Basic SQL

Part 3

Basic Retrieval Queries in SQL

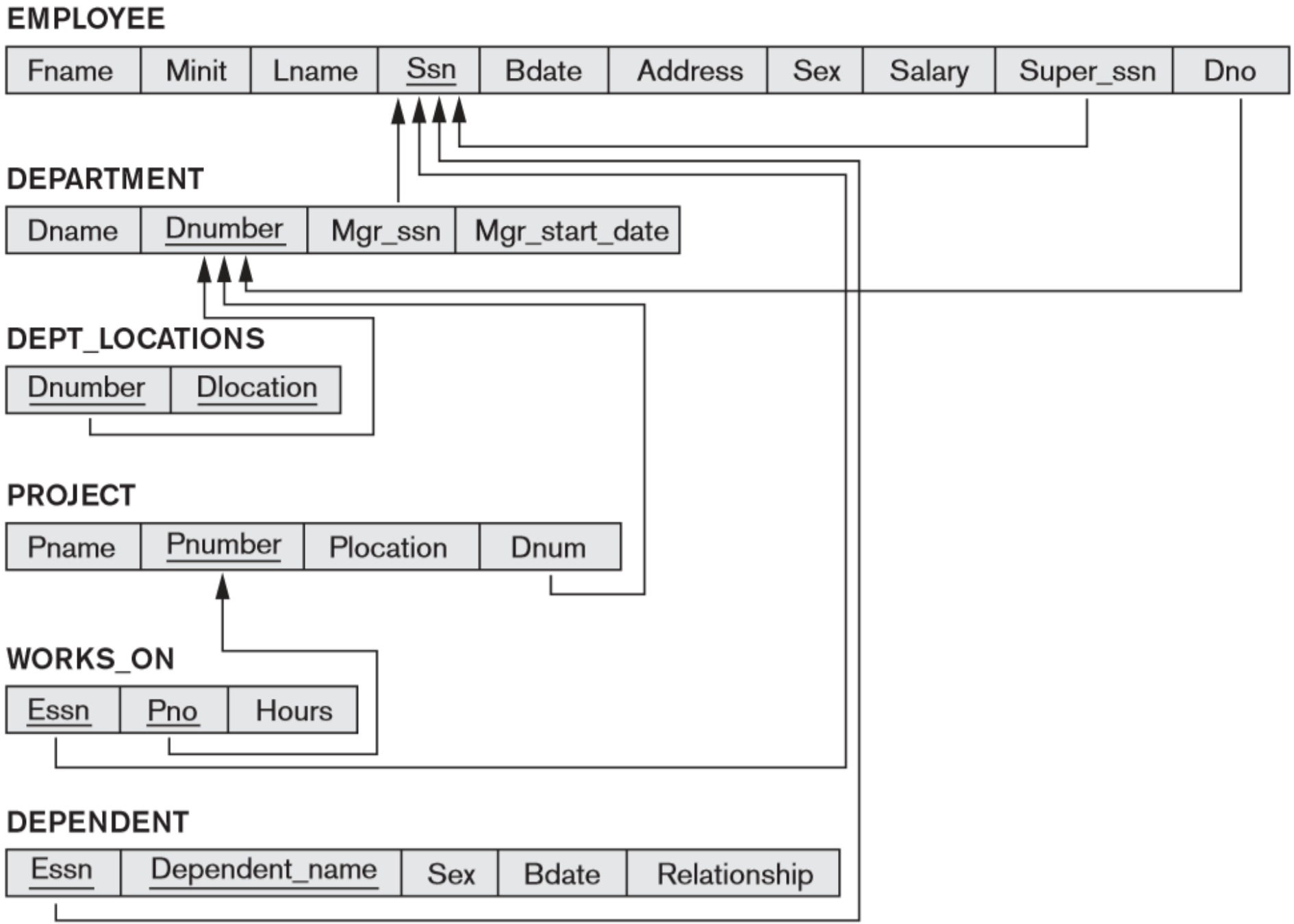
- SQL allows a table (relation) to have two or more tuples that are identical in all their attribute values.
- Hence, in general, an **SQL** table is not a *set of tuples*, because a set does not allow two identical members; rather, it is a **multiset** (sometimes called a *bag*) of tuples.

The SELECT-FROM-WHERE Structure of Basic SQL Queries

- **SELECT** <attribute list>
  FROM <table list>
  WHERE <condition>;

- **Query 0**. Retrieve the birth date and address of the employee(s) whose name is ‘John B. Smith'.
- Q0:

```
SELECT Bdate, Address
FROM   EMPLOYEE
WHERE  Fname='John' AND
       Minit='B' AND
       Lname='Smith';
```


EMPLOYEE
Fname
Minit
Lname
Ssn
Bdate
Address
Sex
Salary
Super_ssn
Dno
DEPARTMENT
Dname
Dnumber
Mgr_ssn
Mgr_start_date
DEPT_LOCATIONS
Dnumber
Dlocation
PROJECT
Pname
Pnumber
Plocation
Dnum
WORKS_ON
Essn
Pno
Hours
DEPENDENT
Essn
Dependent_name
Sex
Bdate
Relationship

EMPLOYEE

Fname	Minit	Lname	Ssn	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	Dnumber	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

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

Essn	Pno	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

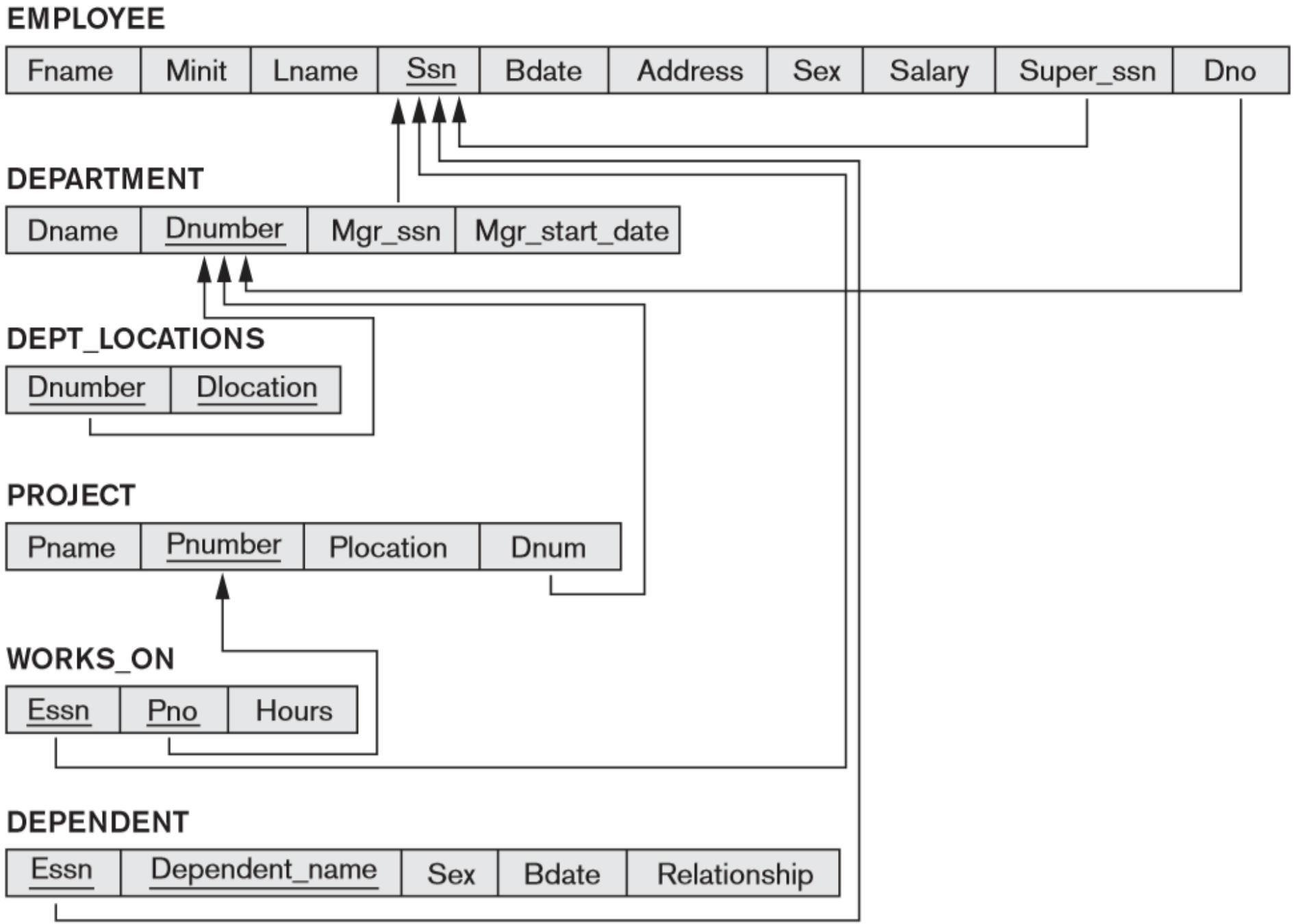
Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

Essn	Dependent_name	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

- **Query 1**. Retrieve the name and address of all employees who work for the 'Research' department.
- Q1:

```
SELECT Fname, Lname, Address
FROM   EMPLOYEE, DEPARTMENT
WHERE  Dname='Research' AND Dnumber=Dno;
```


EMPLOYEE
Fname
Minit
Lname
Ssn
Bdate
Address
Sex
Salary
Super_ssn
Dno
DEPARTMENT
Dname
Dnumber
Mgr_ssn
Mgr_start_date
DEPT_LOCATIONS
Dnumber
Dlocation
PROJECT
Pname
Pnumber
Plocation
Dnum
WORKS_ON
Essn
Pno
Hours
DEPENDENT
Essn
Dependent_name
Sex
Bdate
Relationship

EMPLOYEE

Fname	Minit	Lname	Ssn	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	Dnumber	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

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

Essn	Pno	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	Pnumber	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

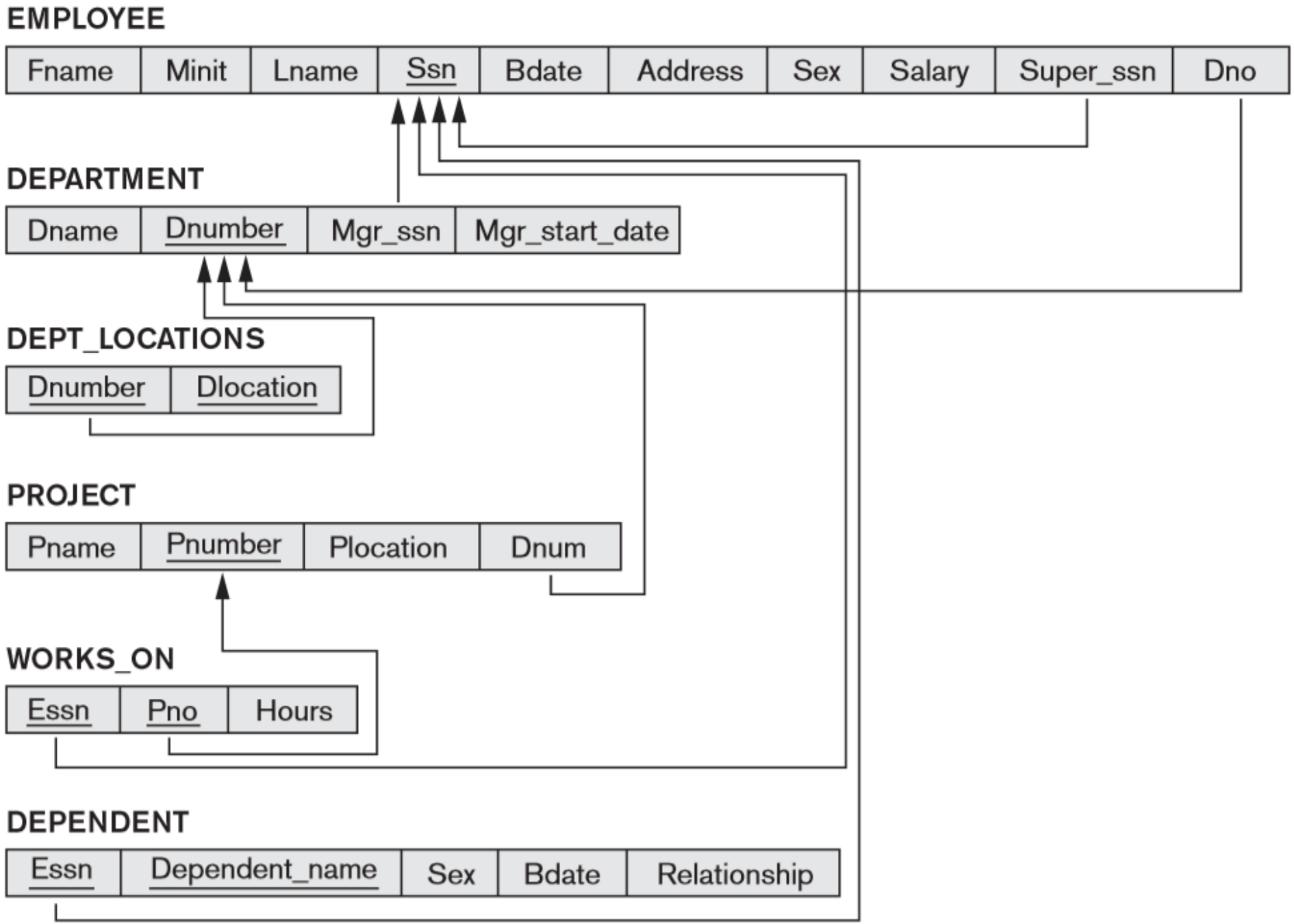
DEPENDENT

Essn	Dependent_name	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

- **Query 2**. For every project located in ‘Stafford’, list the project number, the controlling department number, and the department manager’s last name, address, and birth date.
- Q2:
```
SELECT Pnumber, Dnum, Lname, Address, Bdate
FROM   PROJECT, DEPARTMENT, EMPLOYEE
WHERE  Dnum=Dnumber AND Mgr_ssn=Ssn AND
       Plocation='Stafford';
```

Ambiguous Attribute Names, Aliasing, Renaming, and Tuple Variables

- Suppose that the `Dno` and `Lname` attributes of the `EMPLOYEE` relation were called `Dnumber` and `Name`, and the `Dname` attribute of `DEPARTMENT` was also called `Name`.
- **Query 1**. Retrieve the name and address of all employees who work for the 'Research' department.


EMPLOYEE
Fname
Minit
Lname
Ssn
Bdate
Address
Sex
Salary
Super_ssn
Dno
DEPARTMENT
Dname
Dnumber
Mgr_ssn
Mgr_start_date
DEPT_LOCATIONS
Dnumber
Dlocation
PROJECT
Pname
Pnumber
Plocation
Dnum
WORKS_ON
Essn
Pno
Hours
DEPENDENT
Essn
Dependent_name
Sex
Bdate
Relationship

- Q1A:

```
SELECT Fname, EMPLOYEE.Name, Address
FROM   EMPLOYEE, DEPARTMENT
WHERE  DEPARTMENT.Name='Research' AND
       DEPARTMENT.Dnumber=
       EMPLOYEE.Dnumber;
```

- Q1’
```
SELECT EMPLOYEE.Fname, EMPLOYEE.LName,
       EMPLOYEE.Address
FROM   EMPLOYEE, DEPARTMENT
WHERE  DEPARTMENT.DName='Research' AND
       DEPARTMENT.Dnumber=EMPLOYEE.Dno;
```

- **Query 8**. For each employee, retrieve the employee's first and last name and the first and last name of his or her immediate supervisor.
- Q8:

```
SELECT E.Fname, E.Lname, S.Fname, S.Lname
FROM   EMPLOYEE AS E, EMPLOYEE AS S
WHERE  E.Super_ssn=S.Ssn;
```

- We can use this alias-naming mechanism in any SQL query to specify tuple variables for every table in the WHERE clause, whether or not the same relation needs to be referenced more than once.
- Q1B:

```
SELECT E.Fname, E.LName, E.Address
FROM   EMPLOYEE E, DEPARTMENT D
WHERE  D.DName='Research' AND
       D.Dnumber=E.Dno;
```

Unspecified WHERE Clause and Use of the Asterisk

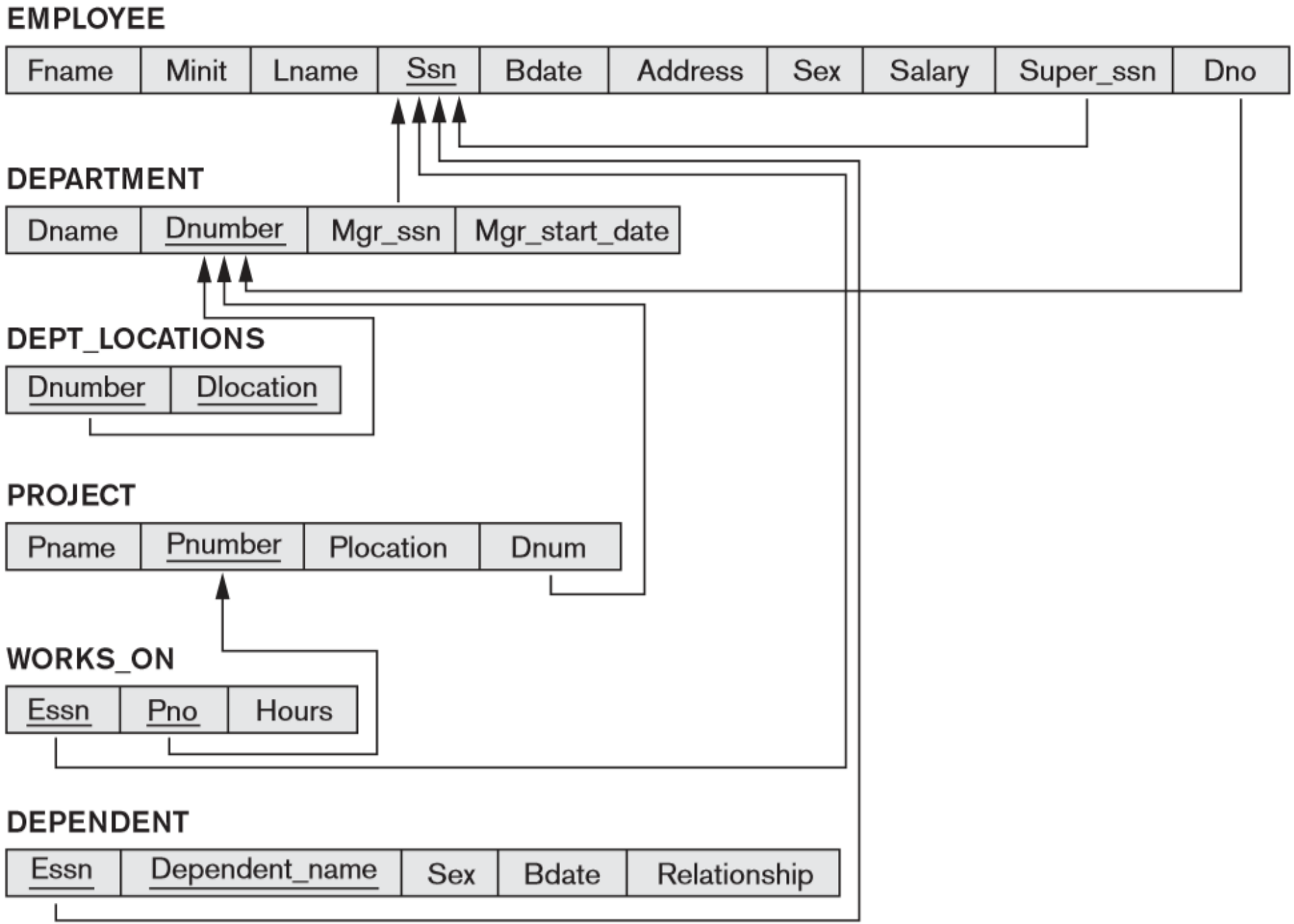
- **Queries 9 and 10**. Select all EMPLOYEE Ssns (Q9) and all combinations of EMPLOYEE Ssn and DEPARTMENT Dname (Q10) in the database.
- Q9:

```
SELECT Ssn
FROM   EMPLOYEE;
```

- Q10:

```
SELECT Ssn, Dname
FROM   EMPLOYEE, DEPARTMENT;
```

- Q1C:
```
SELECT *
FROM   EMPLOYEE WHERE Dno=5;
```
- Q1D:
```
SELECT *
FROM   EMPLOYEE, DEPARTMENT
WHERE  Dname='Research' AND
       Dno=Dnumber;
```
- Q10A:
```
SELECT *
FROM   EMPLOYEE, DEPARTMENT;
```


EMPLOYEE
Fname
Minit
Lname
Ssn
Bdate
Address
Sex
Salary
Super_ssn
Dno
DEPARTMENT
Dname
Dnumber
Mgr_ssn
Mgr_start_date
DEPT_LOCATIONS
Dnumber
Dlocation
PROJECT
Pname
Pnumber
Plocation
Dnum
WORKS_ON
Essn
Pno
Hours
DEPENDENT
Essn
Dependent_name
Sex
Bdate
Relationship

Tables as Sets in SQL

- SQL does not automatically eliminate duplicate tuples in the results of queries, for the following reasons:
  - Duplicate elimination is an expensive operation. One way to implement it is to sort the tuples first and then eliminate duplicates.
  - The user may want to see duplicate tuples in the result of a query.
  - When an aggregate function is applied to tuples, in most cases we do not want to eliminate duplicates.

- **Query 11**. Retrieve the salary of every employee (Q11) and all distinct salary values (Q11A).
- Q11:

```
SELECT ALL Salary
FROM   EMPLOYEE;
```

- Q11A:

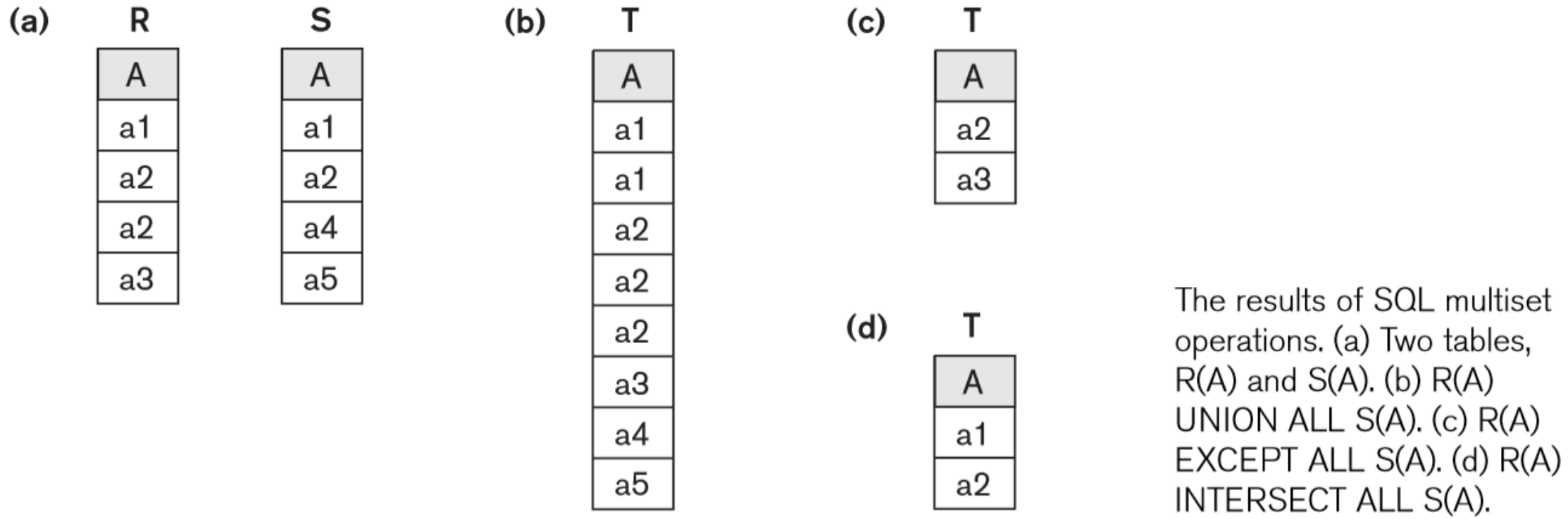
```
SELECT DISTINCT Salary
FROM   EMPLOYEE;
```

- SQL has directly incorporated some of the set operations from mathematical set theory, which are also part of relational algebra.
- There are set union (**UNION**), set difference (**EXCEPT**), and set intersection (**INTERSECT**) operations.
- The relations resulting from these set operations are sets of tuples; that is, *duplicate tuples are eliminated from the result*.
- These set operations apply only to *union-compatible relations*, so we must make sure that the two relations on which we apply the operation have the same attributes and that the attributes appear in the same order in both relations.

- **Query 4**. Make a list of all project numbers for projects that involve an employee whose last name is 'Smith', either as a worker or as a manager of the department that controls the project.
- Q4A:

```
(SELECT DISTINCT Pnumber
FROM    PROJECT, DEPARTMENT, EMPLOYEE
WHERE   Dnum=Dnumber AND
        Mgr_ssn=Ssn AND
        Lname='Smith')
UNION
(SELECT DISTINCT Pnumber
FROM    PROJECT, WORKS_ON, EMPLOYEE
WHERE   Pnumber=Pno AND
        Essn=Ssn AND
        Lname='Smith');
```

- SQL also has corresponding multiset operations, which are followed by the keyword **ALL** (`UNION ALL`, `EXCEPT ALL`, `INTERSECT ALL`).
- Their results are multisets (duplicates are not eliminated).



The results of SQL multiset operations. (a) Two tables, R(A) and S(A). (b) R(A) UNION ALL S(A). (c) R(A) EXCEPT ALL S(A). (d) R(A) INTERSECT ALL S(A).

Substring Pattern Matching and Arithmetic Operators

- The **LIKE** comparison operator allows comparison conditions on only parts of a character string.
- This can be used for string pattern matching.
- Partial strings are specified using two reserved characters: % replaces an arbitrary number of zero or more characters, and the underscore (_) replaces a single character.

- **Query 12**. Retrieve all employees whose address is in Houston, Texas.
- Q12:

```
SELECT Fname, Lname
FROM   EMPLOYEE
WHERE  Address LIKE '%Houston,TX%';
```

- **Query 12A**. Find all employees who were born during the 1950s.
- Q12:

```
SELECT Fname, Lname
FROM   EMPLOYEE
WHERE  Bdate LIKE '_ _ 5 _ _ _ _ _ _ _';
```

- If an underscore or `%` is needed as a literal character in the string, the character should be preceded by an *escape character*, which is specified after the string using the keyword `ESCAPE`.
- For example, `'AB\_CD\%EF' ESCAPE '\'` represents the literal string `'AB_CD%EF'` because `\` is specified as the escape character.
- Any character not used in the string can be chosen as the escape character.

- Also, we need a rule to specify apostrophes or single quotation marks (' ') if they are to be included in a string because they are used to begin and end strings.
- If an apostrophe (') is needed, it is represented as two consecutive apostrophes ('') so that it will not be interpreted as ending the string.

- Another feature allows the use of arithmetic in queries.
- The standard arithmetic operators for addition ($+$), subtraction ($-$), multiplication ($*$), and division ($/$) can be applied to numeric values or attributes with numeric domains.

- **Query 13**. Show the resulting salaries if every employee working on the 'ProductX' project is given a 10 percent raise.
- Q13:

```
SELECT E.Fname, E.Lname,
       1.1 * E.Salary AS Increased_sal
FROM   EMPLOYEE AS E, WORKS_ON AS W,
       PROJECT AS P
WHERE  E.Ssn=W.Essn AND
       W.Pno=P.Pnumber AND
       P.Pname='ProductX';
```

- For string data types, the concatenate operator || can be used in a query to append two string values.
- For date, time, timestamp, and interval data types, operators include incrementing (+) or decrementing (−) a date, time, or timestamp by an interval.
- In addition, an interval value is the result of the difference between two date, time, or timestamp values.
- Another comparison operator, which can be used for convenience, is **BETWEEN**.

- **Query 14**. Retrieve all employees in department 5 whose salary is between $30,000 and $40,000.
- Q14:

```
SELECT *
FROM   EMPLOYEE
WHERE  (Salary BETWEEN 30000 AND 40000)
       AND
       Dno = 5;
```

Ordering of Query Results

- **Query 15**. Retrieve a list of employees and the projects they are working on, ordered by department and, within each department,ordered alphabetically by last name, then first name.
- Q15:

```
SELECT    D.Dname, E.Lname, E.Fname,
          P.Pname
FROM      DEPARTMENT D, EMPLOYEE E,
          WORKS_ON W, PROJECT P
WHERE     D.Dnumber=E.Dno AND
          E.Ssn=W.Essn AND
          W.Pno=P.Pnumber
ORDER BY  D.Dname, E.Lname, E.Fname;
```

- The default order is in ascending order of values.
- We can specify the keyword **DESC** if we want to see the result in a descending order of values.
- The keyword **ASC** can be used to specify ascending order explicitly.

- `ORDER BY D.Dname DESC, E.Lname ASC, E.Fname ASC`