Database Design

Practical Database Design for Relational Database Management Systems

Overview

- A little background and terminology:
  - What is a relational database?
  - What is a primary key?
  - What is a foreign key?
- Things to know about designing a database:
  - The normalization process and how/why use it
  - Relating tables
  - Types of relationships

© 2007 Instructional Services at KU Libraries, The University of Kansas

Relational Database Management System

- Collection of information organized in tables
  - Tables are also “relations”
- Tables are constructed and associated to each other through shared fields–“common” fields
  - Fields are also “columns” or “attributes”
- A set of attributes comprises a record
  - Records are also “rows” or “tuples”
- Tables are related through common fields designated as primary and foreign keys
- Allow us to find, update, and delete data quickly, and help to ensure accuracy

Primary and Foreign Key Fields

- Primary Key
  - Primary key fields must be unique and cannot contain a null value.
  - Each table should have a primary key field.
  - Concatenated keys: using more than one field as a primary key field.
- Foreign Key: Fields in a table that refer to the primary key in another table
  - The data in this field must exactly match data contained in the primary key field.

© 2007 Instructional Services at KU Libraries, The University of Kansas

What is Normalization?

- The process by which we efficiently organize data to achieve these goals:
  - Eliminating redundancy
  - Ensuring data is stored in the correct table
  - Eliminating need for restructuring database when data is added.
- Five levels of normal form
  - In order to achieve one level of normal form, each previous level must be met

Third normal form is sufficient for most typical database applications.

First Normal Form (1NF)

- There are no repeating or duplicate fields.
- Each cell contains only a single value.
- Each record is unique.
  - Identified by primary key

© 2007 Instructional Services at KU Libraries, The University of Kansas

Example

item	colors	price	tax
T-shirt	red, blue	12.00	0.60
polo	red, yellow	12.00	0.60
T-shirt	red, blue	12.00	0.60
sweatshirt	blue, black	25.00	1.25

Table is not in first normal form because:

- Multiple items in color field
- Duplicate records / no primary key

Example

item	color	price	tax
T-shirt	red	12.00	0.60
T-shirt	blue	12.00	0.60
polo	red	12.00	0.60
polo	yellow	12.00	0.60
sweatshirt	blue	25.00	1.25
sweatshirt	black	25.00	1.25

Table is now in first normal form.

© 2007 Instructional Services at KU Libraries, The University of Kansas

Second Normal Form (2NF)

- All non-key fields depend on all components of the primary key.
  - Guaranteed when primary key is a single field.

Example

item	color	price	tax
T-shirt	red	12.00	0.60
T-shirt	blue	12.00	0.60
polo	red	12.00	0.60
polo	yellow	12.00	0.60
sweatshirt	blue	25.00	1.25
sweatshirt	black	25.00	1.25

Table is not in second normal form because:

- **price** and **tax** depend on **item**, but not **color**

© 2007 Instructional Services at KU Libraries, The University of Kansas

Example

item	color
T-shirt	red
T-shirt	blue
polo	red
polo	yellow
sweatshirt	blue
sweatshirt	black

item	price	tax
T-shirt	12.00	0.60
polo	12.00	0.60
sweatshirt	25.00	1.25

Tables are now in second normal form.

Third Normal Form (3NF)

- No non-key field depends upon another.
  - All non-key fields depend only on the primary key.

© 2007 Instructional Services at KU Libraries, The University of Kansas

Example

item	color
T-shirt	red
T-shirt	blue
polo	red
polo	yellow
sweatshirt	blue
sweatshirt	black

item	price	tax
T-shirt	12.00	0.60
polo	12.00	0.60
sweatshirt	25.00	1.25

Tables are not in third normal form because:

- **tax** depends on **price**, not **item**

9/19/07 13

Example

item	color
T-shirt	red
T-shirt	blue
polo	red
polo	yellow
sweatshirt	blue
sweatshirt	black

item	price
T-shirt	12.00
polo	12.00
sweatshirt	25.00

price	tax
12.00	0.60
25.00	1.25

Tables are now in third normal form.

9/19/07 14

© 2007 Instructional Services at KU Libraries, The University of Kansas

Another Example

Name	Assignment 1	Assignment 2
Jeff Smith	Article Summary	Poetry Analysis
Nancy Jones	Article Summary	Reaction Paper
Jane Scott	Article Summary	Poetry Analysis

Table is not in first normal form because:

- Assignment field repeating
- First and last name in one field
- No (guaranteed unique) primary key field

Another Example

Assignment ID	Description
1	Article Summary
2	Poetry Analysis
3	Reaction Paper

Student ID	First Name	Last Name
1	Jeff	Smith
2	Nancy	Jones
3	Jane	Scott

Assignment ID	Student ID
1	1
1	2
1	3
2	1
2	3
3	2

Tables are in third normal form.

© 2007 Instructional Services at KU Libraries, The University of Kansas

Relationships

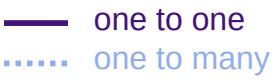
- Relationships are created between tables using the primary key field and a foreign key field
  - One to One Relationship
    - One record in a table relates to one record in another table
  - One to Many Relationship
    - One record in a table can relate to many records in another table
  - Many to Many Relationship
    - Many records in one table can relate to many records in another table

Relationships in First Example

item	color
T-shirt	red
T-shirt	blue
polo	red
polo	yellow
sweatshirt	blue
sweatshirt	black

item	price
T-shirt	12.00
polo	12.00
sweatshirt	25.00

price	tax
12.00	0.60
25.00	1.25



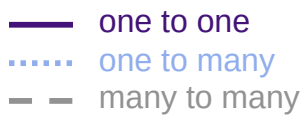
© 2007 Instructional Services at KU Libraries, The University of Kansas

Relationships in Second Example

Assignment ID	Description
1	Article Summary
2	Poetry Analysis
3	Reaction Paper

Assignment ID	Student ID
1	1
1	2
1	3
2	1
2	3
3	2

Student ID	First Name	Last Name
1	Jeff	Smith
2	Nancy	Jones
3	Jane	Scott



9/19/07 19

Bibliography

- Hernandez, Michael J. Database Design for Mere Mortals. San Francisco: Addison-Wesley, 1997.
- Chapple, Mike. " Database Normalization Basics." 5 August 2001. Online. Internet. Available http://databases.about.com/library/weekly/aa080501a.htm
- Association for Geographic Information. GIS Dictionary. 1999. Online. Internet. Available http://www.geo.ed.ac.uk/agidexe/term?821
- Wise, Barry. "Database Normalization and Design Techniques." 1 August 2000. 6 pp. Online. Internet. Available http://www.phpbuilder.com/columns/barry20000731.php3

9/19/07 20

© 2007 Instructional Services at KU Libraries, The University of Kansas

Further Reading

- Harrington, Jan L. Relational Database Design Clearly Explained. San Diego: Academic Express, 1998.
- Chapple, Mike. “Choosing a Database Product.” 6 May 2001. Online. Internet. Available http://databases.about.com/library/weekly/aa050601a.htm
- Gilmore, W.J. “Introduction to Database Normalization.” 27 November 2000. Online. Internet. Available http://www.devshed.com/Server_Side/MySQL/Normal/Normal1/page1.html

© 2007 Instructional Services at KU Libraries, The University of Kansas