School: Computer Science
Institution: University of Windsor
Term: Winter 2021
Course: Comp-3150-1 : Database Management Systems
Instructor: Dr. C. I. Ezeife
Assignment #2 Solution: Total: 50 marks
Handed Out:Thurs. Jan. 28, 2021; Due: Thurs. Feb. 25, 2021

Objective of Assignment: To test on knowledge and design of relational model constraints, relational database schemas, functional dependencies and normalization of relational databases.

Scope: Assignment covers materials from Chapters 5 and 14 of book discussed in class.

Electronic Assignment Submission: Done through http://blackboard.uwindsor.ca

Marking Sheme : The mark for each of the questions is indicated beside each question.

Academic Integrity Statement: Remember to submit only work that is yours and include the following confidentiality agreement and statement at the beginning of your assignment.

CONFIDENTIALITY AGREEMENT & STATEMENT OF HONESTY

I confirm that I will keep the content of this assignment/examination confidential.
I confirm that I have not received any unauthorized assistance in preparing for or doing this assignment/examination. I confirm knowing that a mark of 0 may be assigned for copied work.

_____ _____
Student Signature Student Name (please print)

_____ _____
Student I.D. Number Date

Marking Scheme : The mark for each question and sub question is shown with the question below. Place your solutions in tables where possible.

For office Use only

Question	Mark
1	/20
2	/10
3	/10
4	/10
Total	/50

CHAPTER 5: THE RELATIONAL DATA MODEL AND RELATIONAL DATABASE CONSTRAINTS

1. (total marks 20) Given the same simple SalesRep-Worksin-SalesArea database schema you worked on in assignment 1, which contains three files described as follows:

   SalesRep (SRid: integer, SRname: string, SRage: integer, salary: real)
   Worksin (SRid: integer, Arid: integer, hours: integer)
   SalesArea (Arid: integer, Aname: string, city: string, budget: real, managerid: integer)

   Note : SRid, SRname, SRage, salary are the sales rep id, name, age and salary respectively. Also, hours is the number of hours worked by sales rep in the sales area. The rest of the attributes Arid, Aname, city, budget and managerid are the area id, name, location, budget and managerid respectively. A manager is a sales rep.

   Assume that an update operation (a general term in this chapter, for an insert, a modify or a delete operation) is to be made to this database to enter information about a new Salesrep who is now the new manager of an existing SalesArea (replacing the old manager) and although he is new, he has worked some hours in another existing SalesArea. Answer the following questions on what specific relations, attributes and operations (eg. insert, modify, delete) that need to be done for this update to be implemented in the entire database. This is not SQL query yet.

   Provide your answers both in descriptive sentence and using the formal (not SQL) database operations of INSERT, MODIFY, DELETE as used in Chapter 5 of book with specific attributes and relations when possible. An example formal insert of a Salesrep record into the Salesrep table is:
   INSERT < SRidr, SRname, SRage, salary> into SalesRep // for new SalesRep record
   And an example descriptive sentence of the formal operation above is:
   do an insert operation for a new SalesRep record into the SalesRep table.

   (a) Give the set of needed insert, modify or delete operations for this update. 5 marks

   (b) What types of constraints (explain with data) would you expect to check for this update to be done? 5 marks

   (c) Which of the constraints of key, entity integrity, and referential integrity constraints are checked by each of your operations (e.g., Insert) and which are not? 5 marks

   (d) Specify all the referential integrity (foreign key) constraints on this database. 5 marks

   (Total for que 1 is 20 marks)

Solution:

Question	**Answers**

a. Give the operations for this update.
 	
5 marks
	One possible set of operations for the given update is the following:
	i. Insert operation into SalesRep for the new SalesRep record
	ii. Insert operation into Worksin for the hours new SalesRep record worked in a SalesArea
	iii. Update operation of existing SalesArea tuple to replace the managerid of the SalesArea the new SalesRep now manages with new SRid. These operations can be specified in English (not SQL) as follows:
	INSERT <SRid, SRname, SRage, salary> into SalesRep; // for new SalesRep record
	INSERT <SRid, Arid, hours> into WORKSIN; // for new SalesRep record
	MODIFY the SalesArea tuple with the condition:
	(SalesArea.Arid = given Arid) by setting
	SalesArea.managerid = newSalesRep.eid;

b. What types of constraints would you expect to check? 5 marks	We would check that with the first INSERT operation that the SRid assigned to the new SalesRep does not already exist (key constraint). With the second INSERT operation, we would check that the new SalesRep Arid in Worksin already exists. With the third MODIFY instruction, we would check that the managerid we are writing is already an existing SRid (foreign key constraint).
c.) Which of the constraints of key, entity integrity, and referential integrity constraints are checked by each of your operations (e.g., Insert) and which are not? 5 marks	The INSERT operations into SalesRep and WORKSIN will check all the key, entity integrity, and referential integrity constraints for the relations. The MODIFY operation in DEPT will also check all the key, entity integrity, and referential integrity constraints for the relation
d. Specify all the referential integrity constraints on this database. 5 marks	We will write a referential integrity constraint as R.A --> S (or R.(X) --> T) whenever attribute A (or the set of attributes X) of relation R form a foreign key that references the primary key of relation S (or T). WORKSIN.SRid --> SalesRep WORKSIN.ARid --> SalesArea SalesArea.managerid --> SalesRep.SRid

2. (total marks 10) Using your own SalesRep-Worksin-SalesArea database instance from assignment 1, login to the SQL query processor on our cs server, called Oracle Sqlplus to create the three database tables and insert the tuples in your database state with the following sequence of instructions. Note that this exercise is to get you beginning to connect to SQLplus while preparing to learn full SQL language syntax in Chapters 6 and 7. You will be given the instructions to use now. Show the result of this exercise through a script file you will attach.

(Total for que 2 is 10 marks)

i. First connect to our cs.uwindsor.ca through either SSH client or NoMachine
ii. Then create a script file to capture your Unix session and connect to Sqlplus with:

(Note that SalesArea here does not have city attribute. Thus, just type in only 4 attributes and values for it).

>script username_assn2que2
>sqlplus <username>
>password

```
SQL> CREATE TABLE SalesRep
(SRid      NUMBER(4) NOT NULL,
 SRname    VARCHAR2(15),
 SRage     NUMBER(2),
Salary     NUMBER(8, 2),
PRIMARY    KEY(SRid));

SQL> CREATE TABLE SalesArea
(Arid   NUMBER(4) NOT NULL,
 Aname  VARCHAR2(15),
 city   VARCHAR2(15),
 budget NUMBER(10,2),
managerid       NUMBER(4),
PRIMARY         KEY(Arid),
FOREIGN KEY (managerid) REFERENCES SalesRep (SRid));

SQL> CREATE TABLE WORKSIN
(SRid  NUMBER(4) NOT NULL,
Arid   NUMBER(4) NOT NULL,
Hours  NUMBER(3,1)  ,
PRIMARY KEY(SRid, Arid),
FOREIGN KEY (SRid) REFERENCES SalesRep (SRid),
FOREIGN KEY (Arid) REFERENCES SalesArea(Arid));

SQL> INSERT INTO SalesRep
VALUES (1111, 'John Smith', 22, 20000.60);
SQL> COMMIT;
// Repeat similar INSERT instructions for all the data in your 3 tables
// starting with the entity tables first, eg, SalesRep, SalesArea, before
WORKSIN. After each INSERT, issue a COMMIT as done above.

SQL> select * from cat;   // to show all the objects in your catalogue
SQL> select * from SalesRep;  // to show the contents of this table
```

SQL> exit //to exit sqlplus

exit // to exit and create script file

** More Hint: While in Sqlplus, if you want to delete data from your tables and drop them for before issuing your instructions for your Unix script file for handing in, you can use the following instructions for each table to first delete the data from the table and then drop the table.

```
delete from Worksin;
delete from SalesArea;
delete from SalesRep;

commit;

drop table Worksin cascade constraints;

drop table SalesArea cascade constraints;

drop table SalesRep cascade constraints;

commit;
***
```

Also Note: you can start creating a script file only after you have created your tables correctly and inserted data in the tables. In that case, you cannot re-create existing tables. Then, you can then just run the desc table (eg. Desc Member) commands for each table to show the structure of each table before using (for example), the (select * from Member;) to show the tuples of each table or delete data and drop the tables as explained above so you can re-create the tables more correctly.

Solution: (10 marks)

An attached script showing execution of CREATE TABLE instructions and INSERT INTO table instructions with the few SELECT instructions to show contents of the catalogue and tables. The results of these queries should be displayed in the script file too.

CHAPTER 14: Database Design Theory: Introduction to Normalization Using Functional and Multivalued Dependencies

3. (total marks 10) Consider the following relation:

Enrolled(Studid, Crsid, Ctitle, Score, Lettergrade)

Assume that a student (Studid) may be enrolled in multiple courses (Crsid) and hence {Studid, Crsid } is the primary key.
Thus, the following functional dependency exists:

{Studid, Crsid } -> {Ctitle, Score, Lettergrade}

Additional functional dependencies are:

Crsid -> Ctitle

Score -> Lettergrade

Based on the given primary key and the dependencies,

i. Is this relation in 1NF, 2NF, or 3NF? Why or why not?
ii. If not in 2NF at least, normalize it completely into 2NF and 3NF? Provide your answers using all functional dependencies (FDs) in this database that apply to your argument.

(Total for que 3 is 10 marks)

Solution (i): (5 marks)

Answer:
Given the relation schema
Enrolled(Studid, Crsid, Ctitle, Score, Lettergrade)

{Studid, Crsid } -> {Ctitle, Score, Lettergrade}
Crsid -> Ctitle
Score -> Lettergrade

This relation satisfies 1NF but not 2NF because there is part of the primary key that determines a non-key attribute (Crsid -> Ctitle).
So the attribute (Sname) is not FFD on the primary key {Studid, Crsid } and the relation is not in 3NF because it is not in 2NF and there is transitive dependency between the primary key and some non-key attributes, e.g., Lettergrade (eg., {Studid, Crsid } -> Score;
and Score -> Lettergrade).
.

Solution (ii) (5 marks)

To normalize into 2NF and 3NF, we break the relations into relations that have only FDs that are FFD on the primary key for 2NF and also relations that have only FDs in each relation that are not transitively dependent on the primary key. The results are given below.

2NF: (keep only relations with FDs that are FFD on PK)
Enrolled1(Studid, Crsid, Score, Lettergrade)
Enrolled2(Crsid, Ctitle)

3NF: (Also remove the transitive dependencies)
Enrolled1(Studid, Crsid, Score)
Enrolled2(Crsid, Ctitle)
Enrolled3(Score, Lettergrade)

4. (total marks 10) What (i) update, (ii) delete and (iii) insertion anomalies occur in the DEPARTMENT_PROJECT relation obtained by doing a natural join of the two relations DEPARTMENT and PROJECT of Fig 14.2 on page 463 of book? Explain with examples using this database and the DEPARTMENT_PROJECT relation schema with state given below as Figures 4.1 and 4.2 below.

(Total for que 4 is 10 marks)

Note: 3 marks for correct discussion of each anomaly and 1 marks for attempt.

Figure 14.2 (book): Sample database state for a simplified COMPANY relation DB

EMPLOYEE

Ename	Ssn	Bdate	Address	Dnumber
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4
Wallace, Jennifer S.	987654321	1941-06-20	291Berry, Bellaire, TX	4
Narayan, Ramesh K.	666884444	1962-09-15	975 Fire Oak, Humble, TX	5
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1

DEPARTMENT

Dname	Dnumber	Dmgr_ssn
Research	5	333445555
Administration	4	987654321
Headquarters	1	888665555

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS_ON

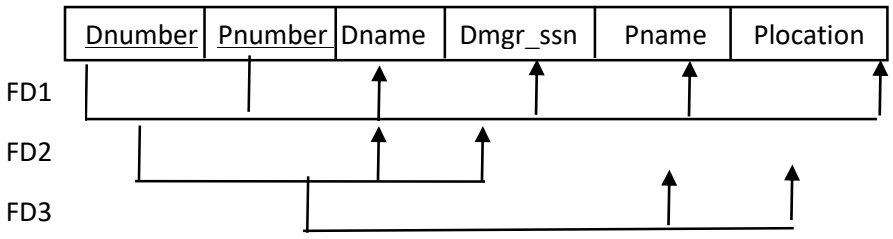
Ssn	Pnumber	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	Null

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

****

Fig 4.1: DEPARTMENT_PROJECT DB schema suffering from update anomalies



***

Fig 4.2: A database state of the DEPARTMENT_PROJECT DATABASE derived from Fig 14.2

DNUMBER	PNUMBER	DNAME	DMGR_SSN	PNAME	PLOCATION
5	3	Research	333445555	ProductZ	Houston
5	10	Research	333445555	Computerize	Stafford
5	20	Research	333445555	Reorganize	Houston
5	30	Research	333445555	Nbenefits	Stafford
5	1	Research	333445555	ProductX	Bellair
5	2	Research	333445555	ProductY	Sugarland
4	3	Administration	987654321	ProductZ	Houston
4	10	Administration	987654321	Computerize	Stafford
4	20	Administration	987654321	Reorganize	Houston
4	30	Administration	987654321	Nbenefits	Stafford
4	1	Administration	987654321	ProductX	Bellair
4	2	Administration	987654321	ProductY	Sugarland
1	3	Headquarters	888665555	ProductZ	Houston
1	10	Headquarters	888665555	Computerize	Stafford
1	20	Headquarters	888665555	Reorganize	Houston
1	30	Headquarters	888665555	Nbenefits	Stafford
1	1	Headquarters	888665555	ProductX	Bellair
1	2	Headquarters	888665555	ProductY	Sugarland

18 rows selected.

Solution: (3 + 3 + 3 + 1 marks)

i.	Update Anomalies: In DEPARTMENT_PROJECT, the partial dependencies {DNUMBER}->{DNAME} and {PNUMBER}->{PNAME,PLOCATION} can cause update anomalies. For example, all PROJECT records whose Pnumber is 2 (about 3 records) all have {PNAME, PLOCATION} as {ProductY, Sugarland}. We cannot update any of these records to something else or the database integrity is violated. If we fail to update some of the records to something else, the database is violated. We must remember these multiple column associations that indicate data redundancy and

change them all in all the associated row records or there is a violation and this is update anomaly.

ii. Delete Anomalies: : For example, if a PROJECT temporarily has no DEPARTMENTS working on it, its information (PNAME, PNUMBER, PLOCATION) will not be represented in the database when the last DEPARTMENT working on it is removed (this is deletion anomaly).

iii. Insertion anomaly: Inserting a new tuple relating an existing DEPARTMENT to an existing PROJECT requires checking both partial dependencies; for example, if a different value is entered for PLOCATION than those values in other tuples with the same value for PNUMBER, we get an update anomaly. Also, if a project is not yet assigned to a department, its record cannot be inserted into the database and this is insert anomaly.