Solutions for the Module 11 Assignment

The graded problems in Module 11 involve falsifying functional dependencies, converting ERDs to table designs, and applying the rules of normalization. You should label the problems in a document so that the grader can easily match your work to the specified problems.

1. You should identify insertion, update, and deletion anomalies in the sample rows of the big patient table shown in Table 1. You should identify one example of each type of anomaly. Each example should have sufficient specificity for the sample rows in the table. The combination of *VisitNo* and *ProvNo* is the only unique column(s) for the table.

Table 1: Sample Rows for the Big Patient Table

VisitNo	VisitDate	PatNo	PatAge	PatCity	PatZip	ProvNo	ProvSpecialty	Diagnosis
V10021	2/13/2021	P1	36	DENVER	80217	D1	INTERNIST	EAR INFECTION
V10021	2/13/2021	P1	36	DENVER	80217	D2	NURSE PRACTITIONER	INFLUENZA
V93030	2/20/2021	P3	17	ENGLEWOOD	80113	D2	NURSE PRACTITIONER	PREGNANCY
V82110	2/18/2021	P2	60	BOULDER	85932	D3	CARDIOLOGIST	MURMUR

Answer

- Insertion anomaly: patients, providers, and visits cannot be entered alone. *VisitNo* and *ProvNo* must be entered for each row because they are the primary key. For example, to insert a new patient P5, a visit number and provider number must be provided.

- Update anomaly: Multiple rows may be needed to be changed to update patient zipcode, visit date, patient city, provider specialty, provider email, and patient age. For example to update the age of patient P1, two rows must be updated.

- Deletion anomaly: deleting a row can inadvertently cause deletion of a patient, visit, provider, and zip code if any of these entities are associated with just a single diagnosis (combination of visit and provider). For example, if row 3 is deleted, details about patient P3 and visit V93030 are lost.

2. Apply the simple BCNF procedure to define BCNF tables using the FD list given in problem
   1. Show the result of each step in your analysis. For the result, you should show the tables,
   columns, primary key of each table, foreign keys, and unique constraints. You do not need to
   provide CREATE TABLE statements.

Table 1: FDs for the Big Patient Table

PatNo → PatAge
PatZip → PatCity
VisitNo → VisitDate
PatNo → PatZip
ProvNo → ProvSpecialty
VisitNo → PatNo
VisitNo, ProvNo → Diagnosis
ProvNo → ProvEmail
ProvEmail → ProvNo

Answer

Step 1: Arrange the remaining FDs into groups by determinant

PatNo → PatAge, PatZip
PatZip → PatCity
ProvNo → ProvSpecialty
VisitNo → PatNo, VisitDate
ProvEmail → ProvNo
VisitNo, ProvNo → Diagnosis

Step 2: For each FD group, make a table with the determinant as the primary key. In the table

list, the primary keys are underlined.

Patient (PatNo, PatAge, PatZip)
        FOREIGN KEY (PatZip) REFERENCES ZipCode

ZipCode(PatZip, PatCity)

Provider(ProvNo, ProvSpecialty, ProvEmail)

ProviderEmail(ProvEmail, ProvNo)
     FOREIGN KEY (ProvNo) REFERENCES Provider

Visit(VisitNo, VisitDate, PatNo)
     FOREIGN KEY (PatNo) REFERENCES Patient

DiagnosisTbl(VisitNo, ProvNo, Diagnosis)
     FOREIGN KEY (VisitNo) REFERENCES Visit
     FOREIGN KEY (ProvNo) REFERENCES Provider

Step 3: The Provider and ProviderEmail tables are merged and a unique constraint is added.

Patient (PatNo, PatAge, PatZip)
     FOREIGN KEY (PatZip) REFERENCES ZipCode

ZipCode(PatZip, PatCity)

Provider(ProvNo, ProvSpecialty, ProvEmail)
  UNIQUE ProvEmail

Visit(VisitNo, VisitDate, PatNo)
     FOREIGN KEY (PatNo) REFERENCES Patient

DiagnosisTbl(VisitNo, ProvNo, Diagnosis)
     FOREIGN KEY (VisitNo) REFERENCES Visit
     FOREIGN KEY (ProvNo) REFERENCES Provider

3. You should determine if the *Student*, *Lender*, and *Institution* tables are in BCNF. In the

    Lender table, *LenderName* is unique. In the *Institution* table, *InstName* is unique. In the

    Student table, *StdEmail* is unique. The primary key of each table is underlined. You should

    explain your decision and modify the table design by splitting tables or adding constraints if

    necessary.

    Student (StdNo, StdName, StdEmail, StdAddress, StdCity, StdState, StdZip)

Lender(LenderNo, LenderName)

Institution(InstNo, InstName, InstMascot)

Answer

LenderName, InstName, and StdEmail are determinants. The Student, Lender, and Institution

tables are in BCNF because these determinants are unique. The tables should not be split

because they are in BCNF. Unique constraints should be added for *LenderName*, *InstName*,

and *StdEmail*. Here is the revised table design with the unique constraints.

Student (StdNo, StdName, StdEmail, StdAddress, StdCity, StdState, StdZip)
     UNIQUE (StdEmail)

Lender(LenderNo, LenderName)
     UNIQUE (LenderName)

Institution(InstNo, InstName, InstMascot)
     UNIQUE (InstName)

4. For the big order database table in Table 2, you should list FDs with the column *OrdNo* as

the determinant. For each FD, you should identify at least one pair of sample rows that

falsify it or indicate that no falsification example exists for the FD. Remember that it takes

two rows to falsify an FD in which the LHS is the same in both rows, but the RHS is

different in both rows.

Table 2: Sample Rows for the Big Order Database Table

OrdNo	ItemNo	QtyOrd	CustNo	CustBal	CustDisc	ItemPrice	OrdDate
O1	I1	10	C1	100	0.10	10	1/15/2021
O1	I2	10	C1	100	0.10	20	1/15/2021
O2	I3	5	C2	200	0.05	30	1/16/2021
O2	I4	10	C2	200	0.05	40	1/16/2021
O3	I1	10	C1	100	0.10	10	1/17/2021

Answer

OrdNo FDs and sample data that violates the FDs. The rows refer to the sample rows above. Only 1 falsification per FD is necessary to identify.

FD	Falsifications
$OrdNo \rightarrow ItemNo$	(1,2), (3,4)
$OrdNo \rightarrow QtyOrd$	(3,4)
$OrdNo \rightarrow CustNo$	none
$OrdNo \rightarrow CustBal$	none
$OrdNo \rightarrow CustDisc$	none
$OrdNo \rightarrow ItemPrice$	(1,2), (3,4)
$OrdNo \rightarrow OrdDate$	None