CS 245, Database Management Systems
Midsem Exam, Winter 2022-2023
Department of Computer Science and Engineering
IIT Guwahati
Time: Two hours

Note: For Q1 and Q2 refer to the tables shown in Figure 1 at the end of the paper. Note that the tables extend from page 6 to page 9 and some of the tables span two pages. The list of tables is as follows: **instructor**, **student**, **department**, **course**, **teaches**, **takes**, **advisor** and **prereq**. Their meaning is the same as in the lectures.

1. First describe in English each of following relational algebra queries. Then write down the relation returned by each query for the tables in Figure 1.

   (a)

   $$\Pi_{\text{dept\_name}}(\text{instructor}$$
   $$-$$
   $$\Pi_{\text{ID,name,dept\_name,instructor.salary}}(\text{instructor}$$
   $$\bowtie_{\text{instructor.salary} > \text{R.salary}} \rho_R(\Pi_{\text{salary}}(\text{instructor}))))$$

   (5)

   > **Solution:** Find the departments which instructors with the least salary belong to.
   >
   > | dept_name |
   > |-----------|
   > | Music |

   (b)

   $$\Pi_{\text{student.ID,student.name}}(\text{student}$$
   $$\bowtie \text{takes}$$
   $$\bowtie \Pi_{\text{course\_id,sec\_id,semester,year}}(\sigma_{\text{name=\"Einstein\"}}(instructor \bowtie \text{teaches})))$$

   (5)

   > **Solution:** Find the IDs and names of students who have taken a course section taught by Einstein.
   >
   > | ID | name |
   > |-------|---------|
   > | 44553 | Peltier |

   (c)

   $$\Pi_{\text{ID,name}}(\text{instructor}$$
   $$\bowtie_{\text{ID=R.i\_ID}}$$
   $$\Pi_{\text{R.i\_ID}}(\rho_R(\text{advisor}) \bowtie_{\text{R.i\_ID = S.i\_ID} \land \text{R.s\_ID} \neq \text{S.s\_ID}} (\rho_S(\text{advisor}))))$$

(5)

> **Solution:** Find the IDs and names of instructors who are advisors for more than one student.
>
> | ID | name |
> |---|---|
> | 22222 | Einstein |
> | 98345 | Kim |
> | 45565 | Katz |

(d)

$$\Pi_{\text{ID,name}}(\text{student}$$
$$\bowtie_{\text{student.ID=takes.ID}} \text{takes}$$
$$\bowtie_{\text{takes.course\_ID=course.course\_ID}} (\sigma_{\text{dept\_name = "Biology"}}(\text{course})))$$

(5)

> **Solution:** Find the IDs and names of students who have taken at least one course offered by the Biology Department.
>
> | ID | name |
> |---|---|
> | 98988 | Tanaka |

2. First describe in English each of following SQL queries. Then write down the relation returned by each query for the tables in Figure 1.

(a)
```
select  dept_name
from  instructor
where  salary >   (select  avg(salary)
                    from  instructor
                    where  dept_name =  'Comp. Sci.')
```

(5)

> **Solution:** Find all departments which have at least one instructor with salary greater than the average salary in the CS department.

dept_name
Finance
Physics
Physics
Finance
Comp. Sci.
Elec. Eng.

(b)
```
select dept_name, count(distinct ID) as instr_count
from instructor natural join teaches
where semester='Spring' and year=2010
group by dept_name;
```

(5)

Solution: Find the number of instructors in each department who teach a course in the Spring 2010 semester.

dept_name	instr_count
Comp. Sci.	3
Finance	1
History	1
Music	1

(c)
```
select dept_name, count (distinct ID) from instructor
where salary > 50000
group by dept_name;
```

(5)

Solution: List the department names along with the number of faculty in the department who earn more than 50000.

dept_name	count(distinct ID)
Biology	1
Comp. Sci.	3
Elec. Eng.	1
Finance	2
History	2
Physics	2

(d)
```
select count (distinct ID)
from takes
where (course_id, sec_id, semester, year) in
              (select course_id, sec_id, semester, year
               from teaches
               where teaches.ID= 10101);
```

(5)

Solution: Find the total number of distinct students who have taken course sections taught by the instructor with ID 10101.

count(distinct ID)
6

3. Consider the following set F of functional dependencies for the relation schema $R = (A, B, C, D, E)$:

$$A \rightarrow BC$$
$$CD \rightarrow E$$
$$B \rightarrow D$$
$$E \rightarrow A$$

(a) List all the functional dependencies in F^+ of the form $\alpha \rightarrow \beta$, where α has a single attribute and $\alpha \cap \beta = \emptyset$. (You can write $X \rightarrow YZ$ instead of $X \rightarrow Y$ and $X \rightarrow Z$, and so on.)

(5)

> **Solution:** $A \to BCDE, B \to D, E \to ABCD$

(b) List all the candidate keys for R.

(5)

> **Solution:** A, E, BC, CD.

(c) Consider the decomposition of the above schema R into $R_1 = (A, B, C)$ and $R_2 = (A, D, E)$. Is this a lossless decomposition? Justify your answer.

(5)

> **Solution:** The decomposition is lossless since $R_1 \cap R_2 = A$, and $A \to R_1$ is in F^+.

(d) Consider the same decomposition of R as above into $R_1 = (A, B, C)$ and $R_2 = (A, D, E)$. Is this decomposition dependency-preserving? Justify your answer.

(5)

> **Solution:** The decomposition is *not* dependency preserving, since the dependency $CD \to E$ in F^+ can only be checked by checking both the relations. The same is true for the dependency $B \to D$ in F^+.

4. Consider the relation schema $R(A, B, C, D)$ with the set F of functional dependencies

$$AB \to C$$
$$C \to D$$
$$D \to A$$

(a) Find all the BCNF violations, if any, for the schema R by the functional dependencies in F^+.

(5)

> **Solution:** The set of superkeys for R with respect to F is $\{AB, BC, BD, ABC, BCD, ABCD\}$. The functional dependencies in F^+

that violate the BCNF condition for R are

$$C \to D$$
$$D \to A$$
$$C \to A$$
$$CD \to A$$
$$AC \to D$$
$$AC \to AD$$
$$CD \to AC$$
$$AC \to CD$$

Note that no trivial dependency (*i.e.*, $\alpha \to \beta$ with $\beta \subseteq \alpha$) can violate the BCNF condition.

(b) Decompose the schema R, if necessary, into a collection of schemas each of which is in BCNF. Write only the final answer. Do not show the intermediate steps.

(5)

Solution: The possible solutions are $R = (R_1, R_2, R_3)$ where the choices for R_1, R_2, R_3 are:

$$R_1 = (C, D), R_2 = (A, C), R_3 = (B, C)$$
$$R_1 = (A, D), R_2 = (C, D), R_3 = (B, C)$$

Note that a decomposition can be listed in any order. For this example, it can be shown that any schema in the decomposition cannot have more than two attributes.

Figure 1: Tables for Q1 and Q2

instructor

ID	name	dept_name	salary
10101	Srinivasan	Comp. Sci.	65000.00
12121	Wu	Finance	90000.00
15151	Mozart	Music	40000.00
22222	Einstein	Physics	95000.00
32343	El Said	History	60000.00
33456	Gold	Physics	87000.00
45565	Katz	Comp. Sci.	75000.00
58583	Califieri	History	62000.00
76543	Singh	Finance	80000.00
76766	Crick	Biology	72000.00
83821	Brandt	Comp. Sci.	92000.00
98345	Kim	Elec. Eng.	80000.00

student

ID	name	dept_name	tot_cred
00128	Zhang	Comp. Sci.	102
12345	Shankar	Comp. Sci.	32
19991	Brandt	History	80
23121	Chavez	Finance	110
44553	Peltier	Physics	56
45678	Levy	Physics	46
54321	Williams	Comp. Sci.	54
55739	Sanchez	Music	38
70557	Snow	Physics	0
76543	Brown	Comp. Sci.	58
76653	Aoi	Elec. Eng.	60
98765	Bourikas	Elec. Eng.	98
98988	Tanaka	Biology	120

department

dept_name	building	budget

_____	_____	_____
Biology	Watson	90000.00
Comp. Sci.	Taylor	100000.00
Elec. Eng.	Taylor	85000.00
Finance	Painter	120000.00
History	Painter	50000.00
Music	Packard	80000.00
Physics	Watson	70000.00

course

course_id	title	dept_name	credits
BIO−101	Intro. to Biology	Biology	4
BIO−301	Genetics	Biology	4
BIO−399	Computational Biology	Biology	3
CS−101	Intro. to Computer Science	Comp. Sci.	4
CS−190	Game Design	Comp. Sci.	4
CS−315	Robotics	Comp. Sci.	3
CS−319	Image Processing	Comp. Sci.	3
CS−347	Database System Concepts	Comp. Sci.	3
EE−181	Intro. to Digital Systems	Elec. Eng.	3
FIN−201	Investment Banking	Finance	3
HIS−351	World History	History	3
MU−199	Music Video Production	Music	3
PHY−101	Physical Principles	Physics	4

teaches

ID	course_id	sec_id	semester	year
10101	CS−101	1	Fall	2009
10101	CS−315	1	Spring	2010
10101	CS−347	1	Fall	2009
12121	FIN−201	1	Spring	2010
15151	MU−199	1	Spring	2010
22222	PHY−101	1	Fall	2009
32343	HIS−351	1	Spring	2010
45565	CS−101	1	Spring	2010
45565	CS−319	1	Spring	2010
76766	BIO−101	1	Summer	2009

76766	BIO−301	1	Summer	2010
83821	CS−190	1	Spring	2009
83821	CS−190	2	Spring	2009
83821	CS−319	2	Spring	2010
98345	EE−181	1	Spring	2009

takes

ID	course_id	sec_id	semester	year	grade
00128	CS−101	1	Fall	2009	A
00128	CS−347	1	Fall	2009	A−
12345	CS−101	1	Fall	2009	C
12345	CS−190	2	Spring	2009	A
12345	CS−315	1	Spring	2010	A
12345	CS−347	1	Fall	2009	A
19991	HIS−351	1	Spring	2010	B
23121	FIN−201	1	Spring	2010	C+
44553	PHY−101	1	Fall	2009	B−
45678	CS−101	1	Fall	2009	F
45678	CS−101	1	Spring	2010	B+
45678	CS−319	1	Spring	2010	B
54321	CS−101	1	Fall	2009	A−
54321	CS−190	2	Spring	2009	B+
55739	MU−199	1	Spring	2010	A−
76543	CS−101	1	Fall	2009	A
76543	CS−319	2	Spring	2010	A
76653	EE−181	1	Spring	2009	C
98765	CS−101	1	Fall	2009	C−
98765	CS−315	1	Spring	2010	B
98988	BIO−101	1	Summer	2009	A
98988	BIO−301	1	Summer	2010	NULL

advisor

s_ID	i_ID
00128	45565
12345	10101
23121	76543
44553	22222

45678 22222
76543 45565
76653 98345
98765 98345
98988 76766

prereq

course_id	prereq_id
BIO−301	BIO−101
BIO−399	BIO−101
CS−190	CS−101
CS−315	CS−101
CS−319	CS−101
CS−347	CS−101
EE−181	PHY−101