Review Questions

3.1 Name three sources for databases.

3.2 What is the basic premise of this and the next chapter?

3.3 Explain what is wrong with the table in Figure 3-2.

3.4 Define each of the terms listed in Figure 3-3.

3.5 Describe the characteristics of a table that make it a relation. Define the term *domain* and explain the significance of the *domain integrity constraint* to a relation.

3.6 Give an example of two tables that are not relations.

3.7 Suppose that two columns in two different tables have the same column name. What convention is used to give each a unique name?

3.8 Must all the values in the same column of a relation have the same length?

3.9 Explain the three different sets of terms used to describe tables, columns, and rows.

3.10 Explain the difference between functional dependencies that arise from equations and those that do not.

3.11 Explain the intuitive meaning of the functional dependency.

PartNumber → PartWeight

3.12 Explain the following statement: "The only reason for having relations is to store instances of functional dependencies."

3.13 Explain the meaning of the expression:

(FirstName, LastName) → Phone

3.14 What is a composite determinant?

3.15 If (A, B) → C , then can we also say that A → C?

3.16 If A → (B, C), then can we also say that A → B?

3.17 For the SKU_DATA table in Figure 3-1, explain why Buyer determines Department, but Department does not determine Buyer.

3.18 For the SKU_DATA table in Figure 3-1, explain why:

SKU_Description → (SKU, Department, Buyer).

3.19 If it is true that

PartNumber → PartWeight

does that mean that PartNumber will be unique in a relation?

3.20 Under what conditions will a determinant be unique in a relation?

3.21 What is the best test for determining whether a determinant is unique?

3.22 What is a composite key?

3.23 What is a candidate key?

3.24 What is a primary key? Explain the significance of the *entity integrity constraint* to a primary key.

3.25 Explain the difference between a candidate key and a primary key.

3.26 What is a surrogate key?

3.27 Where does the value of a surrogate key come from?

3.28 When would you use a surrogate key?

3.29 What is a foreign key? Explain the significance of the *referential integrity constraint* to a foreign primary key.

3.30 The term *domestic key* is not used. If it were used, however, what do you think it would mean?

3.31 What is a normal form?

3.32 Illustrate deletion, modification, and insertion anomalies on the STUDENT_ACTIVITY relation in Figure 3-24.

3.33 Explain why duplicated data lead to data integrity problems.

3.34 What relations are in 1NF?

3.35 Which normal forms are concerned with functional dependencies?

3.36 What conditions are required for a relation to be in 2NF?

3.37 What conditions are required for a relation to be in 3NF?

3.38 What conditions are required for a relation to be in BCNF?

3.39 If a relation is in BCNF, what can we say about it with regard to 2NF and 3NF?

3.40 What normal form is concerned with multivalued dependencies?

3.41 What is the premise of Fagin's work on DK/NF?

3.42 Summarize the three categories of normalization theory.

3.43 In general, how can you transform a relation not in BCNF into ones that are in BCNF?

3.44 What is a referential integrity constraint? Define the term, and give an example of its use. Are null values allowed in foreign key columns with a referential integrity constrain? How does the referential integrity constrain contribute to *database integrity*?

3.45 Explain the role of referential integrity constraints in normalization.

3.46 Why is an un-normalized relation like a paragraph with multiple themes?

3.47 In normalization example 3, why is the EXTENDED_PRICE relation "silly"?

3.48 In normalization example 4, under what conditions is the functional dependency

(StudentID, Activity) → ActivityFee

more accurate than

Activity → ActivityFee

3.49 If a determinant is part of a candidate key, is that good enough for BCNF?

3.50 In normalization example 5, why are the following two tables not correct?

DEPARTMENT (Department, DeptBudgetCode, Buyer)
SKU_DATA_4 (SKU, SKU_Description, *Department*)

3.51 How does a multivalued dependency differ from a functional dependency?

3.52 Consider the relation:

PERSON (Name, Sibling, ShoeSize)

Assume that the following functional dependencies exist:

Name →→ Sibling

Name → ShoeSize

Describe deletion, modification, and insertion anomalies for this relation.

3.53 Place the PERSON relation in Review Question 3.52 into 4NF.

3.54 Consider the relation:

PERSON_2 (Name, Sibling, ShoeSize, Hobby)

Assume that the following functional dependencies exist:

Name →→ Sibling

Name → ShoeSize

Name →→ Hobby

Describe deletion, modification, and insertion anomalies for this relation.

3.55 Place the PERSON_2 relation in Review Question 3.54 into 4NF.

3.56 What is 5NF?

3.57 How do the conditions for DK/NF correspond to the conditions for BCNF?

Project Questions

3.58 Consider the table:

STAFF_MEETING (EmployeeName, ProjectName, Date)

The rows of this table record the fact that an employee from a particular project attended a meeting on a given date. Assume that a project meets at most once per day. Also, assume that only one employee represents a given project, but that employees can be assigned to multiple projects.

A. State the functional dependencies in STAFF_MEETING.

B. Transform this table into one or more tables in BCNF. State the primary keys, candidate keys, foreign keys, and referential integrity constraints.

C. Is your design in part B an improvement over the original table? What advantages and disadvantages does it have?

3.59 Consider the table:

STUDENT (StudentNumber, StudentName, Dorm, RoomType, DormCost, Club, ClubCost, Sibling, Nickname)

Assume that students pay different dorm costs depending on the type of room they have, but that all members of a club pay the same cost. Assume that students can have multiple nicknames.

A. State any multivalued dependencies in STUDENT.

B. State the functional dependencies in STUDENT.

C. Transform this table into two or more tables such that each table is in BCNF and in 4NF. State the primary keys, candidate keys, foreign keys, and referential integrity constraints.

Case Questions

Regional Labs Case Questions

Regional Labs is a company that conducts research and development work on a contract basis for other companies and organizations. Figure 3-32 shows data that Regional Labs collects about projects and the employees assigned to them. This data is stored in a relation (table) named PROJECT:

PROJECT (ProjectID, EmployeeName, EmployeeSalary)

A. Assuming that all functional dependencies are apparent in this data, which of the following are true?

1. ProjectID → EmployeeName
2. ProjectID → EmployeeSalary
3. (ProjectID, EmployeeName) → EmployeeSalary
4. EmployeeName → EmployeeSalary
5. EmployeeSalary → ProjectID
6. EmployeeSalary → (ProjectID, EmployeeName)

B. What is the primary key of PROJECT?

C. Are all the non-key attributes (if any) dependent on the primary key?

D. In what normal form is PROJECT?

E. Describe two modification anomalies that affect PROJECT.

F. Is ProjectID a determinant? If so, based on which functional dependencies in part A?

G. Is EmployeeName a determinant? If so, based on which functional dependencies in part A?

H. Is (ProjectID, EmployeeName) a determinant? If so, based on which functional dependencies in part A?

I. Is EmployeeSalary a determinant? If so, based on which functional dependencies in part A?

J. Does this relation contain a transitive dependency? If so, what is it?

K. Redesign the relation to eliminate modification anomalies.

Figure 3-32
Sample Data for Regional Labs

ProjectID	EmployeeName	EmployeeSalary
100-A	Eric Jones	64,000.00
100-A	Donna Smith	70,000.00
100-B	Donna Smith	70,000.00
200-A	Eric Jones	64,000.00
200-B	Eric Jones	64,000.00
200-C	Eric Parks	58,000.00
200-C	Donna Smith	70,000.00
200-D	Eric Parks	58,000.00

The Queen Anne Curiosity Shop

Figure 3-33 shows typical sales data for The Queen Anne Curiosity Shop, and Figure 3-34 shows typical purchase data.

A. Using these data, state assumptions about functional dependencies among the columns of data. Justify your assumptions on the basis of these sample data and also on the basis of what you know about retail sales.

B. Given your assumptions in part A, comment on the appropriateness of the following designs:

1. CUSTOMER (LastName, FirstName, Phone, Email, InvoiceDate, InvoiceItem, Price, Tax, Total)
2. CUSTOMER (LastName, FirstName, Phone, Email, InvoiceDate, InvoiceItem, Price, Tax, Total)
3. CUSTOMER (LastName, FirstName, Phone, Email, InvoiceDate, InvoiceItem, Price, Tax, Total)
4. CUSTOMER (LastName, FirstName, Phone, Email, InvoiceDate, InvoiceItem, Price, Tax, Total)
5. CUSTOMER (LastName, FirstName, Phone, Email, InvoiceDate, InvoiceItem, Price, Tax, Total)
6. CUSTOMER (LastName, FirstName, Phone, Email)
   and:
   SALE (InvoiceDate, InvoiceItem, Price, Tax, Total)
7. CUSTOMER (LastName, FirstName, Phone, Email, *InvoiceDate*)
   and:
   SALE (InvoiceDate, InvoiceItem, Price, Tax, Total)

Figure 3-33
Sample Sales Data for The Queen Anne Curiosity Shop

LastName	FirstName	Phone	InvoiceDate	InvoiceItem	Price	Tax	Total
Shire	Robert	206-524-2433	14-Dec-13	Antique Desk	3,000.00	249.00	3,249.00
Shire	Robert	206-524-2433	14-Dec-13	Antique Desk Chair	500.00	41.50	541.50
Goodyear	Katherine	206-524-3544	15-Dec-13	Dining Table Linens	1,000.00	83.00	1,083.00
Bancroft	Chris	425-635-9788	15-Dec-13	Candles	50.00	4.15	54.15
Griffith	John	206-524-4655	23-Dec-13	Candles	45.00	3.74	48.74
Shire	Robert	206-524-2433	5-Jan-14	Desk Lamp	250.00	20.75	270.75
Tierney	Doris	425-635-8677	10-Jan-14	Dining Table Linens	750.00	62.25	812.25
Anderson	Donna	360-538-7566	12-Jan-14	Book Shelf	250.00	20.75	270.75
Goodyear	Katherine	206-524-3544	15-Jan-14	Antique Chair	1,250.00	103.75	1,353.75
Goodyear	Katherine	206-524-3544	15-Jan-14	Antique Chair	1,750.00	145.25	1,895.25
Tierney	Doris	425-635-8677	25-Jan-14	Antique Candle Holders	350.00	29.05	379.05

PurchaseItem	PurchasePrice	PurchaseDate	Vendor	Phone
Antique Desk	1,800.00	7-Nov-13	European Specialties	206-325-7866
Antique Desk	1,750.00	7-Nov-13	European Specialties	206-325-7866
Antique Candle Holders	210.00	7-Nov-13	European Specialties	206-325-7866
Antique Candle Holders	200.00	7-Nov-13	European Specialties	206-325-7866
Dining Table Linens	600.00	14-Nov-13	Linens and Things	206-325-6755
Candles	30.00	14-Nov-13	Linens and Things	206-325-6755
Desk Lamp	150.00	14-Nov-13	Lamps and Lighting	206-325-8977
Floor Lamp	300.00	14-Nov-13	Lamps and Lighting	206-325-8977
Dining Table Linens	450.00	21-Nov-13	Linens and Things	206-325-6755
Candles	27.00	21-Nov-13	Linens and Things	206-325-6755
Book Shelf	150.00	21-Nov-13	Harrison, Denise	425-746-4322
Antique Desk	1,000.00	28-Nov-13	Lee, Andrew	425-746-5433
Antique Desk Chair	300.00	28-Nov-13	Lee, Andrew	425-746-5433
Antique Chair	750.00	28-Nov-13	New York Brokerage	206-325-9088
Antique Chair	1,050.00	28-Nov-13	New York Brokerage	206-325-9088

Figure 3-34
Sample Purchase Data for The Queen Anne Curiosity Shop

8. CUSTOMER (LastName, FirstName, Phone, Email, *InvoiceDate*, *InvoiceItem*)
   and:
   SALE (InvoiceDate, Item, Price, Tax, Total)

C. Modify what you consider to be the best design in part B to include surrogate ID columns called CustomerID and SaleID. How does this improve the design?

D. Modify the design in part C by breaking SALE into two relations named SALE and SALE_ITEM. Modify columns and add additional columns as you think necessary. How does this improve the design?

E. Given your assumptions, comment on the appropriateness of the following designs:

1. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate, Vendor, Phone)
2. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate, Vendor, Phone)
3. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate, Vendor, Phone)
4. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate, Vendor, Phone)
5. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate)
   and:
   VENDOR (Vendor, Phone)

6. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate,Vendor)

   and:

   VENDOR (Vendor, Phone)

7. PURCHASE (PurchaseItem, PurchasePrice, PurchaseDate, *Vendor*)

   and:

   VENDOR (Vendor, Phone)

F. Modify what you consider to be the best design in part E to include surrogate ID columns called PurchaseID and VendorID. How does this improve the design?

G. The relations in your design from part D and part F are not connected. Modify the database design so that sales data and purchase data are related.

Morgan Importing

James Morgan keeps a table of data about the stores from which he purchases. The stores are located in different countries and have different specialties. Consider the following relation:

STORE (StoreName, City, Country, OwnerName, Specialty)

A. Explain the conditions under which each of the following is true:

1. StoreName → City
2. City → StoreName
3. City → Country
4. (StoreName, Country) → (City, OwnerName)
5. (City, Specialty) → StoreName
6. OwnerName →→ StoreName
7. StoreName →→ Specialty

B. With regard to the relation in part A:

1. Specify which of the dependencies in part A seem most appropriate for a small import–export business.
2. Given your assumptions in B.1, transform the STORE table into a set of tables that are in both 4NF and BCNF. Indicate the primary keys, candidate keys, foreign keys, and referential integrity constraints.

C. Consider the relation:

SHIPMENT (ShipmentNumber, ShipperName, ShipperContact, ShipperFax, DepartureDate, ArrivalDate, CountryOfOrigin, Destination, ShipmentCost, InsuranceValue, Insurer)

1. Write a functional dependency that expresses the fact that the cost of a shipment between two cities is always the same.
2. Write a functional dependency that expresses the fact that the insurance value is always the same for a given shipper.
3. Write a functional dependency that expresses the fact that the insurance value is always the same for a given shipper and country of origin.
4. Describe two possible multivalued dependencies in SHIPMENT.
5. State what you believe are reasonable functional dependencies for the SHIPMENT relation for a small import–export business.
6. State what you believe are reasonable multivalued dependencies for the SHIPMENT relation.
7. Using your assumptions in 5 and 6, transform SHIPMENT into a set of tables in BCNF and 4NF. Indicate the primary keys, candidate keys, foreign keys, and referential integrity constraints.