A Guide to SQL, Ninth Edition

Chapter Two

Database Design Fundamentals

Objectives

- Understand the terms *entity*, *attribute*, and *relationship*
- Understand the terms *relation* and *relational database*
- Understand functional dependence and be able to identify when one column is functionally dependent on another
- Understand the term *primary key* and identify primary keys in tables

Objectives (continued)

- Design a database to satisfy a set of requirements
- Convert an unnormalized relation to first normal form
- Convert tables from first normal form to second normal form
- Convert tables from second normal form to third normal form

Objectives (continued)

- Create an entity-relationship diagram to represent the design of a database

Introduction

- Database design
  - Process of determining the particular tables and columns that will comprise a database
- Must understand database concepts
- Process of normalization

Database Concepts

- Entity
- Attribute
- Relationship
- Functional dependence
- Primary key

Relational Database

- A collection of tables
- Tables in TAL Distributors Database
  - Rep
  - Customer
  - Orders
  - Item
  - Order_Line

Entities, Attributes, and Relationships

- Entity (like a noun)
  - A person, place, thing, or event
- Attribute (like an adjective or adverb)
  - Property of an entity
- Relationship
  - Association between entities

Entities, Attributes, and Relationships (continued)

- One-to-many relationship
  - One rep is related to many customers
  - Implement by having a common column in two or more tables
    - REP_NUM is a column in the CUSTOMER table and the REP table
- Repeating groups
  - Multiple entries in an individual location

Entities, Attributes, and Relationships (continued)

ORDERS

ORDER_ NUM	ORDER_ DATE	CUSTOMER_ NUM	ITEM_ NUM	NUM_ ORDERED	QUOTED_ PRICE
51608	10/12/2015	126	CD33	5	$86.99
51610	10/12/2015	334	KL78 TR40	25 10	$10.95 $13.99
51613	10/13/2015	386	DL51	5	$104.95
51614	10/13/2015	260	FD11	1	$124.95
51617	10/15/2015	586	NL89 TW35	4 3	$115.99 $116.95
51619	10/15/2015	126	FD11	2	$121.95
51623	10/15/2015	586	DR67 FH24 KD34	5 12 10	$29.95 $36.95 $13.10
51625	10/16/2015	796	MT03	8	$45.79

Figure 2-2 Table with repeating groups

Entities, Attributes, and Relationships (continued)

ORDERS

ORDER_ NUM	ORDER_ DATE	CUSTOMER_ NUM	ITEM_ NUM	NUM_ ORDERED	QUOTED_ PRICE
51608	10/12/2015	126	CD33	5	$86.99
51610	10/12/2015	334	KL78	25	$10.95
51610	10/12/2015	334	TR40	10	$13.99
51613	10/13/2015	386	DL51	5	$104.95
51614	10/13/2015	260	FD11	1	$124.95
51617	10/15/2015	586	NL89	4	$115.99
51617	10/15/2015	586	TW35	3	$116.95
51619	10/15/2015	126	FD11	2	$121.95
51623	10/15/2015	586	DR67	5	$29.95
51623	10/15/2015	586	FH24	12	$36.95
51623	10/15/2015	586	KD34	10	$13.10
51625	10/16/2015	796	MT03	8	$45.79

Figure 2-3 ORDERS table without repeating groups

Entities, Attributes, and Relationships (continued)

- Relation is a two-dimensional table
  - Entries in the table are single-valued
  - Each column has a distinct name
  - All values in a column are values of the same attribute
  - The order of the columns is immaterial
  - Each row is distinct
  - The order of the rows is immaterial

Entities, Attributes, and Relationships (continued)

- Use shorthand representation to show tables and columns

REP (REP_NUM, LAST_NAME, FIRST_NAME, STREET,
CITY, STATE, POSTAL_CODE, COMMISSION, RATE)
CUSTOMER (CUSTOMER_NUM, CUSTOMER_NAME,
STREET, CITY, STATE, POSTAL_CODE, BALANCE,
CREDIT_LIMIT, REP_NUM)
ORDERS (ORDER_NUM, ORDER_DATE, CUSTOMER_NUM)
ORDER_LINE (ORDER_NUM, ITEM_NUM, NUM_ORDERED,
QUOTED_PRICE)
ITEM (ITEM_NUM, DESCRIPTION, ON_HAND, CATEGORY,
STOREHOUSE, PRICE)

Functional Dependence

- An attribute, B, is functionally dependent on another attribute (or collection), A, if a value for A determines a single value for B at any one time
- B is functionally dependent on A
- A $\rightarrow$ B
- A functionally determines B
- Cannot determine from sample data; must know the users' policies

Functional Dependence (continued)

REP

REP_ NUM	LAST_ NAME	FIRST_ NAME	STREET	CITY	STATE	POSTAL_ CODE	COMMISSION	PAY_ CLASS	RATE
15	Campos	Rafael	724 Vinca Dr.	Grove	CA	90092	$23,457.50	1	0.06
30	Gradey	Megan	632 Liatris St.	Fullton	CA	90085	$41,317.00	2	0.08
45	Tian	Hui	1785 Tyler Ave.	Northfield	CA	90098	$27,789.25	1	0.06
60	Sefton	Janet	267 Oakley St.	Congaree	CA	90097	$0.00	1	0.06

Figure 2-4 REP table with a PAY_CLASS column

Primary Keys

- Unique identifier for a table
- Column (attribute) A (or a collection of columns) is the primary key for a table (relation), R, if:
  - All columns in R are functionally dependent on A
  - No subcollection of the columns in A (assuming that A is a collection of columns and not just a single column) also has Property 1

Database Design

- Given a set of requirements that the database must support
- Requirements gathered through a process known as systems analysis

Design Method

1. Read the requirements, identify the entities (objects) involved, and name the entities
2. Identify the unique identifiers for the entities identified in step 1
3. Identify the attributes for all the entities
4. Identify the functional dependencies that exist among the attributes
5. Use the functional dependencies to identify the tables by placing each attribute with the attribute or minimum combination of attributes on which it is functionally dependent
6. Identify any relationships between tables.

Database Design Requirements

- For TAL Distributors
  - Must store data about sales reps, customers, items, orders, and order lines
  - Must enforce certain constraints; for example:
    - There is only customer per order
    - On a given order, there is at most one line item for a given item
    - The quoted price may differ from the actual price

Database Design Process Example

- Apply requirements to six steps in design method

Normalization

- Identify the existence of potential problems
- Provides a method for correcting problems
- Goal
  - Convert unnormalized relations (tables that contain repeating groups) into various types of normal forms

Normalization (continued)

- 1 NF
  - Better than unnormalized
- 2 NF
  - Better than 1 NF
- 3 NF
  - Better than 2 NF

First Normal Form

- A relation is in first normal form (1NF) if it does not contain any repeating groups
- To convert an unnormalized relation to 1NF, expand the PK to include the PK of the repeating group
  - This effectively eliminates the repeating group from the relation

First Normal Form (continued)

ORDERS

ORDER_ NUM	ORDER_ DATE	ITEM_ NUM	NUM_ ORDERED
51608	10/12/2015	CD33	5
51610	10/12/2015	KL78 TR40	25 10
51613	10/13/2015	DL51	5
51614	10/13/2015	FD11	1
51617	10/15/2015	NL89 TW35	4 3
51619	10/15/2015	FD11	2
51623	10/15/2015	DR67 FH24 KD34	5 12 10
51625	10/16/2015	MT03	8

Figure 2-7 Unnormalized order data

First Normal Form (continued)

ORDERS

ORDER_ NUM	ORDER_ DATE	ITEM_ NUM	NUM_ ORDERED
51608	10/12/2015	CD33	5
51610	10/12/2015	KL78	25
51610	10/12/2015	TR40	10
51613	10/13/2015	DL51	5
51614	10/13/2015	FD11	1
51617	10/15/2015	NL89	4
51617	10/15/2015	TW35	3
51619	10/15/2015	FD11	2
51623	10/15/2015	DR67	5
51623	10/15/2015	FH24	12
51623	10/15/2015	KD34	10
51625	10/16/2015	MT03	8

Figure 2-8 Order data converted to first normal form

Second Normal Form

- Redundancy causes problems
- Update Anomalies
  - Update
  - Inconsistent data
  - Additions
  - Deletions

Second Normal Form (continued)

ORDERS

ORDER_ NUM	ORDER_ DATE	ITEM_ NUM	DESCRIPTION	NUM_ ORDERED	QUOTED_ PRICE
51608	10/12/2015	CD33	Wood Block Set (48 piece)	5	$86.99
51610	10/12/2015	KL78	Pick Up Sticks	25	$10.95
51610	10/12/2015	TR40	Tic Tac Toe	10	$13.99
51613	10/13/2015	DL51	Classic Railway Set	5	$104.95
51614	10/13/2015	FD11	Rocking Horse	1	$124.95
51617	10/15/2015	NL89	Wood Block Set (62 piece)	4	$115.99
51617	10/15/2015	TW35	Fire Engine	3	$116.95
51619	10/15/2015	FD11	Rocking Horse	2	$121.95
51623	10/15/2015	DR67	Giant Star Brain Teaser	5	$29.95
51623	10/15/2015	FH24	Puzzle Gift Set	12	$36.95
51623	10/15/2015	KD34	Pentominoes Brain Teaser	10	$13.10
51625	10/16/2015	MT03	Zauberkasten Brain Teaser	8	$45.79

Table is in First Normal Form but not in Second Normal Form

Second Normal Form (continued)

- A relation is in second normal form (2NF) if it is in 1NF and no nonkey attribute is dependent on only a portion of the primary key

or …

- All nonkey attributes are functionally dependent on the entire primary key

Second Normal Form (continued)

- A 1NF relation with a primary key that is a single field is in 2NF automatically

Second Normal Form (continued)

ORDERS

ORDER_ NUM	ORDER_ DATE	ITEM_ NUM	DESCRIPTION	NUM_ ORDERED	QUOTED_ PRICE
51608	10/12/2015	CD33	Wood Block Set (48 piece)	5	$86.99
51610	10/12/2015	KL78	Pick Up Sticks	25	$10.95
51610	10/12/2015	TR40	Tic Tac Toe	10	$13.99
51613	10/13/2015	DL51	Classic Railway Set	5	$104.95
51614	10/13/2015	FD11	Rocking Horse	1	$124.95
51617	10/15/2015	NL89	Wood Block Set (62 piece)	4	$115.99
51617	10/15/2015	TW35	Fire Engine	3	$116.95
51619	10/15/2015	FD11	Rocking Horse	2	$121.95
51623	10/15/2015	DR67	Giant Star Brain Teaser	5	$29.95
51623	10/15/2015	FH24	Puzzle Gift Set	12	$36.95
51623	10/15/2015	KD34	Pentominoes Brain Teaser	10	$13.10
51625	10/16/2015	MT03	Zauberkasten Brain Teaser	8	$45.79

ORDERS

ORDER_ NUM	ORDER_ DATE
51608	10/12/2015
51610	10/12/2015
51613	10/13/2015
51614	10/13/2015
51617	10/15/2015
51619	10/15/2015
51623	10/15/2015
51623	10/16/2015

ITEM

ITEM_ NUM	DESCRIPTION
AH74	Patience
BR23	Skittles
CD33	Wood Block Set (48 piece)
DL51	Classic Railway Set
DR67	Giant Star Brain Teaser
DW23	Mancala
FD11	Rocking Horse
FH24	Puzzle Gift Set
KA12	Cribbage Set
KD34	Pentominoes Brain Teaser
KL78	Pick Up Sticks
MT03	Zauberkasten Brain Teaser
NL89	Wood Block Set (62 piece)
TR40	Tic Tac Toe
TW35	Fire Engine

ORDER_LINE

ORDER_ NUM	ITEM_ NUM	NUM_ ORDERED	QUOTED_ PRICE
51608	CD33	5	$86.99
51610	KL78	25	$10.95
51610	TR40	10	$13.99
51613	DL51	5	$104.95
51614	FD11	1	$124.95
51617	NL89	4	$115.99
51617	TW35	3	$116.95
51619	FD11	2	$121.95
51623	DR67	5	$29.95
51623	FH24	12	$36.95
51623	KD34	10	$13.10
51625	MT03	8	$45.79

Figure 2-10 ORDERS table converted to second normal form

Third Normal Form

- Update anomalies still possible
- Determinant
  - An attribute (or collection) that functionally determines another attribute

Third Normal Form (continued)

CUSTOMER

CUSTOMER_NUM	CUSTOMER_NAME	BALANCE	CREDIT_LIMIT	REP_NUM	LAST_NAME	FIRST_NAME
126	Toys Galore	$1,210.25	$7,500.00	15	Campos	Rafael
502	Cards and More	$5,025.75	$5,000.00	15	Campos	Rafael
713	Cress Store	$4,234.60	$10,000.00	15	Campos	Rafael
893	All Season Gifts	$935.75	$7,500.00	15	Campos	Rafael
260	Brookings Direct	$575.00	$10,000.00	30	Gradey	Megan
386	Johnson's Department Store	$879.25	$7,500.00	30	Gradey	Megan
665	Cricket Gift Shop	$678.90	$7,500.00	30	Gradey	Megan
824	Kline's	$2,475.99	$15,000.00	30	Gradey	Megan
334	The Everything Shop	$2,345.75	$7,500.00	45	Tian	Hui
440	Grove Historical Museum Store	$345.00	$5,000.00	45	Tian	Hui
586	Almondton General Store	$3,456.75	$15,000.00	45	Tian	Hui
796	Unique Gifts	$124.75	$7,500.00	45	Tian	Hui

Table is in Second Normal Form but not in Third Normal Form

Third Normal Form (continued)

- A relation is in third normal form (3NF) if it is in 2NF and the only determinants it contains are candidate keys
- Boyce-Codd normal form (BCNF) is the true name for this version of 3NF

Third Normal Form (continued)

CUSTOMER

CUSTOMER_NUM	CUSTOMER_NAME	BALANCE	CREDIT_LIMIT	REP_NUM	LAST_NAME	FIRST_NAME
126	Toys Galore	$1,210.25	$7,500.00	15	Campos	Rafael
502	Cards and More	$5,025.75	$5,000.00	15	Campos	Rafael
713	Cress Store	$4,234.60	$10,000.00	15	Campos	Rafael
893	All Season Gifts	$935.75	$7,500.00	15	Campos	Rafael
260	Brookings Direct	$575.00	$10,000.00	30	Gradey	Megan
386	Johnson's Department Store	$879.25	$7,500.00	30	Gradey	Megan
665	Cricket Gift Shop	$678.90	$7,500.00	30	Gradey	Megan
824	Kline's	$2,475.99	$15,000.00	30	Gradey	Megan
334	The Everything Shop	$2,345.75	$7,500.00	45	Tian	Hui
440	Grove Historical Museum Store	$345.00	$5,000.00	45	Tian	Hui
586	Almondton General Store	$3,456.75	$15,000.00	45	Tian	Hui
796	Unique Gifts	$124.75	$7,500.00	45	Tian	Hui

CUSTOMER

CUSTOMER_NUM	CUSTOMER_NAME	BALANCE	CREDIT_LIMIT	REP_NUM
126	Toys Galore	$1,210.25	$7,500.00	15
260	Brookings Direct	$575.00	$10,000.00	30
334	The Everything Shop	$2,345.75	$7,500.00	45
386	Johnson's Department Store	$879.25	$7,500.00	30
440	Grove Historical Museum Store	$345.00	$5,000.00	45
502	Cards and More	$5,025.75	$5,000.00	15
586	Almondton General Store	$3,456.75	$15,000.00	45
665	Cricket Gift Shop	$678.90	$7,500.00	30
713	Cress Store	$4,234.60	$10,000.00	15
796	Unique Gifts	$124.75	$7,500.00	45
824	Kline's	$2,475.99	$15,000.00	30
893	All Season Gifts	$935.75	$7,500.00	15

REP

REP_NUM	LAST_NAME	FIRST_NAME
15	Campos	Rafael
30	Gradey	Megan
45	Tian	Hui
60	Sefton	Janet

Figure 2-12 CUSTOMER table converted to third normal form

Diagrams for Database Design

- Graphical illustration
- Entity-relationship (E-R) diagram
  - Rectangles represent entities
  - Arrows represent relationships

Diagrams for Database Design (continued)

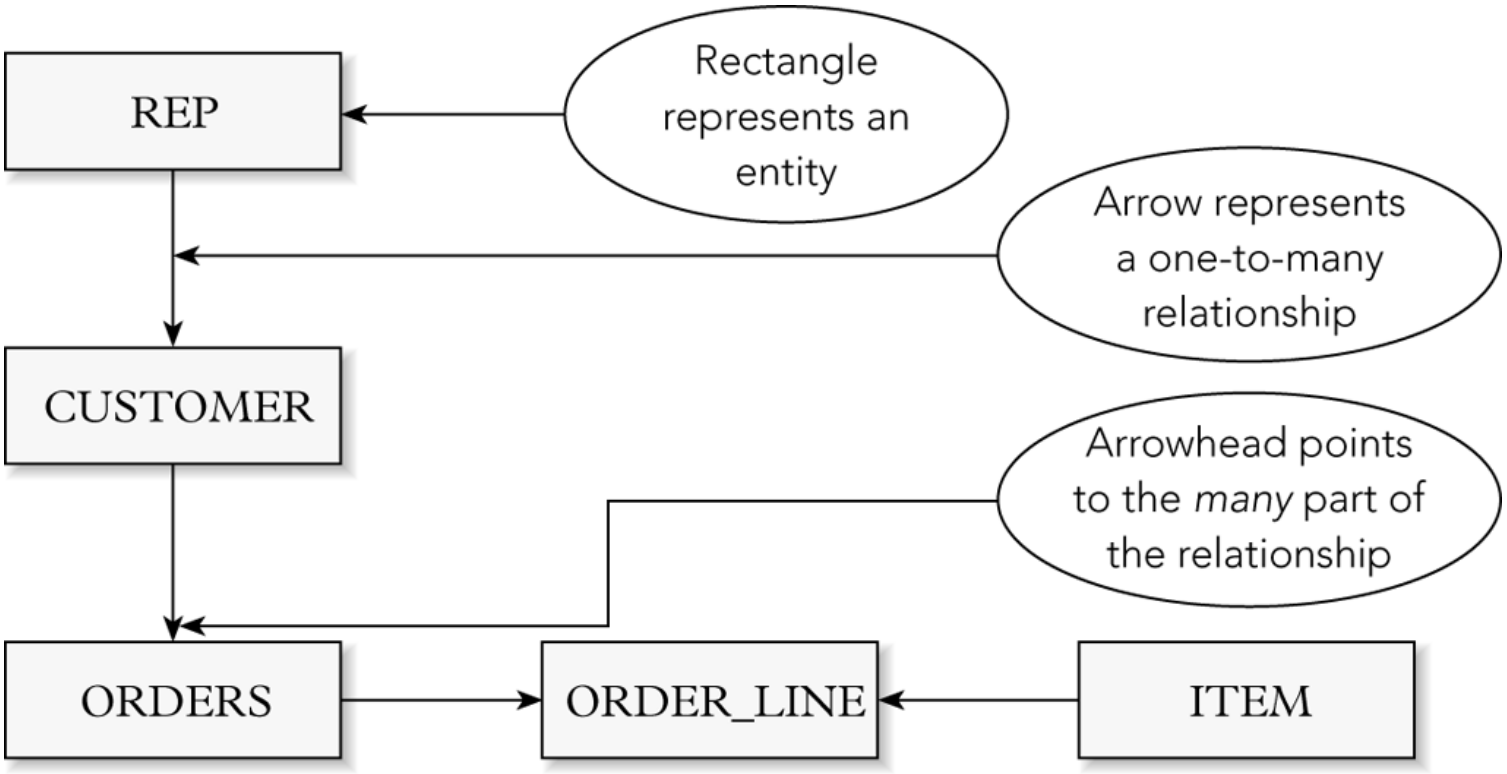

Figure 2-13 E-R diagram for the TAL Distributors database with rectangles and arrows

Diagrams for Database Design (continued)

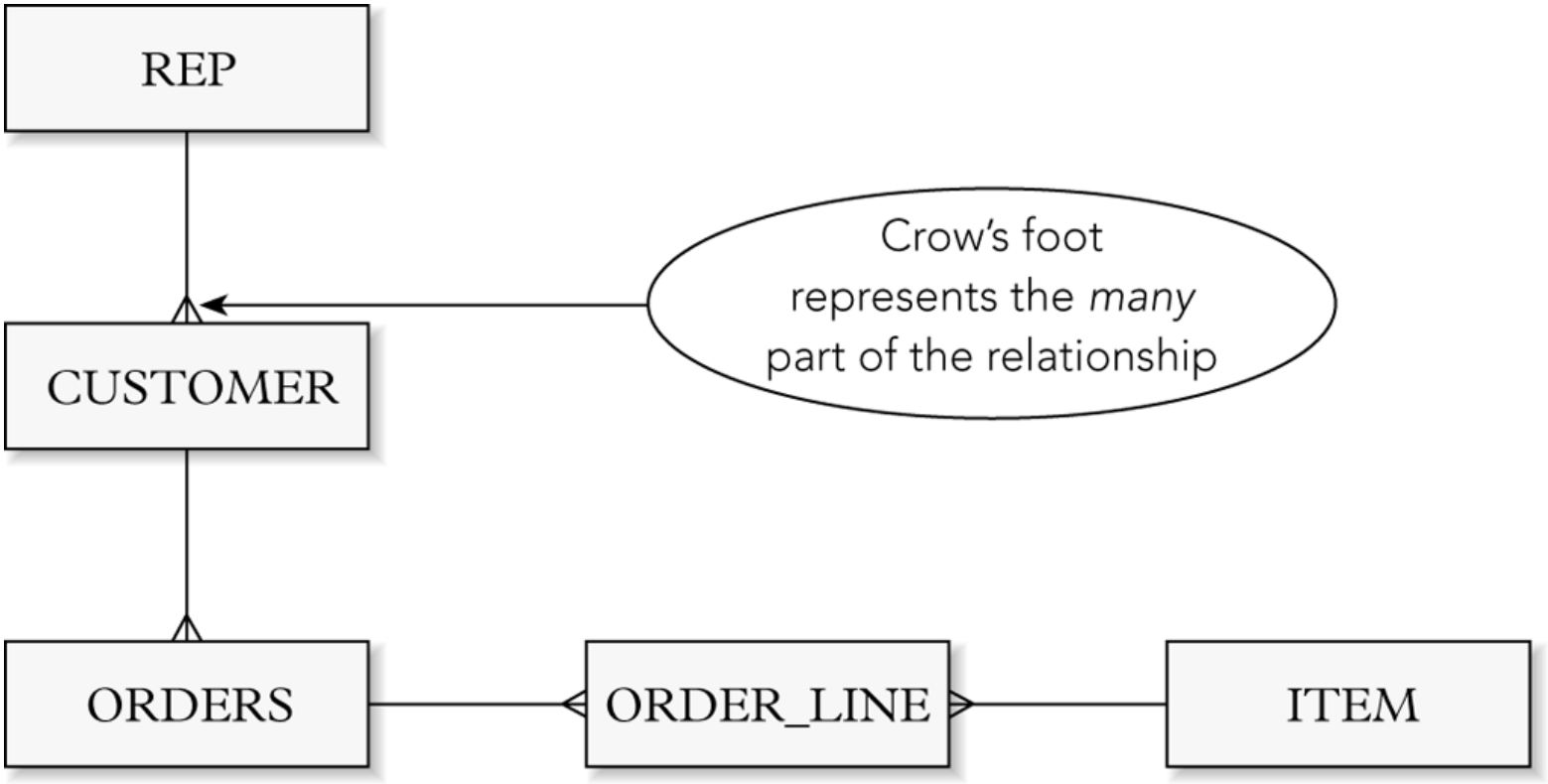

Figure 2-14 E-R diagram for the TAL Distributors database with a crow's foot

Diagrams for Database Design (continued)

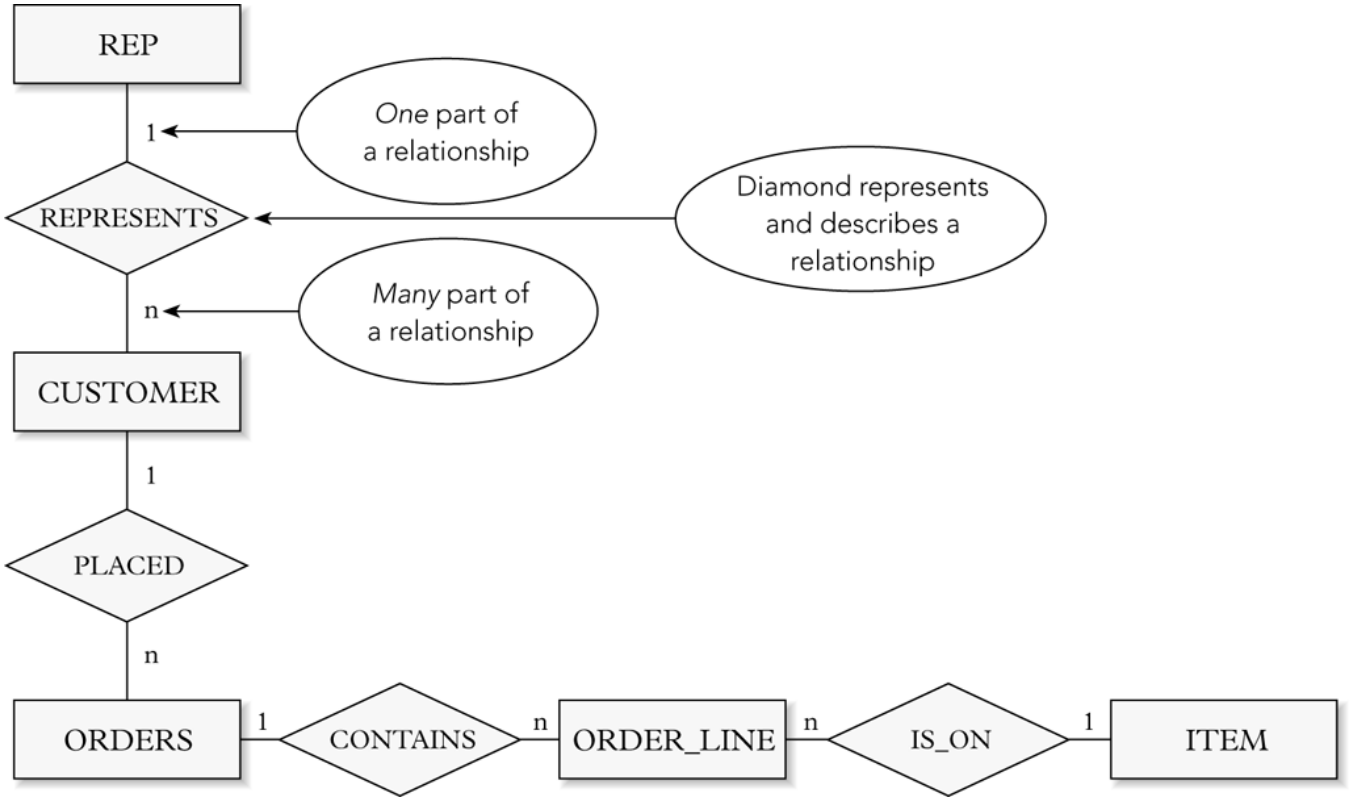

Figure 2-15 E-R diagram for the TAL Distributors database with named relationships

Summary

- Definition of entity
- Definition of attribute
- Definition of relationship
- Definition of relation
- Definition of functional dependence
- Definition of primary key
- Database design method

Summary (continued)

- Normalization
- Unnormalized (repeating groups)
- First normal form (INF)
- Second normal form (2NF)
- Third normal form (3NF)
- Entity-relationship diagram (E-R diagram)