
# DATABASE SYSTEMS

CS – 355/ CE – 373

Instructor: Maria N. Samad

September 23rd, 2024

# SET DIFFERENCE (-)

- The set-difference operation allows us to find tuples that are in one relation but are not in another.
- Notation $r - s$
- Set differences must be taken between **compatible** relations.
  - $r$ and $s$ must have the same arity
  - attribute domains of $r$ and $s$ must be compatible

# SET DIFFERENCE (-)

- Example:
  - Find all courses taught in the Fall 2017 semester, but not in the Spring 2018 semester

- Query:
  - $\Pi_{course\_id}\,(\sigma_{semester=\text{“Fall”} \wedge year=2017}(section))\; -\; \Pi_{course\_id}\,(\sigma_{semester=\text{“Spring”} \wedge year=2018}(section))$

- Result →

| *course_id* |
|---|
| CS-347 |
| PHY-101 |

| *course_id* | *sec_id* | *semester* | *year* | *building* | *room_number* | *time_slot_id* |
|---|---|---|---|---|---|---|
| BIO-101 | 1 | Summer | 2017 | Painter | 514 | B |
| BIO-301 | 1 | Summer | 2018 | Painter | 514 | A |
| CS-101 | 1 | Fall | 2017 | Packard | 101 | H |
| CS-101 | 1 | Spring | 2018 | Packard | 101 | F |
| CS-190 | 1 | Spring | 2017 | Taylor | 3128 | E |
| CS-190 | 2 | Spring | 2017 | Taylor | 3128 | A |
| CS-315 | 1 | Spring | 2018 | Watson | 120 | D |
| CS-319 | 1 | Spring | 2018 | Watson | 100 | B |
| CS-319 | 2 | Spring | 2018 | Taylor | 3128 | C |
| CS-347 | 1 | Fall | 2017 | Taylor | 3128 | A |
| EE-181 | 1 | Spring | 2017 | Taylor | 3128 | C |
| FIN-201 | 1 | Spring | 2018 | Packard | 101 | B |
| HIS-351 | 1 | Spring | 2018 | Painter | 514 | C |
| MU-199 | 1 | Spring | 2018 | Packard | 101 | D |
| PHY-101 | 1 | Fall | 2017 | Watson | 100 | A |

The *section* relation.

# SET DIFFERENCE (-) – EXERCISES

- Activity Sheet:
  - Attempt **Part E**

# SET DIFFERENCE (-) – EXERCISES

- Activity Sheet **Part E** Solution:

| | |
|---|---|
| 1 | $\Pi_{ID}(student) - \Pi_{ID}(\sigma_{(semester = \text{“Spring”} \wedge year = 2024)}(takes))$ |
| 2 | $\Pi_{course\_id}(course) - \Pi_{course\_id}(\sigma_{(semester = \text{“Fall”} \wedge year = 2023)}(teaches) \cup \sigma_{(semester = \text{“Spring”} \wedge year = 2024)}(teaches))$ |
| 3 | $\Pi_{course\_id}(course) - \Pi_{course\_id}(prereq)$ |

# CARTESIAN PRODUCT (x)

- The Cartesian-product operation (denoted by x) allows us to combine information from any two relations.
- Example: The Cartesian product of the relations *instructor* and t*eaches* is written as:

  *instructor* x *teaches*

- We construct a tuple of the result out of each possible pair of tuples: one from the *instructor* relation and one from the *teaches* relation (see next slide)

# CARTESIAN PRODUCT (x)

- Since the instructor *ID* appears in both relations we distinguish between these attributes by attaching to the attribute the name of the relation from which the attribute originally came.
  - *instructor.ID*
  - *teaches.ID*

# CARTESIAN PRODUCT (x)

- The *instructor* relation:

| *ID* | *name* | *dept_name* | *salary* |
|---|---|---|---|
| 22222 | Einstein | Physics | 95000 |
| 12121 | Wu | Finance | 90000 |
| 32343 | El Said | History | 60000 |
| 45565 | Katz | Comp. Sci. | 75000 |
| 98345 | Kim | Elec. Eng. | 80000 |
| 76766 | Crick | Biology | 72000 |
| 10101 | Srinivasan | Comp. Sci. | 65000 |
| 58583 | Califieri | History | 62000 |
| 83821 | Brandt | Comp. Sci. | 92000 |
| 15151 | Mozart | Music | 40000 |
| 33456 | Gold | Physics | 87000 |
| 76543 | Singh | Finance | 80000 |

The *instructor* relation.

- The *teaches* relation:

| *ID* | *course_id* | *sec_id* | *semester* | *year* |
|---|---|---|---|---|
| 10101 | CS-101 | 1 | Fall | 2017 |
| 10101 | CS-315 | 1 | Spring | 2018 |
| 10101 | CS-347 | 1 | Fall | 2017 |
| 12121 | FIN-201 | 1 | Spring | 2018 |
| 15151 | MU-199 | 1 | Spring | 2018 |
| 22222 | PHY-101 | 1 | Fall | 2017 |
| 32343 | HIS-351 | 1 | Spring | 2018 |
| 45565 | CS-101 | 1 | Spring | 2018 |
| 45565 | CS-319 | 1 | Spring | 2018 |
| 76766 | BIO-101 | 1 | Summer | 2017 |
| 76766 | BIO-301 | 1 | Summer | 2018 |
| 83821 | CS-190 | 1 | Spring | 2017 |
| 83821 | CS-190 | 2 | Spring | 2017 |
| 83821 | CS-319 | 2 | Spring | 2018 |
| 98345 | EE-181 | 1 | Spring | 2017 |

The *teaches* relation.

The *cartesian* product, i.e. *instructor* x *teaches* →

| *instructor.ID* | *name* | *dept_name* | *salary* | *teaches.ID* | *course_id* | *sec_id* | *semester* | *year* |
|---|---|---|---|---|---|---|---|---|
| 10101 | Srinivasan | Comp. Sci. | 65000 | 10101 | CS-101 | 1 | Fall | 2017 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 10101 | CS-315 | 1 | Spring | 2018 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 10101 | CS-347 | 1 | Fall | 2017 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 12121 | FIN-201 | 1 | Spring | 2018 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 15151 | MU-199 | 1 | Spring | 2018 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 22222 | PHY-101 | 1 | Fall | 2017 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 12121 | Wu | Finance | 90000 | 10101 | CS-101 | 1 | Fall | 2017 |
| 12121 | Wu | Finance | 90000 | 10101 | CS-315 | 1 | Spring | 2018 |
| 12121 | Wu | Finance | 90000 | 10101 | CS-347 | 1 | Fall | 2017 |
| 12121 | Wu | Finance | 90000 | 12121 | FIN-201 | 1 | Spring | 2018 |
| 12121 | Wu | Finance | 90000 | 15151 | MU-199 | 1 | Spring | 2018 |
| 12121 | Wu | Finance | 90000 | 22222 | PHY-101 | 1 | Fall | 2017 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 15151 | Mozart | Music | 40000 | 10101 | CS-101 | 1 | Fall | 2017 |
| 15151 | Mozart | Music | 40000 | 10101 | CS-315 | 1 | Spring | 2018 |
| 15151 | Mozart | Music | 40000 | 10101 | CS-347 | 1 | Fall | 2017 |
| 15151 | Mozart | Music | 40000 | 12121 | FIN-201 | 1 | Spring | 2018 |
| 15151 | Mozart | Music | 40000 | 15151 | MU-199 | 1 | Spring | 2018 |
| 15151 | Mozart | Music | 40000 | 22222 | PHY-101 | 1 | Fall | 2017 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 22222 | Einstein | Physics | 95000 | 10101 | CS-101 | 1 | Fall | 2017 |
| 22222 | Einstein | Physics | 95000 | 10101 | CS-315 | 1 | Spring | 2018 |
| 22222 | Einstein | Physics | 95000 | 10101 | CS-347 | 1 | Fall | 2017 |
| 22222 | Einstein | Physics | 95000 | 12121 | FIN-201 | 1 | Spring | 2018 |
| 22222 | Einstein | Physics | 95000 | 15151 | MU-199 | 1 | Spring | 2018 |
| 22222 | Einstein | Physics | 95000 | 22222 | PHY-101 | 1 | Fall | 2017 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |

# CARTESIAN PRODUCT (x)

- The Cartesian-Product

  *instructor* x *teaches*

  associates every tuple of instructor with every tuple of teaches.

- Most of the resulting rows have information about instructors who did *not* teach a particular course.
- To get only those tuples of "*instructor* x *teaches* " that pertain to instructors and the courses that they taught, we write:

$$\sigma_{instructor.id = teaches.id} (instructor \times teaches)$$

# CARTESIAN PRODUCT (x)

- The table corresponding to: $\sigma_{instructor.id = teaches.id}$ (*instructor* x *teaches*))

| *instructor.ID* | *name* | *dept_name* | *salary* | *teaches.ID* | *course_id* | *sec_id* | *semester* | *year* |
|---|---|---|---|---|---|---|---|---|
| 10101 | Srinivasan | Comp. Sci. | 65000 | 10101 | CS-101 | 1 | Fall | 2017 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 10101 | CS-315 | 1 | Spring | 2018 |
| 10101 | Srinivasan | Comp. Sci. | 65000 | 10101 | CS-347 | 1 | Fall | 2017 |
| 12121 | Wu | Finance | 90000 | 12121 | FIN-201 | 1 | Spring | 2018 |
| 15151 | Mozart | Music | 40000 | 15151 | MU-199 | 1 | Spring | 2018 |
| 22222 | Einstein | Physics | 95000 | 22222 | PHY-101 | 1 | Fall | 2017 |
| 32343 | El Said | History | 60000 | 32343 | HIS-351 | 1 | Spring | 2018 |
| 45565 | Katz | Comp. Sci. | 75000 | 45565 | CS-101 | 1 | Spring | 2018 |
| 45565 | Katz | Comp. Sci. | 75000 | 45565 | CS-319 | 1 | Spring | 2018 |
| 76766 | Crick | Biology | 72000 | 76766 | BIO-101 | 1 | Summer | 2017 |
| 76766 | Crick | Biology | 72000 | 76766 | BIO-301 | 1 | Summer | 2018 |
| 83821 | Brandt | Comp. Sci. | 92000 | 83821 | CS-190 | 1 | Spring | 2017 |
| 83821 | Brandt | Comp. Sci. | 92000 | 83821 | CS-190 | 2 | Spring | 2017 |
| 83821 | Brandt | Comp. Sci. | 92000 | 83821 | CS-319 | 2 | Spring | 2018 |
| 98345 | Kim | Elec. Eng. | 80000 | 98345 | EE-181 | 1 | Spring | 2017 |

# CARTESIAN PRODUCT (x)

- If we want to further reduce this and get the names and course IDs of instructor teaching a course, instead of complete records, then we will add projection to the query

$$\Pi_{name,\ courseID}\ (\sigma_{instructor.id\ =\ teaches.id}\ (instructor \times teaches))$$

# CARTESIAN PRODUCT (x) – EXERCISES

- Activity Sheet:
  - Attempt **Part F**

# CARTESIAN PRODUCT (x) – EXERCISES

- Activity Sheet **Part F** Solution:

| | |
|---|---|
| 1 | $\Pi_{\text{name}} (\sigma_{(\text{student.id} = \text{takes.id} \wedge \text{course\_id} = \text{“CS-101”} \wedge \text{grade} > 80)} (\textit{student} \times \textit{takes}))$ |
| 2 | $\Pi_{\text{name, student.dept\_name, title}} (\sigma_{\text{course.dept\_name} = \text{student.dept\_name}} (\textit{course} \times \textit{student}))$ |
| 3 | $\Pi_{\text{capacity}} (\sigma_{\text{deparment.building} = \text{classroom.building} \wedge \text{dept\_name} = \text{“ECE”}} (\textit{classroom} \times \textit{department}))$ |

# RENAME(ρ)

- The results of relational-algebra expressions do not have a name that we can use to refer to them.
- The rename operator, $\rho$, is provided for that purpose
- Notation: $\rho_x (R)$
- The rename operation can be denoted by one of the following three forms:
  - $\rho_x (R)$
  - $\rho_{(A1,A2,A3\ldots,An)} (R)$
  - $\rho_{x\ (A1,A2,A3\ldots,An)} (R)$

# RENAME(ρ)

- $\rho_x$ (*R*)
  - Renames the relation *R* under the new name *x*
  - For example, in the given query we want to rename the relation name, *instructor* to *professor*:
    - Query: $\prod_{name}(\sigma_{dept\_name = \text{“Physics”}}$ (*instructor*))
    - Renaming with the query:
      - $\prod_{name}(\sigma_{dept\_name = \text{“Physics”}}$ ($\rho_{professor}$ (*instructor*)))
    - From here onwards until the program ends, the relation will be used with the name *professor*

# RENAME(ρ)

- $\rho_{(A1,A2,A3\ldots,An)}\ (R)$
  - Renames the attributes in the relation *R* to *A1*, *A2*, *A3*, …, *An*
  - For example, in the given query we want to rename only one of the attributes of relation *instructor*, i.e. *salary* should now be called *pay*, and the rest remains the same:
    - Query: $\prod_{salary}(\sigma_{dept\_name\ =\ \text{“Physics”}\ \wedge\ salary\ >\ 90000}\ (instructor))$
    - Renaming with the query:
      - $\prod_{pay}(\sigma_{dept\_name\ =\ \text{“Physics”}\ \wedge\ pay\ >\ 90000}\ (\rho_{(ID,\ name,\ dept\_name,\ pay)}\ (instructor)))$
    - From here onwards until the program ends, the relation variable, *salary* will be used with the name *pay*, hence the use of *pay* in the query

# RENAME(ρ)

- $\rho_{x\ (A1,A2,A3\ldots,An)}\ (R)$
  - Renames both the relation *R* to *x* ,as well as the attributes in the relation to *A1*, *A2*, *A3*, …, *An*
  - For example, the *instructor* relation should now be renamed as *professor*, as well as rename one of the attributes of relation, i.e. *salary* which should now be called *pay*, and the rest of the attributes remain the same:
    - Renaming operation:
      - $\rho_{professor\ (ID,\ name,\ dept\_name,\ pay)}\ (instructor)$
    - From here onwards until the program ends, the relation *instructor* will be accessed by the name *professor*, and variable, *salary* will be used with the name *pay*, instead

# RENAME(ρ) – EXERCISES

- Activity Sheet:
  - Attempt **Part G**

# RENAME(ρ) – EXERCISES

- Activity Sheet **Part G** Solution:

| | |
|---|---|
| 1 | $\sigma_{\text{(semester ="Fall" ^ year = 2022)}} (\rho_{\text{enrolled}} (takes))$ |
| 2 | $\Pi_{\text{s\_ID}} (\rho_{\text{mentors}} (advisors))$ |
| 3 | $\Pi_{\text{instructorName, instructorID}} (\sigma_{\text{dept\_name = "Physics" ^ pay > 90000}} (\rho_{\text{(instructorID, instructorName, dept\_name, pay)}} (instructor)))$ |
| 4 | $\Pi_{\text{instructorName, instructorID}} (\sigma_{\text{dept\_name = "Physics"}} (\rho_{\text{professor(instructorID, instructorName, dept\_name, pay)}} (instructor)))$ |
| 5 | $\rho_{\text{Fall24Enrollment}} (\sigma_{\text{(semester ="Fall" ^ year = 2024)}} (takes))$ |