9. Suppose you have been given the table structure and data shown in Table P6.9, which was imported from an Excel spreadsheet. The data reflects that a professor can have multiple advisees, can serve on multiple committees, and can edit more than one journal.

TABLE P6.9

ATTRIBUTE NAME	SAMPLE VALUE	SAMPLE VALUE	SAMPLE VALUE	SAMPLE VALUE
EMP_NUM	123	104	118	
PROF_RANK	Professor	Asst. Professor	Assoc. Professor	Assoc. Professor
EMP_NAME	Ghee	Rankin	Ortega	Smith
DEPT_CODE	CIS	CHEM	CIS	ENG
DEPT_NAME	Computer Info. Systems	Chemistry	Computer Info. Systems	English
PROF_OFFICE	KDD-567	BLF-119	KDD-562	PRT-345
ADVISEE	1215, 2312, 3233, 2218, 2098	3102, 2782, 3311, 2008, 2876, 2222, 3745, 1783, 2378	2134, 2789, 3456, 2002, 2046, 2018, 2764	2873, 2765, 2238, 2901, 2308
COMMITTEE_CODE	PROMO, TRAF, APPL, DEV	DEV	SPR, TRAF	PROMO, SPR, DEV
JOURNAL_CODE	JMIS, QED, JMGT		JCIS, JMGT	

Given the information in Table P6.9:

a. Draw the dependency diagram.

b. Identify the multivalued dependencies.

c. Create the dependency diagrams to yield a set of table structures in 3NF.

d. Eliminate the multivalued dependencies by converting the affected table structures to 4NF.

e. Draw the Crow's Foot ERD to reflect the dependency diagrams you drew in Problem 9c. (*Note:* You might have to create additional attributes to define the proper PKs and FKs. Make sure that all of your attributes conform to the naming conventions.)

10. The manager of a consulting firm has asked you to evaluate a database that contains the table structure shown in Table P6.10.

Table P6.10 was created to enable the manager to match clients with consultants. The objective is to match a client within a given region with a consultant in that region and to make sure that the client's need for specific consulting services is properly matched to the consultant's expertise. For example, if the client needs help with database design and is located in the Southeast, the objective is to make a match with a consultant who is located in the Southeast and whose expertise is in database design. (Although the consulting company manager tries to match consultant and client locations to minimize travel expense, it is not always possible to do so.) The following basic business rules are maintained:

- Each client is located in one region.
- A region can contain many clients.
- Each consultant can work on many contracts.

TABLE P6.10

ATTRIBUTE NAME	SAMPLE VALUE	SAMPLE VALUE	SAMPLE VALUE
CLIENT_NUM	298	289	289
CLIENT_NAME	Marianne R. Brown	James D. Smith	James D. Smith
CLIENT_REGION	Midwest	Southeast	Southeast
CONTRACT_DATE	10-Feb-2018	15-Feb-2018	12-Mar-2018
CONTRACT_NUMBER	5841	5842	5843
CONTRACT_AMOUNT	$2,985,000.00	$670,300.00	$1,250,000.00
CONSULT_CLASS_1	Database Administration	Internet Services	Database Design
CONSULT_CLASS_2	Web Applications		Database Administration
CONSULT_CLASS_3			Network Installation
CONSULT_CLASS_4			
CONSULTANT_NUM_1	29	34	25
CONSULTANT_NAME_1	Rachel G. Carson	Gerald K. Ricardo	Angela M. Jamison
CONSULTANT_REGION_1	Midwest	Southeast	Southeast
CONSULTANT_NUM_2	56	38	34
CONSULTANT_NAME_2	Karl M. Spenser	Anne T. Dimarco	Gerald K. Ricardo
CONSULTANT_REGION_2	Midwest	Southeast	Southeast
CONSULTANT_NUM_3	22	45	
CONSULTANT_NAME_3	Julian H. Donatello	Geraldo J. Rivera	
CONSULTANT_REGION_3	Midwest	Southeast	
CONSULTANT_NUM_4		18	
CONSULTANT_NAME_4		Donald Chen	
CONSULTANT_REGION_4		West	

- Each contract might require the services of many consultants.
- A client can sign more than one contract, but each contract is signed by only one client.
- Each contract might cover multiple consulting classifications. For example, a contract may list consulting services in database design and networking.
- Each consultant is located in one region.
- A region can contain many consultants.
- Each consultant has one or more areas of expertise (class). For example, a consultant might be classified as an expert in both database design and networking.
- Each area of expertise (class) can have many consultants. For example, the consulting company might employ many consultants who are networking experts.

a. Given this brief description of the requirements and the business rules, write the relational schema and draw the dependency diagram for the preceding (and very poor) table structure. Label all transitive and/or partial dependencies.

b. Break up the dependency diagram you drew in Problem 10a to produce dependency diagrams that are in 3NF and write the relational schema. (*Hint:* You might

have to create a few new attributes. Also make sure that the new dependency diagrams contain attributes that meet proper design criteria; that is, make sure there are no multivalued attributes, that the naming conventions are met, and so on.)

c. Using the results of Problem 10b, draw the Crow's Foot ERD.

11. Given the sample records in the CHARTER table shown in Table P6.11, do the following:

a. Write the relational schema and draw the dependency diagram for the table structure. Make sure that you label all dependencies. CHAR_PAX indicates the number of passengers carried. The CHAR_MILES entry is based on round-trip miles, including pickup points. (*Hint:* Look at the data values to determine the nature of the relationships. For example, note that employee Melton has flown two charter trips as pilot and one trip as copilot.)

b. Decompose the dependency diagram you drew to solve Problem 11a to create table structures that are in 3NF and write the relational schema.

c. Draw the Crow's Foot ERD to reflect the properly decomposed dependency diagrams you created in Problem 11b. Make sure the ERD yields a database that can track all of the data shown in Problem 11. Show all entities, relationships, connectivities, optionalities, and cardinalities.

TABLE P6.11

ATTRIBUTE NAME	SAMPLE VALUE	SAMPLE VALUE	SAMPLE VALUE	SAMPLE VALUE
CHAR_TRIP	10232	10233	10234	10235
CHAR_DATE	15-Jan-2018	15-Jan-2018	16-Jan-2018	17-Jan-2018
CHAR_CITY	STL	MIA	TYS	ATL
CHAR_MILES	580	1,290	524	768
CUST_NUM	784	231	544	784
CUST_LNAME	Brown	Hanson	Bryana	Brown
CHAR_PAX	5	12	2	5
CHAR_CARGO	235 lbs.	18,940 lbs.	348 lbs.	155 lbs.
PILOT	Melton	Chen	Henderson	Melton
COPILOT		Henderson	Melton	
FLT_ENGINEER		O'Shaski		
LOAD_MASTER		Benkasi		
AC_NUMBER	1234Q	3456Y	1234Q	2256W
MODEL_CODE	PA31-350	CV-580	PA31-350	PA31-350
MODEL_SEATS	10	38	10	10
MODEL_CHG_MILE	$2.79	$23.36	$2.79	$2.79

4-56. The following attributes represent data about a movie copy at an online video rental service. Each movie is identified by a movie number and has a title and information about the director and the studio that produced the movie. Each movie has one or several characters, and there is exactly one actor playing the role of each of the characters (but one actor can play multiple roles in each of the movies). The video rental service has multiple licenses for the same movie, and the service differentiates the licenses with a movie copy number, which is unique within a single movie but not unique between different movies. Each movie license has a rental status and return date; in addition, each license has a type (Regular or HD). The rental price depends on the movie and the license type, but the price is the same for all licenses of the same type. The attributes are as follows:

Movie Nbr, Title, Director ID, Director Name, Studio ID, Studio Name, Studio Location, Studio CEO, Character, Actor ID, Name, Movie License Nbr, Movie License Type, Movie Rental Price, License Rental Status, License Return Date

A sample data set regarding a movie would be as follows (the data in the curly brackets are character/actor data, in this case for four different characters):

567, "It's a Wonderful Life", 25, "Frank Capra", 234, "Liberty Films", "Hollywood, CA", "Orson Wells", {"George Bailey", 245, "James Stewart" | "Mary Bailey", 236, "Donna Reed" | "Clarence Oddbody", 765, "Henry Travers" | "Henry F. Potter", 325, "Lionel Barrymore" }, 5434, "HD", 3.95, "Rented", "12/15/2015"

Based on this information,

a. Identify the functional dependencies between the attributes.
b. Identify the reasons why this set of data items is not in 3NF and tell what normal form (if any) it is in.
c. Present the attributes organized into 3NF relations that have been named appropriately.

4-57. A start-up is working on an online personal financial management system. The goal of the system is to provide the users an opportunity to obtain item-level purchase data from as many sources as possible in order to improve the accuracy of budget management and control activities (instead of only at the level of the total of each purchase). For example, let's assume a customer purchases three books from a major online bookseller. For most financial management software systems, the system only receives the total of the purchase from a bank or other financial institution. In the case of this start-up, the intent is to create a link between the financial transaction and the vendor's system data so that the financial management system retrieves product details from the vendor. Now it will be easy for the customer to classify one book as self-help, the other one as a business expense, and the third one as entertainment without having to resort to an analysis of receipts.

To provide this capability, the system maintains the following data regarding the transactions:

TransactionID, CustomerID, CustomerName, CustomerEmail, TransactionDate, TransactionTime, TransactionTotalAmount, TransactionTax, ProductID, ProductDescription, ProductCategory, ProductManufacturerID, ManufacturerName, ProductListPrice, ProductPurchasePrice, ProductQuantity, TransactionProductTotal

Sample data for this set of attributes is as follows:

823434434582, 2434254, Silver Patrick, psilver@mail.net, 9/2/2015, 10.28.34, $167.23, $10.37, {78234, "Achieving One's Fullest Potential," self-help, 145432, Brown and Gray, $29.95, $24.75, 1, $24.75 | 4782349, "Programming Server-side Solutions with Python," Programming, 63453632, Green & Yellow, $47.95, $39.99, 2, $79.98 | 2342343, "Murder at Eleven," fiction, 145432, Brown and Gray, $14.95, $12.50, 5, $62.50}. Note that information regarding specific products is repeated multiple times in the sample data set and each repeated set is separated by the "|" symbol.

Based on the facts stated above,

a. Identify the functional dependencies between the attributes.
b. Identify the reasons why this set of data is not in 3NF and indicate the normal form (if any) it is in.
c. Including all intermediate stages, organize the attributes into a set of 3NF relations.
d. Draw an ER diagram based on the normalized relations.

4-58. A bus company is responsible for offering public transportation in the suburbs of a large metropolitan area. The company has significant data management requirements: It needs to keep track of its 150 vehicles, 400 drivers, 60 bus routes, and hundreds of scheduled departures every day. In addition, it is essential for the company to know which drivers are certified to drive which buses.
The data that the company has available include the following attributes:

RouteID, RouteStartPoint, RouteEndPoint, RouteStandardDrivingTime, ScheduleDate, ScheduledDepTime, ScheduledArrTime, DriverID, DriverFName, DriverLName, DateDriverJoinedCompany, DriverDOB, VehicleID, VehicleMake, VehicleModel, VehiclePassangerCapacity, DriverCertStartDate, DriverCertEndDate.

Sample data for this set of attributes are as follows:

28, Grand Avenue, Madison Street, 38, {9/12/2015, 8.30, 9.18, 8273, Mary, Smith, 5/2/2007, 3/23/1974, 1123, GreatTrucks, CityCoach, 58, 6/10/2015, 6/9/2016 | 9/12/2015, 9.30, 10.12, 7234, John, Jones, 10/12/2011, 12/15/1991, 5673, GreatTrucks, CityCoach 2, 62, 4/12/2015, 4/11/2016 | 9/12/2015, 10.30, 11.08, 2343, Pat, Moore, 2/24/1982, 1/19/1958, 4323, PowerTransport, MidiBus, 32, 8/20/2015, 8/19/2016}

Note that the information for specific bus schedules (starting with the attribute ScheduleDate) is repeated three times in the sample data set and is separated by the "|" symbol. Also, take into account that in this case, the certification is specific to a particular vehicle driver pair.

Based on the facts stated above,

a. Identify the functional dependencies between the attributes.
b. Identify the reasons why this set of data is not in 3NF and indicate the normal form (if any) it is in.
c. Including all intermediate stages, organize the attributes into a set of 3NF relations.
d. Draw an ER diagram based on the normalized relations.
e. Based on the ER diagram you just drew and the case narrative, explore the areas in which there could be opportunities to expand the data model to achieve better tracking of the company's operations or improved clarity, such as maintaining more detailed route information.

Field Exercises

4-59. Interview system designers and database designers at several organizations. Ask them to describe the process they use for logical design. How do they transform their conceptual data models (e.g., E-R diagrams) to relational schema? What is the role of CASE tools in this process? Do they use normalization? If they do, how far in the process do they go, and for what purpose?

4-60. Obtain a common document such as a sales slip, customer invoice from an auto repair shop, credit card statement, and so on. Use the normalization steps (Steps 0 through 4) described in this chapter to convert this user view to a set of relations in third normal form. Also draw a relational schema. List several integrity rules that you would recommend to ensure the quality of the data in this application.

4-61. Using the online Appendix B, available on the book's Web site, as a resource, interview a database analyst/designer to determine whether he or she normalizes relations to higher than 3NF. Why or why not does he or she use normal forms beyond 3NF?

4-62. Find a form or report from a business organization, possibly a statement, bill, or document you have received. Draw an EER diagram of the data in this form or report. Transform the diagram into a set of 3NF relations.

References

Chouinard, P. 1989. "Supertypes, Subtypes, and DB2." *Database Programming & Design* 2,10 (October): 50–57.

Codd, E. F. 1970. "A Relational Model of Data for Large Shared Data Banks." *Communications of the ACM* 13,6 (June): 77–87.

Codd, E. F. 1990. *The Relational Model for Database Management*, Version 2. Reading, MA: Addison-Wesley.

Date, C. J. 2003. *An Introduction to Database Systems*. 8th ed. Reading, MA: Addison-Wesley.

Dutka, A. F., and H. H. Hanson. 1989. *Fundamentals of Data Normalization*. Reading, MA: Addison-Wesley.

Fleming, C. C., and B. von Halle. 1989. *Handbook of Relational Database Design*. Reading, MA: Addison-Wesley.

Hoberman, S. 2006. "To Surrogate Key or Not." *DM Review* 16,8 (August): 29.

Johnston, T. 2000. "Primary Key Reengineering Projects: The Problem" and "Primary Key Reengineering Projects: The Solution." Available at **www.information-management.com.**

Navathe, S., R. Elmasri, and J. Larson. 1986. "Integrating User Views in Database Design." *Computer* 19,1 (January): 50–62.

Further Reading

Elmasri, R., and S. Navathe. 2010. *Fundamentals of Database Systems*. 6th ed. Reading, MA: Addison Wesley.

Hoffer, J. A., J. F. George, and J. S. Valacich. 2014. *Modern Systems Analysis and Design*. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Russell, T., and R. Armstrong. 2002. "13 Reasons Why Normalized Tables Help Your Business." *Database Administrator*, April 20, 2002. Available at **http://searchoracle.techtarget.com/tip/13-reasons-why-normalized-tables-help-your-business**

Storey, V. C. 1991. "Relational Database Design Based on the Entity-Relationship Model." *Data and Knowledge Engineering* 7,1 (November): 47–83.

Web Resources

http://en.wikipedia.org/wiki/Database_normalization Wikipedia entry that provides a thorough explanation of first, second, third, fourth, fifth, and Boyce-Codd normal forms.

www.bkent.net/Doc/simple5.htm Web site that presents a summary paper by William Kent titled "A Simple Guide to Five Normal Forms in Relational Database Theory."

http://www.stevehoberman.com Web site where Steve Hoberman, a leading consultant and lecturer on database design, presents and analyzes database design (conceptual and logical) problem. These are practical (based on real experiences or questions sent to him) situations that make for interesting puzzles to solve.

www.troubleshooters.com/codecorn/norm.htm Web page on normalization on Steve Litt's site that contains various troubleshooting tips for avoiding programming and systems development problems.

JobID	JobDate Time	driverID	driver Name	taxiID	clientID	clientName	jobPickUpAddress
1	25/07/14 10.00	D1	Joe Bull	T1	C1	Anne Woo	1 Storrie Rd, Paisley
2	29/07/14 10.00	D1	Joe Bull	T1	C1	Anne Woo	1 Storrie Rd, Paisley
3	30/07/14 11.00	D2	Tom Win	T2	C1	Anne Woo	3 High Street, Paisley
4	2/08/14 13.00	D3	Jim Jones	T3	C2	Mark Tin	1A Lady Lane, Paisley
5	2/08/14 13.00	D4	Steven Win	T1	C3	John Seal	22 Red Road, Paisley
6	25/08/14 10.00	D2	Tom Win	T2	C4	Karen Bow	17 High Street, Paisley

Figure 14.21 Table displaying sample data for *FastCabs*.

14.18 Applying normalisation to 3NF on the table shown in Figure 14.21 results in the formation of the three 3NF tables shown in Figure 14.22.

(a) Identify the functional dependencies that exist between the columns of each table in Figure 14.22 and identify the primary key and any alternate and foreign key(s) (if present) for each table.

(b) Describe why storing the *FastCabs* data across three 3NF tables avoids the update anomalies described in Exercise 14.17(b).

(c) Describe how the original table shown in Figure 14.21 can be re-created through relational joins between primary key and foreign keys columns of the tables in Figure 14.22.

JobID	JobDateTime	driverID	clientID	jobPickUpAddress
1	25/07/14 10.00	D1	C1	1 Storrier Rd, Paisley
2	29/07/14 10.00	D1	C1	1 Storrier Rd, Paisley
3	30/07/14 11.00	D2	C1	3 High Street, Paisley
4	2/08/14 13.00	D3	C2	1A Lady Lane, Paisley
5	2/08/14 13.00	D4	C3	22 Red Road, Paisley
6	25/08/14 10.00	D2	C4	17 High Street, Paisley

drierID	driverName	taxiD
D1	Joe Bull	T1
D2	Tom Win	T2
D3	Jim Jones	T3
D4	Steven Win	T1

clientID	clientName
C1	Anne Woo
C2	Mark Tin
C3	John Seal
C4	Karen Bow

Figure 14.22 Tables (in 3NF) displaying sample data for *FastCabs*.

14.19 Students can lease university flats and some of the details of leases held by students for places in university flats are shown in Figure 14.23. A place number (placeNo) uniquely identifies each single room in all flats and is used when leasing a room to a student.

(a) Identify the functional dependencies that exist between the columns of the table in Figure 14.23 and identify the primary key and any alternate key(s) (if present) for the table.

(b) Describe why the table in Figure 14.23 is not in 3NF.

(c) The table shown in Figure 14.23 is susceptible to update anomalies. Provide examples of how insertion, deletion, and modification anomalies could occur on this table.

leaseNo	bannerID	placeNo	fName	lName	startDate	finishDate	flatNo	flatAddress
10003	B017706	78	Jane	Watt	01/09/2010	30/06/2011	F56	34 High Street, Paisley
10259	B017706	88	Jane	Watt	01/09/2011	30/06/2012	F78	111 Storrie Road, Paisley
10364	B013399	89	Tom	Jones	01/09/2011	30/06/2012	F78	111 Storrie Road, Paisley
10566	B012124	102	Karen	Black	01/09/2011	30/06/2012	F79	120 Lady Lane, Paisley
11067	B034511	88	Steven	Smith	01/09/2012	30/06/2013	F78	111 Storrie Road, Paisley
11169	B013399	78	Tom	Jones	01/09/2012	30/06/2013	F56	34 High Street, Paisley

Figure 14.23 Table displaying sample data for university accommodation.

14.20 Applying normalisation to 3NF on the table shown in Figure 14.23 results in the formation of the four tables shown in Figure 14.24.

(a) Identify the functional dependencies that exist between the columns of each table in Figure 14.24 and identify the primary key and any alternate and foreign key(s) (if present) for each table.

(b) Describe why storing the university accommodation data across four 3NF tables avoids the update anomalies described in Exercise 14.19(b).

(c) Describe how the original table shown in Figure 14.23 can be re-created through relational joins between primary key and foreign keys columns of the tables in Figure 14.24.

leaseNo	bannerID	placeNo	startDate	finishDate
10003	B017706	78	01/09/2010	30/06/2011
10259	B017706	88	01/09/2011	30/06/2012
10364	B013399	89	01/09/2011	30/06/2012
10566	B012124	102	01/09/2011	30/06/2012
11067	B034511	88	01/09/2012	30/06/2013
11169	B013399	78	01/09/2012	30/06/2013

flatNo	flat Address
F56	34 High Street, Paisley
F78	111 Storrie Road, Paisley
F79	120 Lady Lane, Paisley

bannerID	fName	lName
B017706	Jane	Watt
B013399	Tom	Jones
B012124	Karen	Black
B034511	Steven	Smith

placeNo	flatNo
78	F56
88	F78
89	F78
102	F79

Figure 14.24 Tables (in 3NF) displaying sample data of university accommodation.