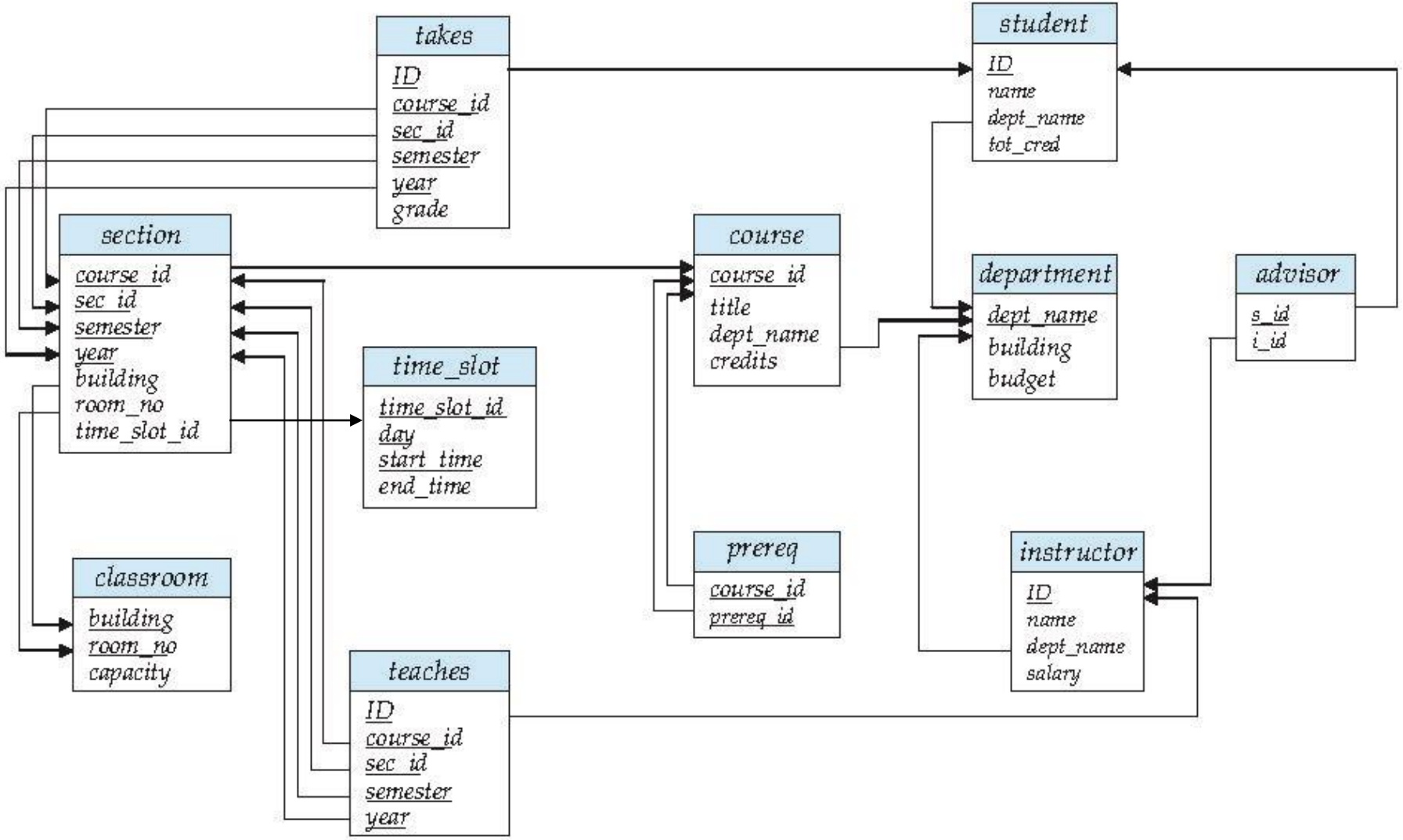
Chapter 3: Introduction to SQL

Edited by Radhika Sukapuram. Original slides by

Database System Concepts, 6th Ed.

Schema Diagram for University Database



Natural Join cont.

- List the names of instructors along with the courses that they teach

  select *name, title*
  from *instructor* **natural join** *teaches, course*
  where *teaches.course_id = course.course_id;*

Does the query below compute the same result ?

  select *name, title*
  from *instructor* **natural join** *teaches* **natural join** *course*

Natural Join cont.

- List the names of instructors along with the courses that they teach

select *name, title*
from *instructor* **natural join** *teaches, course*
where *teaches.course_id = course.course_id;*

Does the query below compute the same result ?

select *name, title*
from *instructor* **natural join** *teaches* **natural join** *course*

No!

To avoid equating attributes wrongly,

select *name, title*
from (*instructor* **natural join** *teaches)* **join** *course* **using** *(course_id);*

The Rename Operation

- The SQL allows renaming relations and attributes using the **as** clause:

  old-name **as** *new-name*

- Find the names of all instructors who have a higher salary than some instructor in 'Comp. Sci'.
  - **select distinct** *T.name*
    from *instructor* **as** *T, instructor* **as** *S*
    where *T.salary* > *S.salary* **and** *S.dept_name* = *'Comp. Sci.'*

- Keyword **as** is optional and may be omitted

  instructor **as** *T* ≡ *instructor T*

Note: T is called a correlation name, table alias, correlation variable or tuple variable.

String Operations

- SQL includes a string-matching operator for comparisons on character strings. The operator **like** uses patterns that are described using two special characters:
  - percent (%). The % character matches any substring.
  - underscore (_). The _ character matches any character.
- Find the names of all instructors whose name includes the substring “dar”.

```
select name
from instructor
where name like '%dar%'
```

- Match the string “100%”

```
like ‘100 \%'  escape  '\'
```

in that above we use backslash (\) as the escape character.

String Operations (Cont.)

- Patterns are case sensitive.
- Pattern matching examples:
  - ‘Intro%’ matches any string beginning with “Intro”.
  - ‘%Comp%’ matches any string containing “Comp” as a substring.
  - ‘_ _ _’ matches any string of exactly three characters.
  - ‘_ _ _ %’ matches any string of at least three characters.

- SQL supports a variety of string operations such as
  - concatenation (using “||”)
  - converting from upper to lower case (and vice versa)
  - finding string length, extracting substrings, etc.

Ordering the Display of Tuples

- List in alphabetic order the names of all instructors

  select distinct *name*
  from *instructor*
  order by *name*

- We may specify **desc** for descending order or **asc** for ascending order, for each attribute; ascending order is the default.
  - Example: **order by** *name* **desc**
- Can sort on multiple attributes
  - Example: **order by** *dept_name, name*

Where Clause Predicates

- SQL includes a **between** comparison operator
- Example: Find the names of all instructors with salary between $90,000 and $100,000 (that is, ≥ $90,000 and ≤ $100,000)
  - **select** *name*
    from *instructor*
    where *salary* **between** 90000 **and** 100000
- Tuple comparison
  - **select** *name*, *course_id*
    from *instructor*, *teaches*
    where (*instructor*.*ID*, *dept_name*) = (*teaches*.*ID*, 'Biology');

section

course_id	*sec_id*	*semester*	*year*	*building*	*room_number*	*time_slot_id*
BIO-101	1	Summer	2009	Painter	514	B
BIO-301	1	Summer	2010	Painter	514	A
CS-101	1	Fall	2009	Packard	101	H
CS-101	1	Spring	2010	Packard	101	F
CS-190	1	Spring	2009	Taylor	3128	E
CS-190	2	Spring	2009	Taylor	3128	A
CS-315	1	Spring	2010	Watson	120	D
CS-319	1	Spring	2010	Watson	100	B
CS-319	2	Spring	2010	Taylor	3128	C
CS-347	1	Fall	2009	Taylor	3128	A
EE-181	1	Spring	2009	Taylor	3128	C
FIN-201	1	Spring	2010	Packard	101	B
HIS-351	1	Spring	2010	Painter	514	C
MU-199	1	Spring	2010	Packard	101	D
PHY-101	1	Fall	2009	Watson	100	A

Set Operations

- Find courses that ran in Fall 2009 or in Spring 2010
- Find courses that ran in Fall 2009 and in Spring 2010
- Find courses that ran in Fall 2009 but not in Spring 2010

Set Operations

- Find courses that ran in Fall 2009 or in Spring 2010

  (**select** *course_id* **from** *section* **where** *sem* = ‘Fall’ **and** *year* = 2009)
  union
  (**select** *course_id* **from** *section* **where** *sem* = ‘Spring’ **and** *year* = 2010)

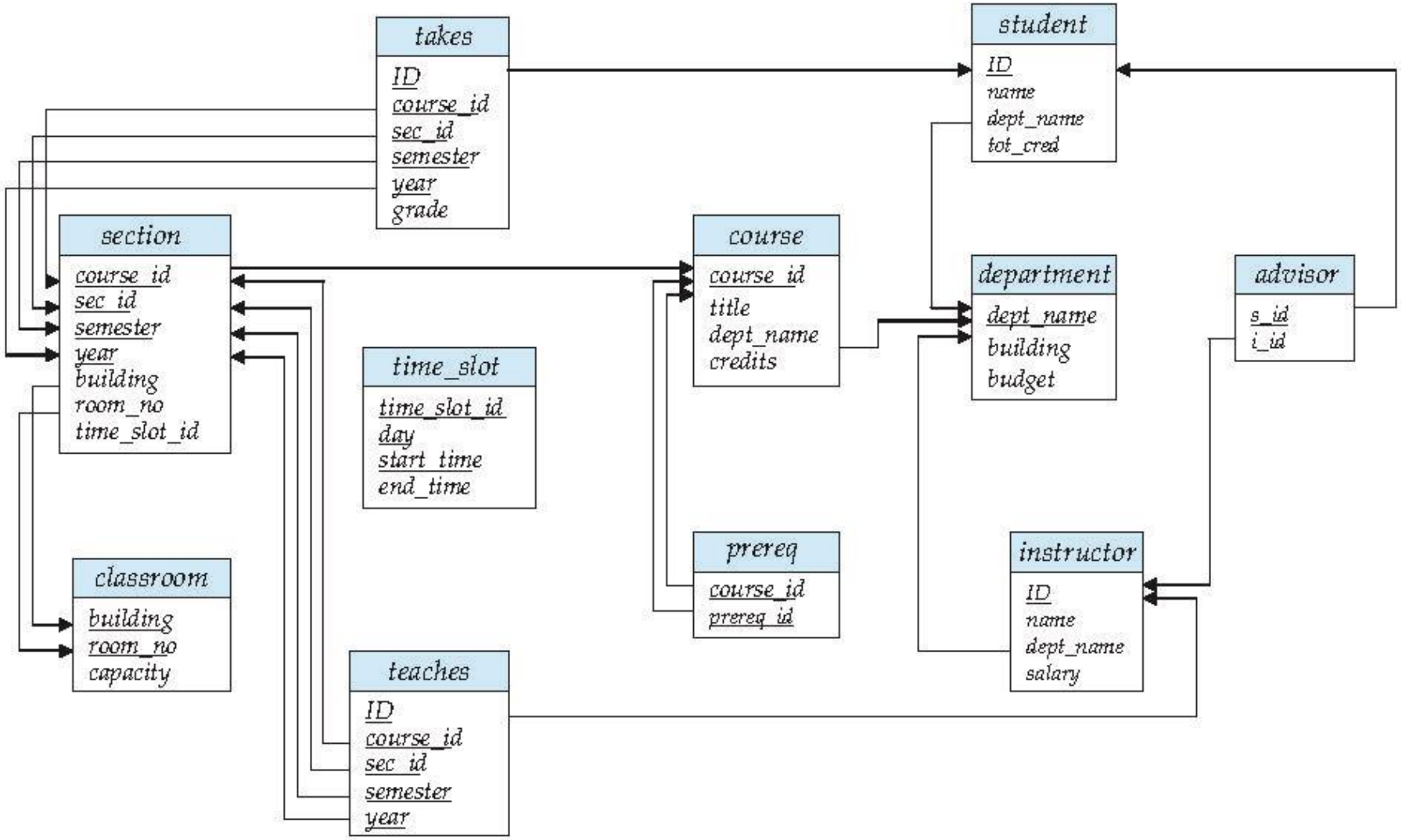
- Find courses that ran in Fall 2009 and in Spring 2010

  (**select** *course_id* **from** *section* **where** *sem* = ‘Fall’ **and** *year* = 2009)
  intersect
  (**select** *course_id* **from** *section* **where** *sem* = ‘Spring’ **and** *year* = 2010)

- Find courses that ran in Fall 2009 but not in Spring 2010

  (**select** *course_id* **from** *section* **where** *sem* = ‘Fall’ **and** *year* = 2009)
  except
  (**select** *course_id* **from** *section* **where** *sem* = ‘Spring’ **and** *year* = 2010)

Set Operations



Set Operations (Cont.)

- Find the largest salary of all instructors.
  - Find the salaries of all instructors that are less than the largest salary.
  - Find all the salaries of all instructors
  - Find the largest salary of all instructors.

Set Operations (Cont.)

- Find the salaries of all instructors that are less than the largest salary.
  - **select distinct** *T.salary*
    from *instructor* **as** *T, instructor* **as** *S*
    where *T.salary* < *S.salary*
- Find all the salaries of all instructors
  - **select distinct** *salary*
    from *instructor*
- Find the largest salary of all instructors.
  - (**select** "second query")
    except
    (**select** "first query")

Set Operations (Cont.)

- Set operations **union**, **intersect**, and **except**
  - Each of the above operations automatically eliminates duplicates
- To retain all duplicates use the corresponding multiset versions **union all**, **intersect all** and **except all**.
- Suppose a tuple occurs *m* times in *r* and *n* times in *s*, then, it occurs:
  - *m* + *n* times in *r* **union all** *s*
  - min(*m*,*n*) times in *r* **intersect all** *s*
  - max(0, *m* – *n*) times in *r* **except all** *s*

Null Values

- It is possible for tuples to have a null value, denoted by *null*, for some of the attributes
- *null* signifies an unknown value or that a value does not exist.
- The result of any arithmetic expression involving *null* is *null*
  - Example: 5 + *null* returns null
- The predicate **is null** can be used to check for null values.
  - Example: Find all instructors whose salary is null.

    select *name*
    from *instructor*
    where *salary* **is null**

Null Values and Three Valued Logic

- Three values – *true*, *false*, *unknown*
- Any comparison with *null* returns *unknown*
  - Example*: 5 < null or null <> null or null = null*
- Three-valued logic using the value *unknown*:
  - OR: (*unknown* **or** *true*) = *true*,
    (*unknown* **or** *false*) = *unknown*
    (*unknown* **or** *unknown*) = *unknown*
  - AND: *(true* **and** *unknown)* = *unknown,*
    (false **and** *unknown) = false,*
    (unknown **and** *unknown) = unknown*
  - NOT*:* *(***not** *unknown) = unknown*
  - "*P* **is unknown**" evaluates to true if predicate *P* evaluates to *unknown*
- Result of **where** clause predicate is treated as *false* if it evaluates to *unknown*

Where predicate with null

- Example: Find all instructors whose salary is greater than 5000.

  select *name*
  from *instructor*
  where *salary > 500*

Suppose a record has salary = null.

What value will the predicate return for that record ?

Will it be treated as true or false ?

Where predicate with null

- Example: Find all instructors whose salary is greater than 5000.

  select *name*
  from *instructor*
  where *salary > 500*

Suppose a record has salary = null.

What value will the predicate return for that record ?

Will it be treated as true or false ?

As false

null: select distinct, set operations

- Eliminating duplicate tuples

  select distinct *name*
  from *instructor*
  where *salary > 500*

Suppose two records have name as null

Will both the records be retained ?

null: select distinct, set operations

- Eliminating duplicate tuples

  select distinct *name*
  from *instructor*
  where *salary > 500*

Suppose two records have name as null

Will both the records be retained ?

No, only one will be retained – treatment different from predicates

- Same treatment for set operations