CS 327E Lecture 12

Shirley Cohen

March 7, 2016

Agenda

- Announcements
- Readings for today
- Reading Quiz
- Concept Questions
- Homework for next time

Reminders

- Midterm 2 will be next class
- Project phase will start after Spring Break

Homework for Today

- Chapters 8 and 9 from the Beginning Database Design book
- Exercises at the end of Chapter 8 and 9

Quiz Question 1

What is the Insertion Problem discussed in Chapter 8?

A. If we insert a record that contains a mistake, fixing it might be impossible
B. Unique IDs may cause confusion if the non-primary key fields in two records are identical
C. Entering repeated data will cause issues referencing a record
D. We can’t enter a record if we don’t have all of the primary key fields

Quiz Question 2

What is an Update Anomaly?

A. Having to update redundant data across multiple records
B. Not being able to update a record due to a foreign key constraint
C. Being required to delete and insert a record, rather than updating it
D. Not being able to determine the primary key of a table

Quiz Question 3

Normalization is the process of decomposing the relations in a schema with the objective of reducing data redundancies.

A.True
B.False

Quiz Question 4

The primary key for a *Customer* table should always be the combination of (*customer_id, customer_name*).

A. True
B. False

Quiz Question 5

A *unique constraint* can help to prevent data integrity problems in situations when an auto-incremented number is used as the primary key (in place of the natural key).

A.True
B.False

Normalization

Students: Unnormalized to 1NF

Rule: A database schema is in 1NF *iff* all attributes have scalar values

Students_Semester

EID	Semester	GPA	Courses
alice1	Fall15	3.9	Stats A DB A Alg A-
bob20	Fall15	3.7	DB A Alg B+
carol30	Fall15	3.5	Stats A- Alg B+

unnormalized

Students_Semester'

EID	Semester	Course	Grade	GPA
alice1	Fall15	Stats	A	3.9
alice1	Fall15	DB	A	3.9
alice1	Fall15	Alg	A-	3.9
bob20	Fall15	DB	A	3.7
bob20	Fall15	Alg	B	3.7
carol30	Fall15	Stats	3.5	3.5
carol30	Fall15	Alg	3.5	3.5

1NF

Drugs: Unnormalized

A pharmaceutical company has an inventory table of drugs and their price changes.

Drugs

drug_nbr	drug_name	drug_qty	drug_price
48	Amoxicillin	500	01/01/13 – 03/31/15 – 0.30 04/01/15 – 01/15/16 – 3.00 01/16/16 – 3.50
50	Lipitor	150	10/01/12 – 03/31/14 – 0.75 04/01/14 – 1.00
72	Singulair	250	01/01/15 – 05/31/15 – 0.20 06/01/15 – 07/31/15 – 0.80 08/01/15 – 2.00

Concept Question 1

The pharma company decides to normalize the table. Is the resulting table in 1NF?

drug_nbr	drug_name	drug_qty	drug_price
48	Amoxicillin	500	01/01/13 | 03/31/15 | 0.30 04/01/15 | 01/15/16 | 3.00 01/16/16 | | 3.50
…	…	…	…

Unnormalized

drug_nbr	drug_name	drug_qty	drug_price	start_date	end_date
48	Amoxicillin	500	0.30	01/01/13	03/31/15
48	Amoxicillin	500	3.00	04/01/15	01/15/16
48	Amoxicillin	500	3.50	01/16/16	
…	…	…	…	…	…

1NF ?

A. Yes B. No C. Not enough information

Functional Dependencies

Definition:

If two records agree on the attributes

$A_1, A_2, \ldots, A_n$

then they must also agree on the attributes

$B_1, B_2, \ldots, B_n$

Formally:

$A_1, A_2, \ldots, A_n \rightarrow B_1, B_2, \ldots, B_n$

FD Example

Which FDs **hold** and **do not hold** on this table?

ID	Name	Phone	City
C0012	Smith	5555	Austin
C3412	Wallace	9876	Houston
C1111	Smith	9876	Dallas
C2323	Johnston	5555	Austin

ID → Name, Phone, City
City → Phone

Not Phone → City
Not Name → Phone

Concept Question 2

Can you find all the FDs that hold for this *Drugs* table?

drug_nbr	drug_name	drug_qty	drug_price	start_date	end_date
48	Amoxicillin	500	0.30	01/01/13	03/31/15
48	Amoxicillin	500	3.00	04/01/15	01/15/16
48	Amoxicillin	500	3.50	01/16/16	
50	Lipitor	150	0.75	10/01/12	03/31/14
50	Lipitor	150	1.00	04/01/14	
72	Singulair	250	0.20	01/01/15	05/31/15
72	Singulair	250	0.80	06/01/15	07/31/15
72	Singulair	250	0.20	08/01/15	

A.
drug_nbr → drug_name, drug_qty
drug_nbr → drug_price

B.
drug_nbr → drug_name, drug_qty
drug_nbr, drug_price → start_date

C.
drug_nbr → drug_name, drug_qty
drug_nbr, start_date → drug_price, end_date

D.
drug_nbr → drug_name, drug_qty
drug_nbr, drug_price → start_date, end_date

Drugs: 1NF to 2NF

Drugs

drug_nbr	drug_name	drug_qty	drug_price	start_date	end_date
48	Amoxicillin	500	0.30	01/01/13	03/31/15
48	Amoxicillin	500	3.00	04/01/15	01/15/16
48	Amoxicillin	500	3.50	01/16/16	
…	…	…	…	…	…

1NF

FDs:
drug_nbr → drug_name, drug_qty
drug_nbr, start_date → drug_price, end_date

Drugs'

drug_nbr	drug_name	drug_qty
48	Amoxicillin	500
50	Lipitor	150
72	Singulair	250

2NF

Prices

drug_nbr	drug_price	start_date	end_date
48	0.30	01/01/13	03/31/15
48	3.00	04/01/15	01/15/16
48	3.50	01/16/16	

2NF

Students: 1NF to 2NF

Rule: A database schema is in 2NF *iff* it is in 1NF and there exists no partial FDs on the primary key (i.e. all non-key attributes must be dependent on the entire PK)

Student_Semesters

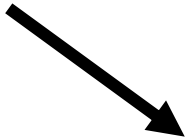
EID	Semester	Course	Grade	GPA
alice1	Fall15	Stats	A	3.9
alice1	Fall15	DB	A	3.9
alice1	Fall15	Alg	A-	3.9
bob20	Fall15	DB	A	3.7
bob20	Fall15	Alg	B+	3.7
carol30	Fall15	Stats	A-	3.5
carol30	Fall15	Alg	B+	3.5

1NF

Student_Semester_Grades

EID	Semester	Course	Grade
alice1	Fall15	Stats	A
alice1	Fall15	DB	A
alice1	Fall15	Alg	A-
bob20	Fall15	DB	A
bob20	Fall15	Alg	B+
carol30	Fall15	Stats	A-
carol30	Fall15	Alg	B+

2NF



Student_Semester_GPAs

EID	Semester	GPA
alice1	Fall15	3.9
bob20	Fall15	3.7
carol30	Fall15	3.5

2NF

FDs:

1. EID, Semester, Course → Grade, GPA
2. EID, Semester → GPA

2NF to 3NF

Rule: A database schema is in 3NF *iff* it is in 2NF and there exists no non-key fields that depend on another non-key field

Student_Majors

EID	Name	Major	College
alice1	Alice	Math	Natural Sciences
bob20	Bob	CS	Natural Sciences
carol30	Carol	Physics	Natural Sciences

FDs:

EID → Name, Major, College

Major → College

2NF

Student_Majors'

EID	Name	Major
alice1	Alice	Math
bob20	Bob	CS
carol30	Carol	Physics

3NF

Major_College

Major	College
Math	Natural Sciences
CS	Natural Sciences
Physics	Natural Sciences

3NF

Concept Question 3

Suppose we added a *drug_description* field to the *Drugs* table as shown below and we discover that drug_name → drug_description. Does this change the normal form for this table? Which normal form is the table in now?

Drugs

drug_nbr	drug_name	drug_description	drug_qty
48	Amoxicillin	Treats bacterial infections	500
50	Lipitor	Reduces cholesterol levels	150
72	Singulair	Prevents asthma symptoms	250

FDs:
drug_nbr → drug_name, drug_description, drug_qty
drug_name → drug_description

A. 1NF
B. 2NF
C. 3NF
D. None of the above

Drugs: 2NF to 3NF

Drugs

drug_nbr	drug_name	drug_description	drug_qty
48	Amoxicillin	Treats bacterial infections	500
50	Lipitor	Reduces cholesterol levels	150
72	Singulair	Prevents asthma symptoms	250

2NF

FDs:
drug_nbr → drug_name, drug_description, drug_qty
drug_name → drug_description

Drugs'

drug_nbr	drug_name	drug_qty
48	Amoxicillin	500
50	Lipitor	150
72	Singulair	250

3NF

Drug_Descriptions

drug_name	drug_description
Amoxicillin	Treats bacterial infections
Lipitor	Reduces cholesterol levels
Singulair	Prevents asthma symptoms

3NF

Midterm 2

- Closed book exam
- 90 minutes
- 3 questions

Midterm 2 Topics

- Conceptual modeling
- Conversion to relations
- Normal forms