

CIS/CSC384

HomeWork 6: Due Apr 8, 2010 before class

You will turn in the assignment via bb. Turn in a single word/pdf document.

1. For the patient table given below, list the FDs with the column PatZip as the determinant that are not true due to the sample data. With each FD that does not hold, list 2 rows that contradict it. [1.5 pts]

<u>VisitNo</u>	<u>VisitDate</u>	<u>PatNo</u>	<u>PatAge</u>	<u>PatCity</u>	<u>PatZip</u>	<u>ProvNo</u>	<u>ProvSpecialty</u>	<u>Diagnosis</u>
V10020	1/13/2007	P1	35	DENVER	80217	D1	INTERNIST	EAR INFECTION
V10020	1/13/2007	P1	35	DENVER	80217	D2	NURSE PRACTITIONER	INFLUENZA
V93030	1/20/2007	P3	17	ENGLEWOOD	80113	D2	NURSE PRACTITIONER	PREGNANCY
V82110	1/18/2007	P2	60	BOULDER	85932	D3	CARDIOLOGIST	MURMUR

Solution:

PatZip -> ProvNo is violated (first 2 rows have same PatZip, but different ProvNo)

PatZip -> ProvSpecialty is violated (first 2 rows have same PatZip, but different ProvSpecialty)

PatZip -> Diagnosis is violated (first 2 rows have same PatZip, but different Diagnosis)

Grading: If students miss 1 or 2 FD violations, they specify incorrect violations, or they do not specify the reason correctly, reduce 0.5 pts/1 pt.

2. Consider the following FDs on the patient table above. [2 pts]

PatNo -> PatAge, PatCity, PatZip

PatZip -> PatCity

ProvNo -> ProvSpecialty

VisitNo -> PatNo, VisitDate, PatAge, PatCity, PatZip

VisitNo, ProvNo -> Diagnosis

Apply the simple synthesis procedure using the above FDs to come up with the 3NF schema.

Note that, after Step 2 (after removing derived FDs), the minimal set of FDs is given below:

PatNo -> PatAge, PatZip

PatZip -> PatCity

ProvNo -> ProvSpecialty

VisitNo -> PatNo, VisitDate

VisitNo, ProvNo -> Diagnosis [2 pts]

Solution:

Patient(PatNo, PatAge, PatZip)

ZipCode(PatZip, PatCity)
Provider(ProvNo, ProvSpecialty)
Visit(VisitNo, PatNo, VisitDate)
DiagnosisTbl(VisitNo, ProvNo, Diagnosis)

Grading: Give partial credits based on how much progress the students are making. Make sure they specify the correct key constraints. The book also asks them to specify foreign key constraints – they can do so, but I do not require them to specify FKs. (Do not decrease points for incorrect FKs, or non-specification of FKs, but do correct them if incorrect FKs).

FKs if they specify include:

Patient.PatZip REFERENCES ZipCode.PatZip
Visit.PatNo REFERENCES Patient.PatNo
DiagnosisTbl.VisitNo REFERENCES Visit.VisitNo
DiagnosisTbl.ProvNo REFERENCES Provider.ProvNo

3. Consider the FDs listed below for an order entry table. Apply the synthesis procedure to come up with the 3NF design. (Note that there are no derived FDs, so the following lists the FDs after step 2). You must identify all candidate keys for every table. [2 pts]

CustNo → CustBal, CustDiscount
OrderNo → CustNo, ShipAddr, OrderDate
ItemNo → ItemDesc
ItemNo, PlantNo → ReorderPoint, QtyOnHand
OrderNo, ItemNo → LineNo, QtyOrdered, QtyOutstanding
OrderNo, LineNo → ItemNo, QtyOrdered, QtyOutstanding

Solution:

Customer(CustNo, CustBal, CustDiscount)
Order(OrderNo, CustNo, ShipAddr, OrderDate)
Item(ItemNo, ItemDesc)
PlantStock(PlantNo, ItemNo, ReorderPoint, QtyOnHand)
OrderLine(OrderNo, ItemNo, LineNo, QtyOrdered, QtyOutstanding)
(OrderNo, LineNo) is a candidate key for OrderLine

Grading: Make sure students get only one OrderLine table, that is the one which has an additional candidate key. This is the result of merging (Step 5).

4. Consider the following FDs for a simplified expense report table. Note that there are three candidate keys: <ExpItemNo>, the combination <ERNo, CatNo>, and the combination <ERNo, CatName>

ERNo -> UserNo, ERSubmitDate, ERStatusDate

ExplItemNo -> ExplItemDesc, ExplItemDate, ExplItemAmt, CatNo, ERNo

UserNo -> UserFirstName, UserLastName, UserPhone, UserEmail

CatNo -> CatName, CatLimit

ERNo, CatNo -> ExplItemNo

UserEmail -> UserNo

CatName -> CatNo

- (a) If we have all the columns in 1 table, is the design in 2NF? (Remember: a schema violates 2NF, if there is a FD $X \rightarrow Y$, where X is part of a candidate key, and Y is a non-key column). If the design is not in 2NF, you must specify all the FDs that violate 2NF. [1 pt]

Solution:

ERNo -> UserNo, ERSubmitDate, ERStatusDate

CatNo -> CatLimit

Grading: Note that

ERNo -> UserNo, ERSubmitDate, ERStatusDate can also be written as:

ERNo -> UserNo

ERNo -> ERSubmitDate

ERNo -> ERStatusDate

Basically, there are 4 FDs that violate 2NF. You must decrease points, if they miss some, or they specify FDs incorrectly.

- (b) If we have all the columns in 1 table, is the design in 3NF? (Remember: a schema violates 3NF, if there is a FD $X \rightarrow Y$, where Y is a non-key column, and X is not part of any key). If the design is not in 3NF, you must specify all the FDs that violate 3NF. [1 pt]

Solution:

UserNo -> UserFirstName, UserLastName, UserPhone, UserEmail

UserEmail -> UserNo

- (c) Come up with the set of 3NF tables for this scenario, using the synthesis procedure. Note that the FDs given are minimal (result of Step 2 of the synthesis procedure). [2.5 pts]

Solution:

After Step 4, we get:

ExpenseReport(ERNo, UserNo, ERSubmitDate, ERStatusDate)

ExpenseItem(ExplItemNo, ExplItemDesc, ExplItemDate, ExplItemAmt, CatNo, ERNo)

ExpenseDet(ERNo, CatNo, ExplItemNo)

Note that Step 5 will merge ExpenseItem and ExpenseDet tables.

User(UserNo, UserFirstName, UserLastName, UserPhone, UserEmail)

UserEm(UserEmail, UserNo)

Note: Step 5 will merge User and UserEm tables.

Category(CatNo, CatName, CatLimit)

CategoryN(CatName, CatNo)

Note that Step 5 will merge Category and CategoryN tables.

Resulting set of tables after merging (Step 5) are:

ExpeseReport(ERNo, UserNo, ERSubmitDate, ERStatusDate)

ExpenseItem(ExpItemNo, ExpItemDesc, ExpItemDate, ExpItemAmt, CatNo, ERNo)

<ERNo, CatNo> is a candidate key for ExpenseItem

User(UserNo, UserFirstName, UserLastName, UserPhone, UserEmail)

UserEmail is a candidate key for User

Category(CatNo, CatName, CatLimit)

CatName is a candidate key for Category.