
           E.DNO=D.NUMBER AND D.MGRSSN=M.SSN) )
```

assertion condition

REVIEW: CREATE ASSERTION Example

Example: The salary of an employee must not be greater than the salary of the manager of the department that the employee works for

assertion name

```
CREATE ASSERTION SALARY_CONSTRAINT
CHECK (NOT EXISTS (
    SELECT *
    FROM   EMPLOYEE E, EMPLOYEE M, DEPARTMENT D
    WHERE  E.SALARY > M.SALARY AND
           E.DNO=D.NUMBER AND D.MGRSSN=M.SSN) )
```

assertion condition

REVIEW: Views in SQL

- A view is a “*virtual*” table that is derived from other tables
- Two ways they are implemented in implementation:
 - Query modification - *copy and paste queries*
 - View materialization - *short-term physical implementation*
- Limited for UPDATE operations
 - Table not physical stored
- Allows full query operations
- A convenience for expressing certain operations
- Useful for security and authorization
- Prevents redundant storage of data

SQL View Example

Example: A “friendlier” view of WORKS_ON

- SQL command: CREATE VIEW

view name

CREATE VIEW WORKS_ON1 AS

```
SELECT FNAME, LNAME, PNAME, HOURS
FROM   EMPLOYEE, PROJECT, WORKS_ON
WHERE  SSN=ESSN AND PNO=PNUMBER
```

query to specify the contents of view

WORKS_ON1

Fname	Lname	Pname	Hours
-------	-------	-------	-------

view name option: specify attribute names

WORKS_ON1 (FIRST_NAME, LAST_NAME, PROJECT, HOURS)

Using a Virtual Table (a View)

Example: A “friendlier” view of WORKS_ON

- We can specify SQL queries on a newly create table (view):

```
SELECT FNAME, LNAME  
FROM WORKS_ON1  
WHERE PNAME= 'ProductX' ;
```

- When no longer needed, a view can be dropped:

```
DROP WORKS_ON1 ;
```

- Dropping a View does NOT modify the data!

Relational Algebra Overview

- Relational algebra is the basic set of operations for the relational model
- These operations enable a user to specify **basic retrieval requests** (or **queries**)
- The result of an operation is a *new relation*, which may have been formed from one or more *input relations*
 - This property makes the algebra “closed” (all objects in relational algebra are relations)

Relational Algebra Overview (continued)

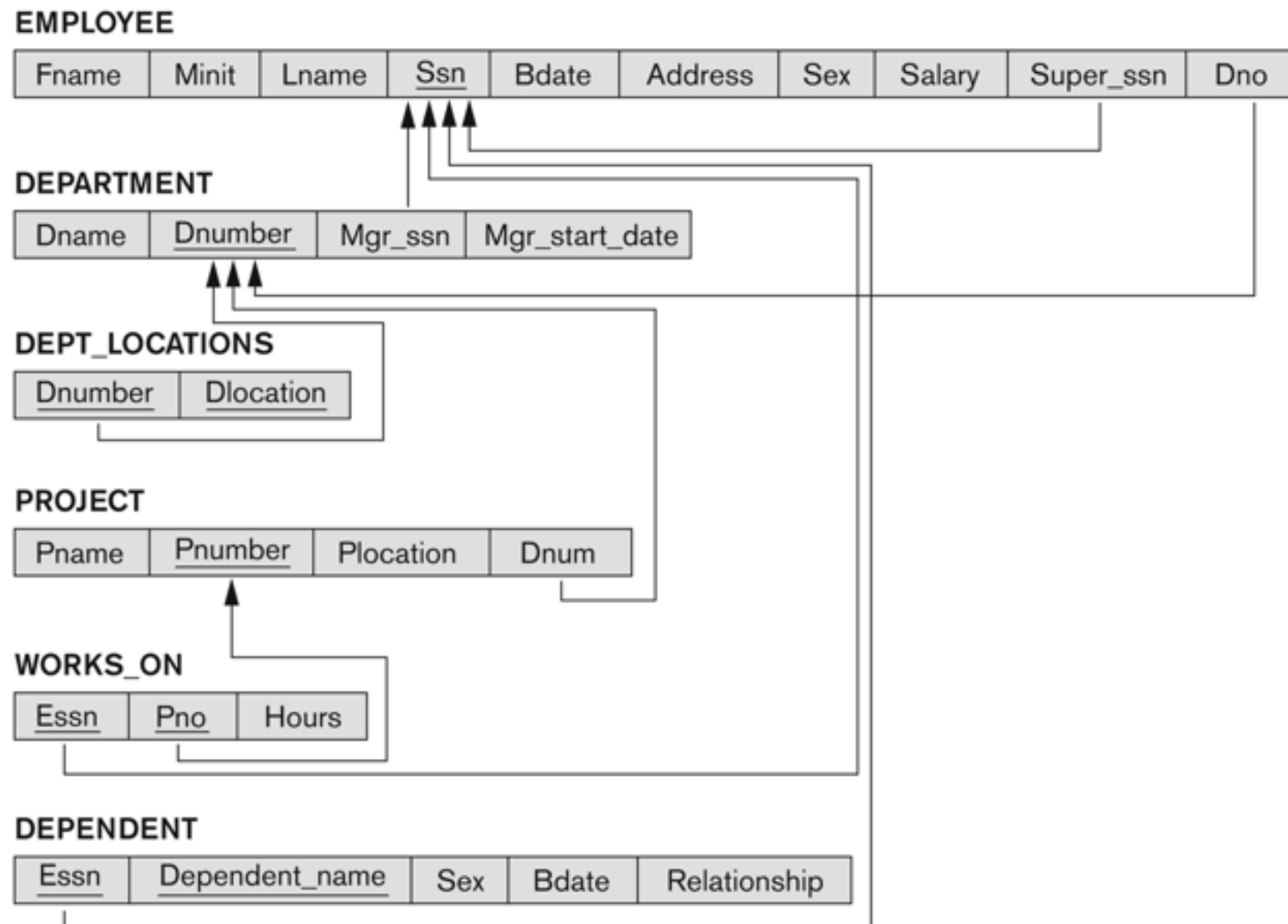
- The **algebra operations** thus produce new relations
 - These can be further manipulated using operations of the same algebra
- A sequence of relational algebra operations forms a **relational algebra expression**
 - The result of a relational algebra expression is also a relation that represents the result of a database query (or retrieval request)

Relational Algebra Overview

- Relational Algebra consists of several groups of operations
 - **Unary Relational Operations**
 - SELECT (symbol: σ (sigma))
 - PROJECT (symbol: π (pi))
 - RENAME (symbol: ρ (rho))
 - **Relational Algebra Operations From Set Theory**
 - UNION (\cup), INTERSECTION (\cap), DIFFERENCE (or MINUS, $-$)
 - CARTESIAN PRODUCT (\times)
 - **Binary Relational Operations**
 - JOIN (several variations of JOIN exist)
 - DIVISION
 - **Additional Relational Operations**
 - OUTER JOINS, OUTER UNION
 - AGGREGATE FUNCTIONS (These compute summary of information: for example, SUM, COUNT, AVG, MIN, MAX)

Database Schema for COMPANY

- Refer to the COMPANY database shown here.



EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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

DEPARTMENT

Dname	<u>Dnumber</u>	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

<u>Dnumber</u>	<u>Dlocation</u>
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

<u>Essn</u>	<u>Pno</u>	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	<u>Pnumber</u>	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

<u>Essn</u>	<u>Dependent_name</u>	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	M	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	M	1942-02-28	Spouse
123456789	Michael	M	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

Unary Relational Operations: SELECT

- The SELECT operation (denoted by σ (sigma)) is used to select a *subset* of the tuples from a relation based on a **selection condition**.
 - The selection condition acts as a **filter**
 - Keeps only those tuples that satisfy the qualifying condition
 - Tuples satisfying the condition are *selected* whereas the other tuples are discarded (*filtered out*)

Unary Relational Operations: SELECT

- Examples:

- Select the EMPLOYEE tuples whose department number is 4:

$$\sigma_{DNO = 4} (EMPLOYEE)$$

- Select the employee tuples whose salary is greater than \$30,000:

$$\sigma_{SALARY > 30,000} (EMPLOYEE)$$

- What are the resulting relations?

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	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

Unary Relational Operations: SELECT

- In general, the *select* operation is denoted by
$$\sigma_{\langle \text{selection condition} \rangle}(R)$$
 - the symbol σ (sigma) is used to denote the *select* operator
 - the selection condition is a Boolean (conditional) expression specified on the attributes of relation R
 - tuples that make the condition **true** are selected
 - appear in the result of the operation
 - tuples that make the condition **false** are filtered out
 - discarded from the result of the operation

Unary Relational Operations: SELECT (contd.)

- **SELECT Operation Properties**

- The SELECT operation $\sigma_{\langle \text{selection condition} \rangle}(R)$ produces a relation S that has the same schema (same attributes) as R
- SELECT σ is commutative:
 - $\sigma_{\langle \text{condition1} \rangle}(\sigma_{\langle \text{condition2} \rangle}(R)) = \sigma_{\langle \text{condition2} \rangle}(\sigma_{\langle \text{condition1} \rangle}(R))$
- Because of commutativity property, a cascade (sequence) of SELECT operations may be applied in any order:
 - $\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(R))) = \sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(\sigma_{\langle \text{cond1} \rangle}(R)))$

Unary Relational Operations: SELECT (contd.)

- **SELECT Operation Properties**
 - A cascade of SELECT operations may be replaced by a single selection with a conjunction of all the conditions:
 - $\sigma_{\langle \text{cond1} \rangle}(\sigma_{\langle \text{cond2} \rangle}(\sigma_{\langle \text{cond3} \rangle}(R))) = \sigma_{\langle \text{cond1} \rangle \text{ AND } \langle \text{cond2} \rangle \text{ AND } \langle \text{cond3} \rangle}(R)$
 - The number of tuples in the result of a SELECT is less than (or equal to) the number of tuples in the input relation R

Example of the SELECT Operation

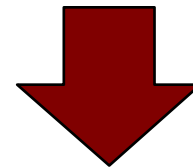
σ (DNO = 4 AND SALARY > 25,000) OR (DNO = 5 AND SALARY > 30,000) (EMPLOYEE)

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	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

Example of the SELECT Operation

σ (DNO = 4 AND SALARY > 25,000) OR (DNO = 5 AND SALARY > 30,000) (EMPLOYEE)

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	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



Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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 Fire Oak, Humble, TX	M	38000	333445555	5

Unary Relational Operations: PROJECT

- PROJECT Operation is denoted by π (pi)
- This operation keeps certain *columns* (attributes) from a relation and discards the other columns.
 - **PROJECT creates a vertical partitioning**
 - The list of specified columns (attributes) is kept in each tuple
 - The other attributes in each tuple are discarded

Unary Relational Operations: PROJECT

- Example: To list each employee's first and last name and salary, the following is used:

$\pi_{\text{LNAME, FNAME, SALARY}}(\text{EMPLOYEE})$

- What is the resulting relation?

EMPLOYEE	FNAME	MINIT	LNAME	<u>SSN</u>	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

Unary Relational Operations: PROJECT (cont.)

- The general form of the *project* operation is:
$$\pi_{\langle \text{attribute list} \rangle}(R)$$
 - π (pi) is the symbol used to represent the *project* operation
 - $\langle \text{attribute list} \rangle$ is the desired list of attributes from relation R.
- The project operation *removes any duplicate tuples*
 - This is because the result of the *project* operation must be a *set of tuples*
 - Mathematical sets *do not allow* duplicate elements.

Unary Relational Operations: PROJECT (contd.)

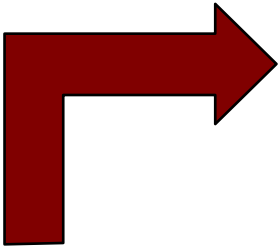
- PROJECT Operation Properties
 - The number of tuples in the result of projection $\pi_{\langle \text{list} \rangle}(R)$ is **always less or equal** to the number of tuples in R
 - If the list of attributes includes a *key* of R, then the number of tuples in the result of PROJECT is *equal* to the number of tuples in R

Examples of Applying SELECT and PROJECT Operations

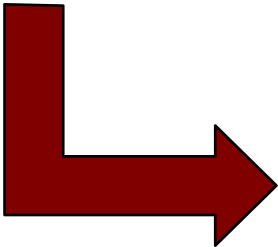
$\pi_{\text{Lname, Fname, Salary}}(\text{EMPLOYEE})$

EMPLOYEE

Fname	Minit	Lname	<u>Ssn</u>	Bdate	Address	Sex	Salary	Super_ssn	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	1968-01-19	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



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



$\pi_{\text{Sex, Salary}}(\text{EMPLOYEE})$

Sex	Salary
M	30000
M	40000
F	25000
F	43000
M	38000
M	25000
M	55000

Relational Algebra Expressions

- We may want to apply several relational algebra operations one after the other
 - Either we can write the operations as a single **relational algebra expression** by nesting the operations, or
 - We can apply one operation at a time and create **intermediate result relations**.
- In the latter case, we must give names to the relations that hold the intermediate results.

Single Expression vs. Sequence of Relational Operations: Example

- Retrieve the first name, last name, and salary of all employees who work in department number 5.
- Apply a SELECT and a PROJECT operation
- We can write a *single relational algebra expression* as follows:
 - $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$

Single Expression vs. Sequence of Relational Operations: Example

- Retrieve the first name, last name, and salary of all employees who work in department number 5.
- Apply a SELECT and a PROJECT operation
- We can write a *single relational algebra expression* as follows:
 - $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$
 - ```
SELECT FNAME, LNAME, SALARY
FROM EMPLOYEE
WHERE DNO=5;
```

- Retrieve the first name, last name, and salary of all employees who work in department number 5:
  - $\pi_{\text{FNAME, LNAME, SALARY}}(\sigma_{\text{DNO}=5}(\text{EMPLOYEE}))$
- OR:
- We can explicitly show the *sequence of operations*, giving a name to each intermediate relation:
  - $\text{DEP5\_EMPS} \leftarrow \sigma_{\text{DNO}=5}(\text{EMPLOYEE})$
  - $\text{RESULT} \leftarrow \pi_{\text{FNAME, LNAME, SALARY}}(\text{DEP5\_EMPS})$


# Unary Relational Operations: RENAME

- The RENAME operator is denoted by  $\rho$  (rho)
- In some cases, we may want to *rename* the attributes of a relation or the relation name or both
  - Useful when a query requires multiple operations
  - Necessary in some cases (see JOIN operation later)

# Unary Relational Operations: RENAME (contd.)

- The general RENAME operation  $\rho$  can be expressed by any of the following forms:
  - $\rho_S(R)$  changes:
    - the *relation name* to  $S$
  - $\rho_{(B1, B2, \dots, Bn)}(R)$  changes:
    - the *column (attribute) names* to  $B1, B2, \dots, Bn$
  - $\rho_{S (B1, B2, \dots, Bn)}(R)$  changes both:
    - the relation name to  $S$ , *and*
    - the column (attribute) names to  $B1, B2, \dots, Bn$

# Example of Applying Multiple Operations and RENAME

$\pi_{\text{FNAME, LNAME, SALARY}}(\text{EMPLOYEE})$  

| Fname    | Lname   | Salary |
|----------|---------|--------|
| John     | Smith   | 30000  |
| Franklin | Wong    | 40000  |
| Ramesh   | Narayan | 38000  |
| Joyce    | English | 25000  |

**EMPLOYEE**

| Fname    | Minit | Lname   | <u>Ssn</u> | Bdate      | Address                  | Sex | Salary | Super_ssn | 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  | 1968-01-19 | 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   |

# Example of applying multiple operations and RENAME

- Using:  $S \leftarrow R$

$TEMP \leftarrow \sigma_{DNO=5}(EMPLOYEE)$

$R(\text{First\_name}, \text{Last\_name}, \text{Salary}) \leftarrow \pi_{FNAME, LNAME, SALARY}(TEMP)$

TEMP

| Fname    | Minit | Lname   | <u>Ssn</u> | Bdate      | Address                 | Sex | Salary | Super_ssn | 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   |
| 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   |

R

| First_name | Last_name | Salary |
|------------|-----------|--------|
| John       | Smith     | 30000  |
| Franklin   | Wong      | 40000  |
| Ramesh     | Narayan   | 38000  |
| Joyce      | English   | 25000  |

# Example of applying multiple operations and RENAME

- Using  $\rho_S(B_1, B_2, \dots, B_n)(R)$

$TEMP \leftarrow \sigma_{DNO=5}(EMPLOYEE)$

$\rho_R(\text{First\_name}, \text{Last\_name}, \text{Salary})(\pi_{FNAME, LNAME, SALARY}(TEMP))$

TEMP

| Fname    | Minit | Lname   | <u>Ssn</u> | Bdate      | Address                 | Sex | Salary | Super_ssn | 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   |
| 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   |

R

| First_name | Last_name | Salary |
|------------|-----------|--------|
| John       | Smith     | 30000  |
| Franklin   | Wong      | 40000  |
| Ramesh     | Narayan   | 38000  |
| Joyce      | English   | 25000  |



# Relational Algebra Operations from Set Theory: UNION

- UNION Operation

- Binary operation, denoted by  $\cup$
- The result of  $R \cup S$ , is a relation that includes all tuples that are either in  $R$  or in  $S$  or in both  $R$  and  $S$
- Duplicate tuples are eliminated
- The two operand relations  $R$  and  $S$  must be “type compatible” (or UNION compatible):
  - $R$  and  $S$  must have same number of attributes
  - each pair of corresponding attributes must be type compatible (have same or compatible domains)

# Relational Algebra Operations from Set Theory: UNION

- **Example:**
  - To retrieve the social security numbers of all employees who either *work in department 5 or directly supervise an employee who works in department 5*

EMPLOYEE

| Fname    | Minit | Lname   | <u>Ssn</u> | Bdate      | Address                  | Sex | Salary | Super_ssn | 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  | 1968-01-19 | 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   |

# Relational Algebra Operations from Set Theory: UNION

- **Example:**

- To retrieve the social security numbers of all employees who either *work in department 5* (RESULT1 below) or *directly supervise an employee who works in department 5* (RESULT2 below)
- We can use the UNION operation as follows:

$\text{DEP5\_EMPS} \leftarrow \sigma_{\text{DNO}=5} (\text{EMPLOYEE})$

$\text{RESULT1} \leftarrow \pi_{\text{SSN}}(\text{DEP5\_EMPS})$

$\text{RESULT2}(\text{SSN}) \leftarrow \pi_{\text{SUPERSSN}}(\text{DEP5\_EMPS})$

$\text{RESULT} \leftarrow \text{RESULT1} \cup \text{RESULT2}$

- The union operation produces the tuples that are in either RESULT1 or RESULT2 or both

# Example of the Result of a UNION Operation

## ■ UNION Example

**EMPLOYEE**

| Fname    | Minit | Lname   | <u>Ssn</u> | Bdate      | Address                  | Sex | Salary | Super_ssn | 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  | 1968-01-19 | 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   |

Result of the  
UNION operation  
 $\text{RESULT} \leftarrow \text{RESULT1} \cup \text{RESULT2}$ .

**RESULT1**

| Ssn       |
|-----------|
| 123456789 |
| 333445555 |
| 666884444 |
| 453453453 |

**RESULT2**

| Ssn       |
|-----------|
| 333445555 |
| 888665555 |

**RESULT**

| Ssn       |
|-----------|
| 123456789 |
| 333445555 |
| 666884444 |
| 453453453 |
| 888665555 |

# Relational Algebra Operations from Set Theory

- Type Compatibility of operands is required for the binary set operation UNION  $\cup$ , (also for INTERSECTION  $\cap$ , and SET DIFFERENCE  $-$ )
- $R1(A1, A2, \dots, An)$  and  $R2(B1, B2, \dots, Bn)$  are type compatible if:
  - they have the same number of attributes, and
  - the domains of corresponding attributes are type compatible (i.e.  $\text{dom}(Ai) = \text{dom}(Bi)$  for  $i=1, 2, \dots, n$ ).
- The resulting relation for  $R1 \cup R2$  (also for  $R1 \cap R2$ , or  $R1 - R2$ ) has the same attribute names as the *first* operand relation  $R1$  (by convention)

# Relational Algebra Operations from Set Theory: INTERSECTION

- INTERSECTION is denoted by  $\cap$
- The result of the operation  $R \cap S$ , is a relation that includes all tuples that are in both R and S
  - The attribute names in the result will be the same as the attribute names in R
- The two operand relations R and S must be “type compatible”

- SET DIFFERENCE (also called MINUS or EXCEPT) is denoted by –
- The result of  $R - S$ , is a relation that includes all tuples that are in  $R$  but not in  $S$ 
  - The attribute names in the result will be the same as the attribute names in  $R$
- The two operand relations  $R$  and  $S$  must be “type compatible”

# Example to Illustrate the Result of UNION

STUDENT  $\cup$  INSTRUCTOR

**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    |

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



# Example to Illustrate the Result of INTERSECT

**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  $\cap$  INSTRUCTOR

| Fn     | Ln   |
|--------|------|
| Susan  | Yao  |
| Ramesh | Shah |

# Example to Illustrate the Result of SET DIFFERENCE

## STUDENT – INSTRUCTOR

**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    |

| Fn      | Ln      |
|---------|---------|
| Johnny  | Kohler  |
| Barbara | Jones   |
| Amy     | Ford    |
| Jimmy   | Wang    |
| Ernest  | Gilbert |

## INSTRUCTOR – STUDENT

| Fname   | Lname   |
|---------|---------|
| John    | Smith   |
| Ricardo | Browne  |
| Francis | Johnson |

# Some properties of UNION, INTERSECT, and SET DIFFERENCE

- Notice that both union and intersection are *commutative* operations; that is
  - $R \cup S = S \cup R$ , and  $R \cap S = S \cap R$
- Both union and intersection can be treated as **n-ary operations** applicable to any number of relations as both are *associative* operations; that is
  - $R \cup (S \cup T) = (R \cup S) \cup T$
  - can be written as:  $R \cup S \cup T$
  - $(R \cap S) \cap T = R \cap (S \cap T)$
  - can be written as:  $R \cap S \cap T$
- The minus operation is not commutative; that is, in general
  - $R - S \neq S - R$

- **CARTESIAN (or CROSS) PRODUCT Operation**
  - This operation is used to combine tuples from two relations in a combinatorial fashion.
  - Denoted by  $R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$
  - Result is a relation  $Q$  with degree  $n + m$  attributes:
    - **$Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$ , in that order.**
  - The resulting relation state has one tuple for each combination of tuples—one from  $R$  and one from  $S$ .
  - Hence, if  $R$  has  $n_R$  tuples (denoted as  $|R| = n_R$ ), and  $S$  has  $n_S$  tuples, then  $R \times S$  will have  $n_R * n_S$  tuples.
  - The two operands do NOT have to be "type compatible"

# Example of Applying CARTESIAN PRODUCT

EMPNAMES

| Fname    | Lname   | Ssn       |
|----------|---------|-----------|
| Alicia   | Zelaya  | 999887777 |
| Jennifer | Wallace | 987654321 |
| Joyce    | English | 453453453 |

DEPENDENT

| <u>Essn</u> | <u>Dependent_name</u> | Sex | Bdate      | Relationship |
|-------------|-----------------------|-----|------------|--------------|
| 333445555   | Alice                 | F   | 1986-04-05 | Daughter     |
| 333445555   | Theodore              | M   | 1983-10-25 | Son          |
| 333445555   | Joy                   | F   | 1958-05-03 | Spouse       |
| 987654321   | Abner                 | M   | 1942-02-28 | Spouse       |
| 123456789   | Michael               | M   | 1988-01-04 | Son          |
| 123456789   | Alice                 | F   | 1988-12-30 | Daughter     |
| 123456789   | Elizabeth             | F   | 1967-05-05 | Spouse       |

EMPNAMES x DEPENDENT

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 | ... |

- Generally, CROSS PRODUCT is not a meaningful operation
  - Can become meaningful when followed by other operations
- Example (not meaningful):
  - $FEMALE\_EMPS \leftarrow \sigma_{SEX='F'}(EMPLOYEE)$
  - $EMP\_NAMES \leftarrow \pi_{FNAME, LNAME, SSN}(FEMALE\_EMPS)$
  - $EMP\_DEPENDENTS \leftarrow EMP\_NAMES \times DEPENDENT$
- EMP\_DEPENDENTS will contain every combination of EMP\_NAMES and DEPENDENT
  - whether or not they are actually related

- To keep only combinations where the DEPENDENT is related to the EMPLOYEE, we add a SELECT operation as follows
- Example (meaningful):
  - $FEMALE\_EMPS \leftarrow \sigma_{SEX='F'}(EMPLOYEE)$
  - $EMP\_NAMES \leftarrow \pi_{FNAME, LNAME, SSN}(FEMALE\_EMPS)$
  - $EMP\_DEPENDENTS \leftarrow EMP\_NAMES \times DEPENDENT$
  - $ACTUAL\_DEPS \leftarrow \sigma_{SSN=ESSN}(EMP\_DEPENDENTS)$
  - $RESULT \leftarrow \pi_{FNAME, LNAME, DEPENDENT\_NAME}(ACTUAL\_DEPS)$
- RESULT will now contain the name of female employees and their dependents



# Example

## FEMALE\_EMPS

| Fname    | Minit | Lname   | Ssn       | Bdate      | Address                 | Sex | Salary | Super_ssn | 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 | ... |

## 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          |



# Binary Relational Operations: JOIN

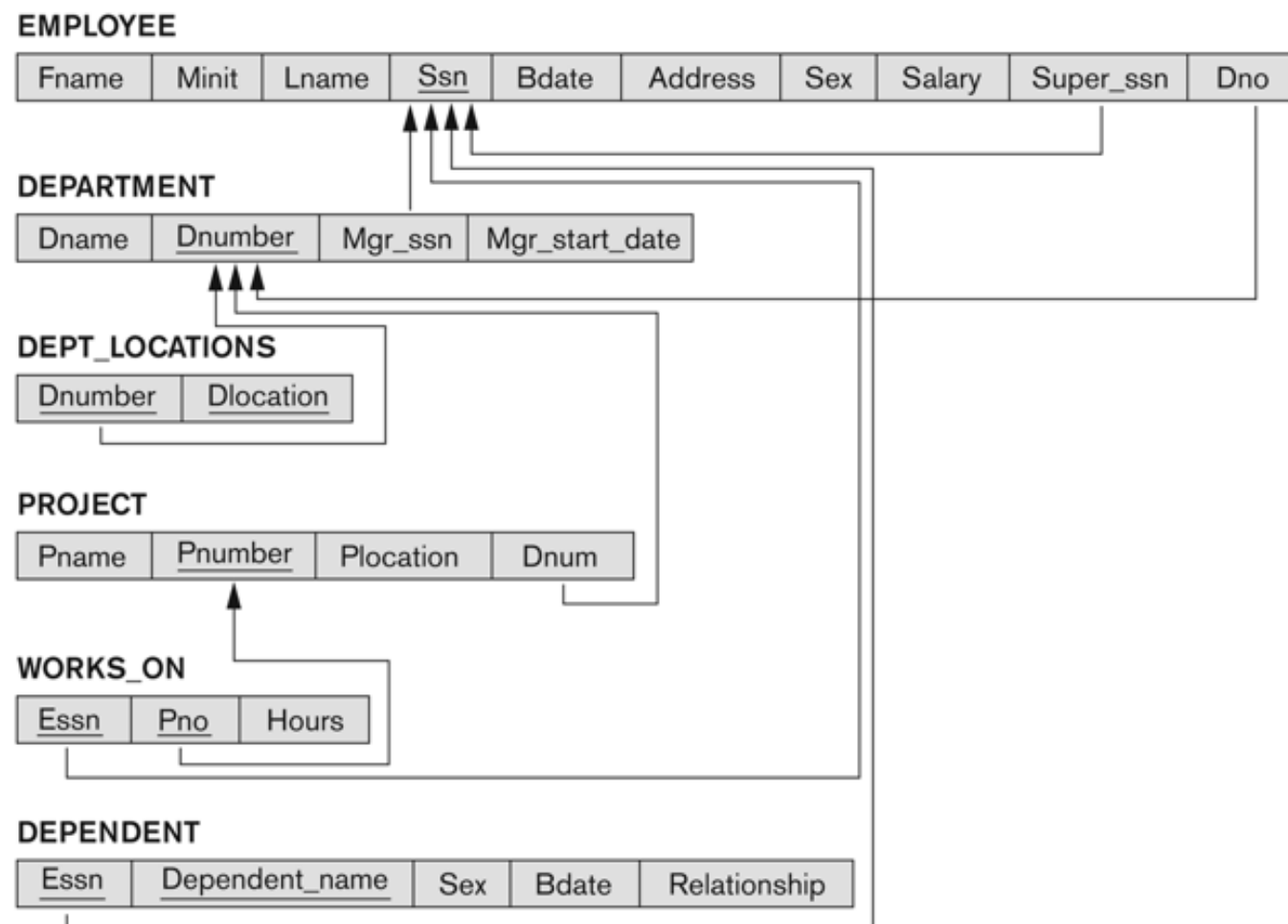
- JOIN Operation (denoted by  $\bowtie$ )
  - The sequence of CARTESIAN PRODUCT followed by SELECT is used quite commonly to identify and select related tuples from two relations
  - A special operation, called JOIN combines this sequence into a single operation
  - This operation is very important for any relational database with more than a single relation, because it allows us to *combine related tuples* from various relations

# Binary Relational Operations: JOIN

- JOIN Operation (denoted by  $\bowtie$  )
  - The general form of a join operation on two relations  $R(A_1, A_2, \dots, A_n)$  and  $S(B_1, B_2, \dots, B_m)$  is:
$$R \bowtie_{\langle \text{join condition} \rangle} S$$
  - where  $R$  and  $S$  can be any relations that result from general *relational algebra expressions*.

# Binary Relational Operations: JOIN (cont.)

- Example: Suppose that we want to retrieve the name of the manager of each department.
  - To get the manager's name, we need to combine each DEPARTMENT tuple with the EMPLOYEE tuple whose SSN value matches the MGRSSN value in the department tuple.
  - We do this by using the join  $\bowtie$  operation.



# Binary Relational Operations: JOIN (cont.)

- Example: Suppose that we want to retrieve the name of the manager of each department.
  - To get the manager's name, we need to combine each DEPARTMENT tuple with the EMPLOYEE tuple whose SSN value matches the MGRSSN value in the department tuple.
  - We do this by using the join  $\bowtie$  operation.
- $\text{DEPT\_MGR} \leftarrow \text{DEPARTMENT} \bowtie_{\text{MGRSSN}=\text{SSN}} \text{EMPLOYEE}$
- MGRSSN=SSN is the join condition
  - Combines each department record with the employee who manages the department
  - The join condition can also be specified as  $\text{DEPARTMENT.MGRSSN} = \text{EMPLOYEE.SSN}$

# Example of Applying the JOIN Operation

**EMPLOYEE**

| Fname    | Minit | Lname   | <u>Ssn</u> | Bdate      | Address                  | Sex | Salary | Super_ssn | 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  | 1968-01-19 | 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   |

**DEPARTMENT**

| Dname          | <u>Dnumber</u> | Mgr_ssn   | Mgr_start_date |
|----------------|----------------|-----------|----------------|
| Research       | 5              | 333445555 | 1988-05-22     |
| Administration | 4              | 987654321 | 1995-01-01     |
| Headquarters   | 1              | 888665555 | 1981-06-19     |

DEPT\_MGR ← DEPARTMENT  MGRSSN=SSN EMPLOYEE

**DEPT\_MGR**

| Dname          | Dnumber | Mgr_ssn   | ... | 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 | ... |

# Relational Algebra Overview

- Relational Algebra consists of several groups of operations
  - **Unary Relational Operations**
    - SELECT (symbol:  $\sigma$  (sigma))
    - PROJECT (symbol:  $\pi$  (pi))
    - RENAME (symbol:  $\rho$  (rho))
  - **Relational Algebra Operations From Set Theory**
    - UNION ( $\cup$ ), INTERSECTION ( $\cap$ ), SET DIFFERENCE (or MINUS,  $-$ )
    - CARTESIAN PRODUCT ( $\times$ )
  - **Binary Relational Operations**
    - JOIN (several variations of JOIN exist)
    - DIVISION
  - Additional Relational Operations
    - OUTER JOINS, OUTER UNION
    - AGGREGATE FUNCTIONS (These compute summary of information: for example, SUM, COUNT, AVG, MIN, MAX)

# Quiz Preparation

- You will only need: a **Black Pen** and **Student ID Card**  
**Answer Sheet on Last Page**

## TURN IN ONLY YOUR ANSWER SHEET!

Student ID, filled in and written

→ *ID goes this way* →

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|   | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 |   | 1 | 1 | 1 | 1 | 1 |
| 2 | 2 |   | 2 | 2 | 2 | 2 |
| 3 | 3 | 3 |   | 3 | 3 | 3 |
| 4 | 4 | 4 | 4 |   | 4 | 4 |
| 5 | 5 | 5 | 5 | 5 |   | 5 |
| 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| 9 | 9 | 9 | 9 | 9 | 9 |   |

Please write your full name and student number below.

Full name:

Chipp Jansen

Student number:

0123459

Please answer the questions in the grid below:

Question 1: ☐ B ☐ C ☐ D ☐ E

Question 2: ☐ A ☐ ☐ C ☐ D ☐ E

~~Question 3: ☐ A ☐ B ☒ ☒ ☐ E~~ Q3 is D

Question 4: ☐ A ☐ B ☐ C ☐ D ☐ E

Question 5: ☐ A ☐ B ☐ C ☐ D ☐ E

OOPS!

Strike through Question  
and Write Answer