Chapter 9

How to design a database
Objectives

Applied

• Given the specifications for a data structure, identify the tables, columns, keys, relationships, and indexes for the structure.

• Given the tables for an unnormalized database, normalize the structure to the third normal form. We’ll do this later.

Knowledge

• Give three criteria for when a column should be indexed.

• Describe referential integrity.

• Explain how Oracle uses declarative referential integrity to prevent deletion, insertion, and update problems.

• Explain how normalizing a database to the third normal form affects database performance. (Later)
A database system is modeled after a real-world system.
The six basic steps for designing a data structure

Step 1: Identify the data elements
Step 2: Subdivide each element into its smallest useful components
Step 3: Identify the tables and assign columns
Step 4: Identify the primary and foreign keys
Step 5: Review whether the data structure is normalized (later)
Step 6: Identify the indexes (later)
An invoice that’s used to identify data elements

---

### Acme Fabrication, Inc.

*Custom Contraptions, Contrivances and Confabulations*

<table>
<thead>
<tr>
<th>1234 West Industrial Way</th>
<th>East Los Angeles</th>
<th>California 90022</th>
</tr>
</thead>
<tbody>
<tr>
<td>800.555.1212</td>
<td>fax 562.555.1213</td>
<td><a href="http://www.acmefabrication.com">www.acmefabrication.com</a></td>
</tr>
</tbody>
</table>

**Invoice Number:** I01-1088  
**Invoice Date:** 10/05/06  
**Terms:** Net 30

<table>
<thead>
<tr>
<th>Part No.</th>
<th>Qty</th>
<th>Description</th>
<th>Unit Price</th>
<th>Extension</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUST345</td>
<td>12</td>
<td>Design service, hr</td>
<td>100.00</td>
<td>1200.00</td>
</tr>
<tr>
<td>457332</td>
<td>7</td>
<td>Baling wire, 25x3ft roll</td>
<td>79.90</td>
<td>559.30</td>
</tr>
<tr>
<td>50173</td>
<td>4375</td>
<td>Duct tape, black, yd</td>
<td>1.09</td>
<td>4768.75</td>
</tr>
<tr>
<td>328771</td>
<td>2</td>
<td>Rubber tubing, 100ft roll</td>
<td>4.79</td>
<td>9.58</td>
</tr>
<tr>
<td>CUST281</td>
<td>7</td>
<td>Assembly, hr</td>
<td>75.00</td>
<td>525.00</td>
</tr>
<tr>
<td>CUST917</td>
<td>2</td>
<td>Testing, hr</td>
<td>125.00</td>
<td>250.00</td>
</tr>
</tbody>
</table>

**Sales Tax**

| 245.20 |

**Total**

$7,557.83

**PLEASE PAY THIS AMOUNT**

*Thanks for your business!*

---

Your salesperson: Ruben Goldberg, ext 4512  
Accounts receivable: Inigo Jones, ext 4901

---

Murach’s Oracle SQL and PL/SQL, C9  
© 2014, Mike Murach & Associates, Inc.  
Slide 5
The data elements on the invoice document

Vendor name
Vendor address
Vendor phone number
Vendor fax number
Vendor web address
Invoice number
Invoice date
Invoice terms
Item part number
Item quantity
Item description
Item unit price

Item extension
Vendor sales contact name
Vendor sales contact extension
Vendor AR contact name
Vendor AR contact extension
Invoice total
A name that’s divided into first and last names

Vendor sales contact name
Ruben Goldberg

Vendor sales contact first name
Ruben

Vendor sales contact last name
Goldberg
An address that’s divided into its components

Vendor address
1234 West Industrial Way, East Los Angeles, California 90022

Street and number
1234 West Industrial Way

City
East Los Angeles

State
California

Zip
90022
## Possible tables and columns for an A/P system

<table>
<thead>
<tr>
<th>Vendors</th>
<th>Invoices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vendor name</td>
<td>Invoice number*</td>
</tr>
<tr>
<td>Vendor address</td>
<td>Invoice date</td>
</tr>
<tr>
<td>Vendor city</td>
<td>Terms*</td>
</tr>
<tr>
<td>Vendor state</td>
<td>Invoice total</td>
</tr>
<tr>
<td>Vendor zip code</td>
<td>Payment date</td>
</tr>
<tr>
<td>Vendor phone number</td>
<td>Payment total</td>
</tr>
<tr>
<td>Vendor fax number</td>
<td>Invoice due date</td>
</tr>
<tr>
<td>Vendor web address</td>
<td>Credit total</td>
</tr>
<tr>
<td>Vendor contact first name</td>
<td>Account number*</td>
</tr>
<tr>
<td>Vendor contact last name</td>
<td></td>
</tr>
<tr>
<td>Vendor contact phone</td>
<td></td>
</tr>
<tr>
<td>Vendor AR first name</td>
<td></td>
</tr>
</tbody>
</table>
### Tables and columns for an A/P system (cont.)

<table>
<thead>
<tr>
<th>Vendors (cont.)</th>
<th>Invoice line items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vendor AR last name</td>
<td>Invoice number*</td>
</tr>
<tr>
<td>Vendor AR phone</td>
<td>Item part number</td>
</tr>
<tr>
<td>Terms*</td>
<td>Item quantity</td>
</tr>
<tr>
<td>Account number*</td>
<td>Item description</td>
</tr>
<tr>
<td></td>
<td>Item unit price</td>
</tr>
<tr>
<td></td>
<td>Item extension</td>
</tr>
<tr>
<td></td>
<td>Account number*</td>
</tr>
<tr>
<td></td>
<td>Sequence number</td>
</tr>
</tbody>
</table>
The notation

- Data elements that were previously identified but aren’t needed are crossed out.
- Data elements that were added are displayed in italics.
- Data elements that are related to two or more entities are followed by an asterisk.
The relationships between the tables

This shows two many-to-one relationships, with a “crow’s foot” at the many end.
Two tables with a many-to-many relationship

- Note that R&G would show “memberships” as a diamond implementing the many-to-many relationship here.
- Murach is showing what tables are needed.
- Murach says the linking table doesn’t need a primary key (pg. 294-295), but R&G (and I agree) use a PK to make sure no duplicate links occur. See enrolled, pg. 71 for example, and in our createdb.sql.
Two tables with a one-to-one relationship

employees

- employee_id
- first_name
- last_name

employee_photos

- employee_id
- employee_photo
Operations that violate referential integrity

Deleting a row from the primary key table
If the foreign key table contains one or more rows related to the deleted row

Inserting a row in the foreign key table
If the foreign key value doesn’t have a matching primary key value in the related table

Updating the value of a foreign key
If the new foreign key value doesn’t have a matching primary key value in the related table

Updating the value of a primary key
If the foreign key table contains one or more rows related to the row that’s changed
Terms to know

- Entity
- Attribute
- Instance
- Entity-relationship (ER) modeling
- Referential integrity
- Declarative referential integrity (DRI)
- Foreign key constraints