Practice Problem Set #2 Solution

CSC 261/461 (Database Systems), Spring 2017,
University of Rochester
02/04/2017

Problem 1: Null Values and Aggregate Functions

t1	
GROUP-KEY	VAL
Group-1	(null)
Group-1	(null)
Group-2	a
Group-2	a
Group-2	z
Group-2	z
Group-2	(null)
Group-3	A
Group-3	A
Group-3	Z

For table t1 above, what's the result of the following query?

```
select
group-key ,
MAX( VAL )            max-val ,
MIN( VAL )            min-val ,
COUNT( * )            count-all-rows ,
COUNT( VAL )          count-val ,
COUNT( DISTINCT VAL ) count-distinct-val
from t1
group by group-key
order by group-key ;
```

Answer:

aggregate functions like MAX , MIN , and COUNT will return values that for the most part ignore nulls, like these.

GROUP-KEY	MAX-VAL	MIN-VAL	COUNT-ALL-ROWS	COUNT-VAL	COUNT-DISTINCT-VAL
Group-1	(null)	(null)	2	0	0
Group-2	z	a	5	4	2
Group-3	Z	A	3	3	2

Note how MAX-VAL contains the same results for Group-2 and Group-3, even though Group-2 contains null VAL values and Group-3 does not. Note also that only COUNT-ALL-ROWS returned a count that included null values. The other two versions of COUNT() ignored null values.

source: http://www.sqlsnippets.com/en/topic-12656.html

Problem 2: Join

Company		
cname	stockprice	country
GizmoWorks	25	USA
Canon	65	Japan
Hitachi	15	Japan
NULL	64	Japan

Product			
pname	price	category	manufacturer
Gizmo	19.99	Gadgets	GizmoWorks
MultiTouch	203.99	Household	Hitachi
OneTouch	203.99	Household	**NULL**
PowerGizmo	29.99	Gadgets	GizmoWorks
SingleTouch	149.99	Photography	Canon

For the relations Company and Product, what do the following queries output?

1. select pname, country from product **join** company on manufacturer = cname;

```
+-------------+---------+
| pname       | country |
+-------------+---------+
| Gizmo       | USA     |
| MultiTouch  | Japan   |
| PowerGizmo  | USA     |
| SingleTouch | Japan   |
+-------------+---------+
```

2. select pname, country from product **left outer join** company on manufacturer = cname;

```
+-------------+---------+
| pname       | country |
+-------------+---------+
| Gizmo       | USA     |
| MultiTouch  | Japan   |
| OneTouch    | NULL    |
| PowerGizmo  | USA     |
| SingleTouch | Japan   |
+-------------+---------+
```

3. select pname, country from product **right outer join** company on manufacturer = cname;

```
+-------------+---------+
| pname       | country |
+-------------+---------+
| Gizmo       | USA     |
| PowerGizmo  | USA     |
| SingleTouch | Japan   |
| MultiTouch  | Japan   |
| NULL        | Japan   |
+-------------+---------+
```

Problem 3: Database concepts

1. What is the difference between a key and a superkey?

Answer:

A Key is a minimal Candidate Key, which is to say all constituent columns are strictly required in order to ensure uniqueness.

A Super Key is simply a non-minimal Candidate Key, that is to say one with additional columns not strictly required to ensure uniqueness of the row.

2. Refer Figure 1

   Suppose each of the following Update operations is applied directly to the database of Figure 1. Discuss all integrity constraints violated by each operation, if any, and the different ways of enforcing these constraints:

   (a) Insert < 'Robert', 'F', 'Scott', '943775543', '21-JUN-42', '2365 Newcastle Rd, Bellaire, TX', M, 58000, '888665555', 1 > into EMPLOYEE.

   (b) Insert < 'ProductA', 4, 'Bellaire', 2 > into PROJECT.

   (c) Insert < 'Production', 4, '943775543', '01-OCT-88' > into DEPARTMENT.

   (d) Insert < '677678989', null, '40.0' > into WORKS_ON.

   (e) Insert < '453453453', 'John', M, '12-DEC-60', 'SPOUSE' > into DEPENDENT.

   (f) Delete the WORKS_ON tuples with ESSN= '333445555'.

   (g) Delete the EMPLOYEE tuple with SSN= '987654321'.

   (h) Delete the PROJECT tuple with PNAME= 'ProductX'.

   (i) Modify the MGRSSN and MGRSTARTDATE of the DEPARTMENT tuple with DNUMBER=5 to '123456789' and '01-OCT-88', respectively.

   (j) Modify the SUPERSSN attribute of the EMPLOYEE tuple with SSN= '999887777' to '943775543'.

   (k) Modify the HOURS attribute of the WORKS_ON tuple with ESSN= '999887777' and PNO= 10 to '5.0'.

Answer:

(a) No constraint violations.

(b) Violates referential integrity because DNUM=2 and there is no tuple in the DEPARTMENT relation with DNUMBER=2. We may enforce the constraint by: (i) rejecting the insertion of the new PROJECT tuple, (ii) changing the value of DNUM in the new PROJECT tuple to an existing DNUMBER value in the DEPARTMENT relation, or (iii) inserting a new DEPARTMENT tuple with DNUMBER=2.

(c) Violates both the key constraint and referential integrity. Violates the key constraint because there already exists a DEPARTMENT tuple with DNUMBER=4. We may enforce this constraint by: (i) rejecting the insertion, or (ii) changing the value of DNUMBER in the new DEPARTMENT tuple to a value that does not violate the key constraint. Violates referential integrity because MGRSSN='943775543' and there is no tuple in the EMPLOYEE relation with SSN='943775543'. We may enforce the constraint by: (i) rejecting the insertion, (ii) changing the value of MGRSSN to an existing SSN value in EMPLOYEE, or (iii) inserting a new EMPLOYEE tuple with SSN='943775543'.

(d) Violates both the entity integrity and referential integrity. Violates entity integrity because PNO, which is part of the primary key of WORKS_ON, is null. We may enforce this constraint by: (i) rejecting the insertion, or (ii) changing the value of PNO in the new WORKS_ON tuple to a value of PNUMBER that exists in the PROJECT relation. Violates referential integrity because ESSN='677678989' and there is no tuple in the EMPLOYEE relation with SSN='677678989'. We may enforce the constraint by: (i) rejecting the insertion, (ii) changing the value of ESSN to an existing SSN value in EMPLOYEE, or (iii) inserting a new EMPLOYEE tuple with SSN='677678989'.

(e) No constraint violations.

(f) No constraint violations.

(g) Violates referential integrity because several tuples exist in the WORKS_ON, DEPENDENT, DEPARTMENT, and EMPLOYEE relations that reference the tuple being deleted from EMPLOYEE. We may enforce the constraint by: (i) rejecting the deletion, or (ii) deleting all tuples in the WORKS_ON, DEPENDENT, DEPARTMENT, and EMPLOYEE relations whose values for ESSN, ESSN, MGRSSN, and SUPERSSN, respectively, is equal to'987654321'.

(h) Violates referential integrity because two tuples exist in the WORKS_ON relations that reference the tuple being deleted from PROJECT. We may enforce the constraint by: (i) rejecting the deletion, or (ii) deleting the tuples in the WORKS_ON relation whose value for PNO=1 (the value for the primary key PNUMBER for the tuple being deleted from PROJECT).

(i) No constraint violations.

(j) Violates referential integrity because the new value of SUPERSSN='943775543' and there is no tuple in the EMPLOYEE relation with SSN='943775543'. We may enforce the constraint by: (i) rejecting the deletion, or (ii) inserting a new EMPLOYEE tuple with SSN='943775543'.

(k) No constraint violations.

3. Consider the following relations for a database that keeps track of business trips of salespersons in a sales office:

```
SALESPERSON (SSN, Name, Start_Year, Dept_No)
TRIP (SSN, From_City, To_City, Departure_Date, Return_Date, Trip_ID)
EXPENSE (Trip_ID, Account#, Amount)
```

Specify the foreign keys for this schema, stating any assumptions you make.

Answer:

The schema of this question has the following two foreign keys: 1. the attribute SSN of relation TRIP that references relation SALESPERSON, and 2. the attribute Trip_ID of relation EXPENSE that references relation TRIP. In addition, the attributes Dept_No of relation SALESPERSON and Account# of relation EXPENSE are probably also foreign keys referencing other relations of the database not mentioned in the question. We now give the queries in relational algebra:

4. Consider the EMPLOYEE table has a constraint EMPSUPERFK as follows: CONSTRAINT EMPSUPERFK FOREIGN KEY (SUPERSSN) REFERNCES EMPLOYEE(SSN) ON DELETE CASCADE ON UPDATE CASCADE, Answer the following questions: a. What happens when the following command is run on the database state shown in Figure 5.6? DELETE EMPLOYEE WHERE LNAME = 'Borg' b. Is it better to CASCADE or SET NULL in case of EMPSUPERFK constraint ON DELETE?

Answer:

a) The James E. Borg entry is deleted from the table, and each employee with him as a supervisor is also (and their supervisees, and so on). In total, 8 rows are deleted and the table is empty. b) It is better to SET NULL, since an employee is not fired (DELETED) when their supervisor is deleted. Instead, their SUPERSSN should be SET NULL so that they can later get a new supervisor.

Note: RESTRICT causes the attempted DELETE of a parent row to fail. CASCADE will propagate the change when the parent changes. (If you delete a row, rows in constrained tables that reference that row will also be deleted, etc.) SET NULL sets the column value to NULL when a parent row goes away.

Problem 4: SQL Query

1. For each department whose average employee salary is more than $30,000, retrieve the department name and the number of male employees working for that department.

Answer:

```
SELECT DNAME, DNUMBER, COUNT (*)
FROM DEPARTMENT, EMPLOYEE
WHERE DNUMBER=DNO AND SEX='M' AND DNO IN ( SELECT DNO
FROM EMPLOYEE
GROUP BY DNO
HAVING AVG (SALARY) > 30000 )
GROUP BY DNAME

Result:
DNAME DNUMBER COUNT(*)
Research 5 3
Administration 4 1
Headquarters 1 1
```

2. Please read chapter 7 and run sample queries on betaweb. Make sure you read and practice problems on:

   - view
   - trigger
   - create vs insert
   - delete vs drop
   - update vs alter

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

Figure 1: Company Database