SQL FOR DB2

Chapter 5
Logical Data Modeling

Chapter Objectives

- Explain logical data modeling
- Explain why data modeling is important
- Identify and explain the steps for developing a data model

Introduction to Logical data Modeling

- Involves talking to end users and researching manuals, forms, and other documents that the organization uses
- This information is first collected from the individual user or department perspective and then integrated into a unified, organization-wide data model
- During the integration, conflicts, redundancy, and ambiguity need to be resolved
- The logical data model should describe entities, attributes, relationships, and integrity constraints
- The data model is stored in a data dictionary, which is a collection of documents, tables, diagrams, and other items that contain all the descriptions that are part of the model
- Provides the basis for subsequent physical database design

Steps for Developing a Data Modeling

- Step 1: Identify Entities(Tables)
- Step 2: Determine Relationships Between Entities
- Step 3: Determine Attributes(Columns)
- Step 4: Determine the Primary Key for Each Entity
- Step 5: Determine Foreign Key Rules
- Step 6: Eliminate Redundancy in Tables using Normalization
- Step 7: Determine Additional Integrity Constraints
- Step 8: Determine Security Rules

Step 1: Identify Entities(Tables)

- Entities are things about which facts needs to captured and stored
- Entities are usually nouns that could fall into the following categories: objects, events, places, and people
- Entities should have more than one row or instance
- When an entity is discovered, be sure to enter it in the data dictionary
- As entities are identified, also identify entity hierarchies

Step 2: Determine Relationships

- Entities do not exist in isolation; they are interrelated
- When a relationship is identified, document it and give it a name
- Relationships can be recorded in textual, table form or an Entity Relationship Diagram(ERD)
- In addition, for each relationship, its cardinality should be determined:
  - One-to-one
  - One-to-many
  - Many-to-many
- When identifying relationships, also determine the degree of each relationship

Step 3: Determine Attributes(Columns)

- Determine which attributes are single-valued and which are multi-valued
- Determine which attributes are time-dependent
- Document each attribute with the following:
  - Official name used in the data model
  - Short description
  - The source in the business process
  - Allowable values (domain)
  - Required or optional
  - Default value
  - Primary key, foreign key or candidate key

Step 4: Determine the Primary Key

- Each primary and candidate key must meet the following criteria:
  - Candidate key values must be unique for each row in the table
  - Candidate key values must never be missing or incomplete for a row
  - Each candidate key must use no columns other than those necessary to identify a row uniquely
- A primary key should also meet the following criteria:
  - It should be meaningless, other than as an identifier
  - A row's primary key value should never change
  - There should be no practical limit to the number of primary key values available
  - Only one primary key should be specified for each table

Step 5: Determine Foreign Key Rules

- The delete-update-insert rules between the primary and foreign keys in each relationship specify the way referential integrity will be enforced
- Delete-Update-Insert rules:
  - Reject the delete, update or insert
  - Handle the delete, update or insert with a custom procedure
  - Cascade the delete or update to all related dependent tables
  - Set the related dependent foreign key value to null or default value
  - On insert create a default row in the target table

Step 6: Normalize the Tables

- Make certain that the entities do not contain redundant facts
- One approach to eliminating redundancy is to check each table to see that is does not violate any of the rules of normalization
- Another way to eliminate redundancy is to check all columns to see whether they can be derived from other columns

Step 7: Determine Integrity Constraints

- Business rules must be accurate and complete, or the database may become contaminated
- Determine the domain of each attribute: allowable values and allowable operations
  - Format or picture checks
  - Data type checks
  - Range checks
  - Limit checks
  - Presence checks
  - Spelling checks
  - Consistency Checks
  - Cross-system Consistency Checks

Step 8: Determine Security Rules

- *Value-independent* rules, which are restrictions on tables, views, and columns based solely on the type of data, not on particular values
- *Value-dependent* rules, which are based on specific values
- *Statistical* rules, which limit the types of inferences that can be made about values in the database using statistical functions such as Sum and Count
- *Context-dependent* rules, which are defined in terms of functions or system values, such as time of day

Chapter Summary

- To complete a data model, you normally work on several steps at the same time, and you may need to repeat some steps to refine the model.
- A process model is also developed in parallel with the data model. The process model references the data model, and vice versa.
- The following steps are a guideline for developing a data model:
  - Step 1. Identify Entities (Tables).
  - Step 2. Determine relationships between entities.
  - Step 3. Determine attributes (columns).
  - Step 4. Determine the primary key for each entity.
  - Step 5. Determine foreign key rules for each relationship.
  - Step 6. Eliminate redundancy in tables using normalization.
  - Step 7. Determine additional integrity constraints.
  - Step 8. Determine security rules.