Database Management Systems

Vijaya Saradhi

IIT Guwahati

Thu, 05^{th} Mar 2020

Note

Multiple Solutions

- Questions about
  - Relational algebraic expressions
  - SQL query

  have multiple answers. Every alternate answer will be checked for its correctness
- Solutions provided here must be treated as model solutions
- Questions involving
  - True or False
  - Computing output for an SQL given relational instances
  - Parse trees for a given SQL query
  - Theoretical questions

  have unique solution and they all must be adhering to the solution listed here

Q01

Q01- Problem

Consider the following schema:

Suppliers (<u>sid int</u>, sname char(50), address char(50))

Parts (<u>pid int</u>, pname char(50), color char(50))

Catalog (<u>sid int, pid int</u>, cost float)

Key attributes are denoted using underline notation. Write the following queries in relational algebra

Q01 - (a)

Find the `sids` of suppliers who supply some red part or at 221 Packer Ave.

$\rho(R1, \pi_{sid}(\sigma_{color='red'}(Parts) \bowtie Catalog))$

$\rho(R2, \pi_{sid}(\sigma_{address='221 PackerStreet'}(Suppliers) \bowtie Catalog))$

$R1 \cup R2$

Q01

Q01- Problem

Consider the following schema:

Suppliers (<u>sid int</u>, sname char(50), address char(50))

Parts (<u>pid int</u>, pname char(50), color char(50))

Catalog (<u>sid int, pid int</u>, cost float)

Key attributes are denoted using underline notation. Write the following queries in relational algebra

Q01 - (b)

Find the sids of suppliers who supply some red part and some green part.

$\rho(R1, \pi_{sid}(\sigma_{color='red'}(Parts) \bowtie Catalog))$

$\rho(R2, \pi_{sid}(\sigma_{color='green'}(Parts) \bowtie Catalog))$

$R1 \cap R2$

Q01

Q01- Problem

Consider the following schema:

Suppliers (<u>`sid int`</u>, `sname char(50)`, `address char(50)`)

Parts (<u>`pid int`</u>, `pname char(50)`, `color char(50)`)

Catalog (<u>`sid int`, `pid int`</u>, `cost float`)

Key attributes are denoted using underline notation. Write the following queries in relational algebra

Q01 - (c)

Find the `sid`s of suppliers who supply every part.

$\pi_{sid,pid}(Catalog)/\pi_{pid}(Parts)$

Q01

Q01- Problem

Consider the following schema:

Suppliers (<u>`sid int`</u>, `sname char(50)`, `address char(50)`)

Parts (<u>`pid int`</u>, `pname char(50)`, `color char(50)`)

Catalog (<u>`sid int`, `pid int`</u>, `cost float`)

Key attributes are denoted using underline notation. Write the following queries in relational algebra

Q01 - (d)

Find the `pids` of parts supplied by at least two different suppliers.

$\rho(R1, Catalog)$

$\rho(R2, Catalog)$

$\pi_{R1.pid}(\sigma_{R1.pid=R2.pid\ AND\ R1.sid \neq R2.sid}(R1 \times R2))$

Q02

Q02 - Problem

Consider the following schema:

Sailors (<u>sid int</u>, sname char(50), rating int, age float)

Reserves (<u>sid int, bid int, rday char(10)</u>)

Boats (<u>bid int</u>, bname char(50), color char(50))

rating take values from the set {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}. Write an SQL query to find the name and age of the oldest sailor for every rating value.

Q02- Solution

```
SELECT   sname, age
FROM     Sailors
WHERE    (rating, age)
IN       (SELECT     rating, MAX(age) AS ag
          FROM       Sailors
          GROUP BY   rating)
ORDER BY rating;
```

Q02

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

```
SELECT  rating, MAX(age) AS ag
    FROM        Sailors
    GROUP BY    rating
```

Q02

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (7, 45.0) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (1, 33.0) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (8, 55.5) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (8, 25.5) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5), (Rusty, 35.0) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (10, 35.0) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5), (Rusty, 35.0) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (7, 35.0) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5), (Rusty, 35.0) }

Sailors

sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (10, 16.0) member of the relation?

Sailors

rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5), (Rusty, 35.0), (Horatio, 35.0) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (9, 35.0) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5), (Rusty, 35.0), (Horatio, 35.0) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (3, 25.5) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q02

Result: {(Dustin, 45.0), (Brutus, 33.0), (Lubber, 55.5), (Rusty, 35.0), (Horatio, 35.0), (Bob, 63.5) }

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

is (3, 63.5) member of the relation?

Sailors	
rating	age
7	45.0
1	33.0
8	55.5
10	35.0
9	35.0
3	63.5

Q03

Q03 - Problem

For the schema given in the Question 2 write SQL statements in four distinct ways for the following query `find the names of sailors who have reserved boat number 103`

Q03(a) - Problem

Using `cross product` operation in SELECT statement

```
SELECT  S.sname
FROM    Sailors AS S, Reserves AS R
WHERE   (S.sid = R.sid AND R.bid = 103)
```

Q03

Q03 - Problem

For the schema given in the Question 2 write SQL statements in four distinct ways for the following query `find the names of sailors who have reserved boat number 103`

Q03(b) - Problem

Using natural join operation in SELECT statement

```
SELECT  sname
FROM    Sailors
JOIN    Reserves
ON      Sailors.sid = Reserves.sid
WHERE   (Reserves.bid = 103)
```

Q03

Q03 - Problem

For the schema given in the Question 2 write SQL statements in four distinct ways for the following query `find the names of sailors who have reserved boat number 103`

Q03(c) - Problem

Using nested SELECT statements using IN

```
SELECT   S.sname
FROM     Sailors AS S
WHERE    S.sid
IN       (SELECT     R.sid
          FROM       Reserves AS R
          WHERE      R.bid = 103)
```

Q03

Q03 - Problem

For the schema given in the Question 2 write SQL statements in four distinct ways for the following query `find the names of sailors who have reserved boat number 103`

Q03(d) - Problem

Using correlated nested SELECT statements

```
SELECT  S1.sname
FROM    Sailors AS S1
WHERE   EXISTS
    (SELECT *
        FROM    Reserves AS R1
        WHERE   R1.bid = 103
        AND     S1.sid = R1.sid
    )
```

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

What is the output of the following query?

```
SELECT ta.player
FROM   top_scorer AS ta
WHERE  ta.goals >ALL (
          SELECT  tb.goals
          FROM    top_scorer AS tb
          WHERE   tb.country='Spain')
AND ta.goals >ANY (
          SELECT  tc.goals
          FROM    top_scorer AS tc
          WHERE   tc.country='Germany')
```

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First subquery `SELECT tb.goals FROM top_scorer AS tb WHERE tb.country='Spain'` result EMPTY set { }
- Second subquery `SELECT tc.goals FROM top_scorer AS tc WHERE tc.country='Germany'`
  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $16 >$ ALL $\{\}$?
- Second test: $16 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $15 >$ ALL $\{\}$?
- Second test: $15 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $14 >$ ALL $\{\}$?
- Second test: $14 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)

- Result: {Klose, Ronaldo, Gmuller}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $13 >$ ALL $\{\}$?
- Second test: $13 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo, Gmuller, Fontaine}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $12 >$ ALL $\{\}$?
- Second test: $12 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $11 >$ ALL $\{\}$?
- Second test: $11 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)

- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: 11 > ALL {} ?
- Second test: 11 > ANY {16, 14, 11, 10, 10}

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $10 >$ ALL $\{\}$?
- Second test: $10 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)

- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $10 >$ ALL $\{\}$?
- Second test: $10 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)

- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $10 >$ ALL $\{\}$?
- Second test: $10 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $10 >$ ALL $\{\}$?
- Second test: $10 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $10 >$ ALL $\{\}$?
- Second test: $10 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)
- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q04

Q04 - Problem

Consider the following database table named top_scorer

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
Gmuller	Germany	14
Fontaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
Tmuller	Germany	10
Rahn	Germany	10

- First test: $10 >$ ALL $\{\}$?
- Second test: $10 >$ ANY $\{16, 14, 11, 10, 10\}$

  (Klose, Germany, 16)
  (Gmuller, Germany, 14)
  (Klinsmann, Germany, 11)
  (Tmuller, Germany, 10)
  (Rahn, Germany, 10)

- Result: {Klose, Ronaldo, Gmuller, Fontaine, Pele, Klinsmann, Kocsis}

Q05

Q05- Problem

Consider the following schema:

Movie (title, year, length, inColor, studioName, producerC)

MovieExec (name, address, cert, networth)

Using this schema, we have created a view named MovieProd given below

```
CREATE VIEW MovieProd AS
    (SELECT title , name
    FROM     Movie , MovieExec
    WHERE    producerC = cert)
```

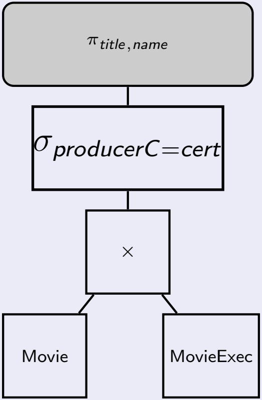
And a query associated with the view Find the producer of 'Gone With the Wind' movie given as

```
SELECT   name
FROM     MovieProd
WHERE    title = 'Gone With the Wind';
```

Q05

Q05- Problem

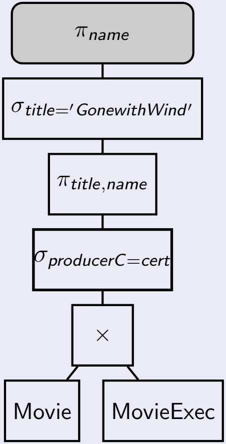
Draw the parse tree (or expression tree) for the view `MovieProd`



Q05

Q05- Problem

Draw the parse tree (or expression tree) for the query involving the `MovieProd`. This parse tree must refer to base tables `Movie` and `MovieExec`.



Q06

Q06(a) - Problem

Assume that you are given a relation R with attributes ABCD. Assume that no record has NULL values. Write an SQL query that checks whether the functional dependency A $\rightarrow$ B holds.

Q06(a)- Solution

```
CHECK(
    (SELECT count(distinct A, B) FROM  R) < 2;
)
```

Q06

Q06(b) - Problem

Assume that you are given a relation R with attributes ABCD. Assume that no record has NULL values. Write an SQL assertion that enforces the functional dependency $A \rightarrow B$.

Q06(b)- Solution

```sql
CREATE ASSERTION ABFD
(
    CHECK(
        (SELECT count(distinct A, B) FROM   R) < 2;
    )
)
```

Q07

Q07(a) - Problem

Suppose you are given a relation R with four attributes ABCD. Given the set of FDs $\{C \rightarrow D, C \rightarrow A, B \rightarrow C\}$

Q07(a)- Solution

B

Q07

Q07(b) - Problem

Does R satisfies 1NF?

Q07(b)- Solution

Yes

Q07

Q07(c) - Problem

Does R satisfies 2NF?

Q07(c)- Solution

Yes

Q07

Q07(d) - Problem

Does R satisfies 3NF?

Q07(d)- Solution

No

Q07

Q07(e) - Problem

Does R satisfies BCNF?

Q07(e)- Solution

No

Q08 - Problem

FDs

Decompose the relation R(`class, section, student, major, exam, year, instructor, rank, salary, text, day, room`) given the set of functional dependencies

FD #	**X**	$\rightarrow$	**Y**
F1	{class, section}	$\rightarrow$	instructor
F2	{class, section, day}	$\rightarrow$	room
F3	student	$\rightarrow$	{major, year}
F4	instructor	$\rightarrow$	{rank, salary}

MVDs

MVD #	**X**	$\rightarrow\rightarrow$	**Y**
M1	{class, section}	$\rightarrow\rightarrow$	{student, major, exam, year}
M2	{class, section}	$\rightarrow\rightarrow$	{instructor, rank, salary}
M3	{class, section}	$\rightarrow\rightarrow$	text
M4	{class, section}	$\rightarrow\rightarrow$	{day, room}
M5	{class, section, student}	$\rightarrow\rightarrow$	exam
M6	class	$\rightarrow\rightarrow$	text

Q08 - Solution

Solution

R_{11}(class, section, student, exam)
R_{121}(student, major, year)
R_{211}(instructor, rank, salary)
R_{212}(class, section, instructor)
R_{221}(class, text)
R_{222}(class, section, day, room)

Q09 - Problem

Sailors			
sid	sname	rating	age
22	Dustin	7	45.0
29	Brutus	1	33.0
31	Lubber	8	55.5
32	Andy	8	25.5
58	Rusty	10	35.0
64	Horatio	7	35.0
71	Zorba	10	16.0
74	Horatio	9	35.0
85	Art	3	25.5
95	Bob	3	63.5

Reserves		
sid	bid	day
22	101	10-Oct-2019
22	102	10-Oct-2019
22	103	08-Oct-2019
22	104	07-Oct-2019
31	102	10-Nov-2019
31	103	06-Nov-2019
31	104	12-Nov-2019
64	101	05-Sep-2019
64	102	08-Sep-2019
74	103	08-Sep-2019

What is the output the following SQL query:

```sql
SELECT              *
FROM                Sailors
LEFT OUTER JOIN     Reserves
ON                  Sailors.sid = Reserves.sid
WHERE               ISNULL(bid);
```

Q09 - Problem

Query

What is the output the following SQL query:

```
SELECT           *
FROM             Sailors
LEFT OUTER JOIN  Reserves
ON               Sailors.sid = Reserves.sid
WHERE            ISNULL(bid);
```

Q09 - Output

sid	sname	rating	age	sid	bid	day
29	Brutus	1	33	NULL	NULL	NULL
32	Andy	8	25.5	NULL	NULL	NULL
58	Rusty	10	35	NULL	NULL	NULL
71	Zorba	10	16	NULL	NULL	NULL
85	Art	3	25.5	NULL	NULL	NULL
95	Bob	3	63.5	NULL	NULL	NULL

Decomposition - Lossy/Loss-less?

Introduction

- Let R be a relation
- Let R be decomposed into two relations R_1 and R_2
- Let n be the number of tuples in R
- Decomposition should be performed in such a way that complete R with n tuples can be recovered using R_1 and R_2
- If we can recover original relation we say the decomposition is lossless
- Otherwise the decomposition is lossy

Decomposition - Lossy/Loss-less?

Example - 01

R		
X	Y	Z
x_1	y_1	z_1
x_2	y_2	z_2
x_3	y_2	z_3
x_4	y_3	z_4

R_1		R_2	
X	Y	Y	Z
x_1	y_1	y_1	z_1
x_2	y_2	y_2	z_2
x_3	y_2	y_2	z_3
x_4	y_3	y_3	z_4

Decomposition - Lossy/Loss-less?

Example - 01

$R_1 \bowtie R_2$		
X	Y	Z
x_1	y_1	z_1
x_2	y_2	z_2
x_2	y_2	z_3
x_3	y_2	z_2
x_3	y_2	z_3
x_4	y_3	z_4

Decomposition - Lossy/Loss-less?

Example - 02

	R	
A	B	C
a	b	c

Example - 02

- Let the FD B $\rightarrow$ C exists on R;
- Let R be decomposed into R_1(A, B) and R_2(B, C)
- The relatoin is not in BCNF ({B} is not the super key)
- If there is another FD: A $\rightarrow$ B then there is transitive dependency and {A} will be the key
- If no other FD exists, then {A, B} would be the key
- B is still not the superkey!

Decomposition - Lossy/Loss-less?

Example - 02

R_1		R_2	
A	B	B	C
a	b	b	c

Decomposition - Lossy/Loss-less?

Example - 02

R_1		R_2	
A	B	B	C
a	b	b	c

- $R_1 \bowtie R_2$ will yield original R
- The decomposition is lossless

Decomposition - Lossy/Loss-less?

Example - 03

If R contains the following tuples

R		
A	B	C
a	b	c
a	b	e

Example - 03

R_1		R_2	
A	B	B	C
a	b	b	c
a	b	b	e

Decomposition - Lossy/Loss-less?

Example - 03

If R contains the following tuples

R		
A	B	C
a	b	c
a	b	e

Example - 03

$R_1 \bowtie R_2$		
A	B	C
a	b	c
a	b	e
a	b	c
a	b	e

Decomposition - Lossy/Loss-less?

Example - 03

If R contains the following tuples

A	B	C
a	b	c
a	b	e

R

Example - 03

- However, R cannot contain the tuple (a, b, e)
- As the FD: B $\rightarrow$ C is in place
- That is c $=$ e
- When relations are decomposed according to FDs, then original relation can be recovered

Functional Dependency

Definition

Let $\mathbf{R} = \{R_1, R_2, R_3, \cdots R_p\}$ be a set of relation schemas over **U**. A relation r(**U**) satisfies join dependency $*[R_1, R_2, \cdots, R_p]$ if r decomposes losslessly onto $R_1, R_2, \cdots, R_p$

$$r = \pi_{R_1}(r) \bowtie \pi_{R_2}(r) \bowtie \cdots \bowtie \pi_{R_p}(r)$$

Functional Dependency

Example

R		
A	B	C
a_1	b_1	c_1
a_1	b_2	c_2
a_3	b_3	c_3
a_4	b_3	c_4
a_5	b_5	c_5
a_6	b_6	c_5

R $= R_1 \bowtie R_2 \bowtie R_3$; *[AB, AC, BC]

R_1	
A	B
a_1	b_1
a_1	b_2
a_3	b_3
a_4	b_3
a_5	b_5
a_6	b_6

R_2	
A	C
a_1	c_1
a_1	c_2
a_3	c_3
a_4	c_4
a_5	c_5
a_6	c_5

R_3	
B	C
b_1	c_1
b_2	c_2
b_3	c_3
b_3	c_4
b_5	c_5
b_6	c_5

Trivial FD

Definition

A JD $*[R_1, R_2, \cdots, R_p]$ over R is trivial if it is satisfied by every relation r(R)

Project-Join Normal Form (PJNF)

Definition (5NF)

Let R be a relation scheme and let F be a set of FDs and JDs over R. R is in PJNF if every JD is trivial or R_i is a superkey for R.

Project-Join Normal Form (PJNF)

Example

- Let F = {*[ABCD, CDE, BDI], *[AB, BCD, AD], A $\rightarrow$ BCDE, BC $\rightarrow$ AI }
- R = A B C D E I
- R is not in PJNF with respect to F because of *[ABCD, CDE, BDI]
- Let $R_1 = ABCD$; $R_2 = CDE$; and $R_3 = BDI$
- The JD *[AB, BCD, AD]: each set of attributes is a superkey for R_1 due to FDs {A $\rightarrow$ BCDE, BC $\rightarrow$ AI }
- The FDs are either trivial or have keys as left sides

Preserving FDs

Introduction

- For a relation r to be recoverable from its projects its decomposition must be lossless
- In addition, decomposition should satisfy dependency preservation
- That is the decompositions satisfy all the FDs that are satisfied by the original relation
- Any decomposition that does not preserve the dependencies of the original relations imposes burden on RDBMS

Preserving FDs

Example

Let $r(X, Y, Z)$ satisfies FDs: $\{$XY $\rightarrow$ Z, Z $\rightarrow$ X$\}$. Let r(X, Y, Z) be decomposed into $R_1(YZ)$ and $R_2(ZX)$.

Preserving FDs

Introduction

R_1	
Y	Z
y_1	z_1
y_1	z_2

R_2	
Z	X
z_1	x_1
z_2	x_1

Join

$R_1 \bowtie R_2$		
X	Y	Z
x_1	y_1	z_1
x_1	y_1	z_2

R_2 satisfies Z $\rightarrow$ X; but $R_1 \bowtie R_2$ does not satisfy XY $\rightarrow$ Z

Preserving FDs

Central Idea

- In addition to lossless decomposition, dependency preserving property must be satisified by decomposition
- The decmpositions statisfy all the FDs that are satisfied by the original relation
- Reason FDs satisified by a relation define integrity constraints that the relation needs to meet

Preserving FDs

Example

- Let R(X, Y, Z) satisfies the FDs: $\{XY \rightarrow Z, Z \rightarrow X\}$
- Let R be decomposed into two relations R_1(Y, Z) and R_2(Z, X)
- R_1 and R_2 are lossless decompositions; that is $R = R_1 \bowtie R_2$
- However, $R_1 \bowtie R_2$ do not preserve the FD XY $\rightarrow$ Z

R_1

Y	Z
y_1	z_1
y_1	z_2

R_2

Z	X
z_1	x_1
z_2	x_1

Projection of set of Dependencies

Onto set of attributes & algorithm

- Let r(R) has been recomposed into $(R_1, R_2, \cdots, R_k)$
- Let F be the set of FDs satisfied by r
- Define projection of F onto Z $\pi_Z(F)$ as

$$\pi_Z(F) = \{X \to Y \subset F | XY \subset Z\}$$

-

Projection of set of Dependencies

Onto set of attributes & algorithm

- To compute $\pi_Z(F)$, consider the proper subsets X of Z that appear as determinant of an FD
- Do
  - Calcluate X^+
  - For P $\in X^+$ that satisfy
    - P $\subset$ Z
    - P $\subset X^+$
    - P $\not\subset$ X

  include X $\rightarrow$ P in $\pi_Z(F)$

Projection of set of Dependencies

Example

- Let r(X, Y, W, Z, Q) and F $= \{X \rightarrow Z, Y \rightarrow Q, ZQ \rightarrow W\}$
- Compute: $\pi_{\{X,Y,W\}}(F)$. That is Z $= \{X, Y, W\}$

Projection of set of Dependencies

Example

- {X, Y} are subsets of {X, Y, W}
- Both {X, Y} appear in LHS of F
- Proper subsets of {X, Y} are: {X}, {Y}, {X, Y}
- Consider {X}; Compute $\{X\}^+$
- $X^+ = \{X, Z\}$
- For each P in X^+ do
    - is P $\subset$ {X, Y, W}? Yes
    - is P $\subset$ $X^+ =$ {X ,Y}; Yes
    - is P $\not\subset$ X

Projection of set of Dependencies

Example

- is (P $=$) X $\subset$ {X, Y, W}; Yes
- is (P $=$) X $\subset$ $X^+ =$ {X, Y}
- is (P $=$) X $\not\subset$ X No
- Do not include X $\rightarrow$ X in $\pi_{\{X,Y,W\}}(F)$

Projection of set of Dependencies

Example

- is $(P =)Z \subset \{X, Y, W\}$; No
- Do not include $X \rightarrow Z$ in $\pi_{\{X,Y,W\}}(F)$

Projection of set of Dependencies

Example

- is Y $\subset$ {X, Y, W}; Yes
- is Y $\subset$ $Y^+ =$ {Y, Q}
- is Y $\not\subset X$ No
- Do not include X $\rightarrow$ X in $\pi_{\{X,Y,W\}}(F)$

Projection of set of Dependencies

Example

- {X, Y} are subsets of {X, Y, W}
- Both {X, Y} appear in LHS of F
- Proper subsets of {X, Y} are: {X}, {Y}, {X, Y}
- Consider each of the above three (i) say X = Y ; Compute Y^+
- $Y^+ = \{Y, Q\}$
- For each P in Y^+ do
  - is P $\subset$ {X, Y, W}? Yes
  - is P $\subset$ Y^+ = {Y, Q};
  - is P $\not\subset$ X

is P $\subset$ {X, Y, W}?

is Y $\subset$ {X, Y, W}; Yes

Projection of set of Dependencies

Example

- {X, Y} are subsets of {X, Y, W}
- Both {X, Y} appear in LHS of F
- Proper subsets of {X, Y} are: {X}, {Y}, {X, Y}
- Consider each of the above three (i) say X = Y ; Compute Y^+
- $Y^+ = \{Y, Q\}$
- For each P in Y^+ do
  - is P $\subset$ {X, Y, W}? Yes
  - is P $\subset$ Y^+ = {Y, Q};
  - is P $\not\subset$ X

is P $\subset$ {X, Y, W}?

is Q $\subset$ {X, Y, W}; No; Do not include X $\rightarrow$ Y in $\pi_{\{X,Y,W\}}$(F)

Projection of set of Dependencies

Example

- {X, Y} are subsets of {X, Y, W}
- Both {X, Y} appear in LHS of F
- Proper subsets of {X, Y} are: {X}, {Y}, {X, Y}
- Now consider {X, Y} Y ; Compute $\{X, Y\}^+$
- $\{X, Y\}^+ = \{X, Y, Z, Q, W\}$
- For each P in $\{X, Y\}^+$ do
  - is P $\subset$ {X, Y, W}? Yes
  - is P $\subset$ $Y^+ =$ {Y, Q};
  - is P $\not\subset$ X

is P $\subset$ {X, Y, W}?

excluding {X, Y} $\rightarrow$ X and
{X, Y} $\rightarrow$ Y we have

Dependency Preservation Testing

Method

- Given r(R), a decomposition $(R_1, R_2, \cdots, R_k)$
- A set F of FDs satisfied by r(R)
- Compute $\pi_{R_i}(F)$
- G $= \cup_{i=1}^{k} \pi_{R_i}(F)$
- Is G = F? if yes, then $(R_1, R_2, \cdots, R_k)$ is dependency preserving decomposition