Computer Science 202

Database Systems:

Database Design

Objectives

- To learn what an information system is.
- To learn what a Database Life Cycle (DBLC) is.
- To learn what a Systems Development Life Cycle (SDLC) is.
- To demonstrate the relationship between the DBLC and the SDLC.
- To learn how database design fits into the design of an information system.

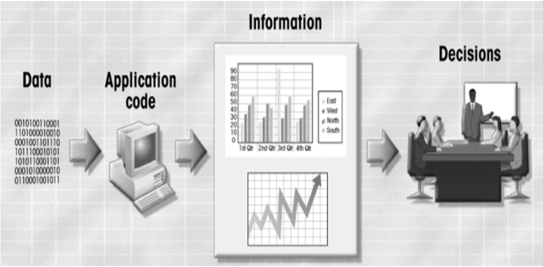
Data *vs.* Information

- Data
  - Raw facts stored in a database
  - Little to no meaning in isolation
  - Require organisation in order to be useful
- Information
  - Used for decision making
  - Result of a transformation of data
  - Consists of data processed and organised in a useful way – e.g. tables, graphs

The Information System

- A system containing a repository of data, which is able to process the data and produce informational output
- Manages data collection, storage and retrieval
- Also consists if the people and processes related to the management of data

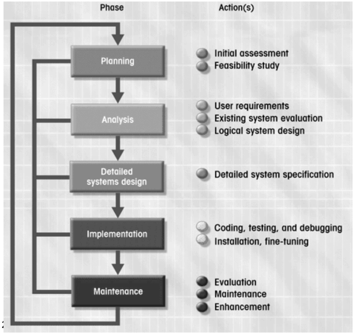
Data processing



Systems Analysis

- **Systems Analysis** is the process that establishes the need for and the extent of a Information System
- The process of creating an information system is **Systems Development**
- Database design and development are an integral part of the systems development process
- Systems development follows a lifecycle (SDLC)

Systems Development Life Cycle



SDLC

- Planning
- Analysis
- Detailed systems design
- Implementation
- Maintenance

SDLC - Planning

- Planning of the project
- Definition of
  - Scope
  - Budget
  - Resources

SDLC - Analysis

- Break down the project into manageable components
- Identify specifics regarding components
- Gather Requirements

SDLC –Detailed Design

- In detail specification of
  - HCI components
  - Data structures
  - API
  - Internal Component functionality
- Testing Framework designed

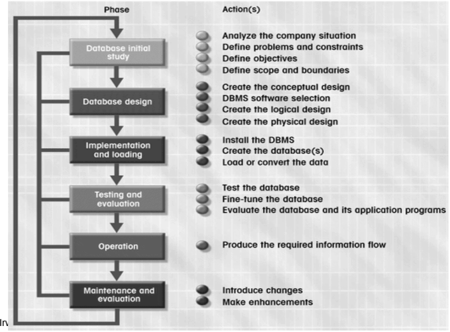
SDLC - Implementation

- Implementation of the System
- Can be on a component by component level (eXtreme Programming)
- Implementation includes development of
  - Test harnesses
  - Unit Tests
  - System Tests

SDLC – Maintenance

- Corrective process
- Remedy errors found in testing
- Iterative
- Can also follow post deployment

DB Development Lifecycle



Phase
Action(s)

Database initial study
- Analyze the company situation
- Define problems and constraints
- Define objectives
- Define scope and boundaries

Database design
- Create the conceptual design
- DBMS software selection
- Create the logical design
- Create the physical design

Implementation and loading
- Install the DBMS
- Create the database(s)
- Load or convert the data

Testing and evaluation
- Test the database
- Fine-tune the database
- Evaluate the database and its application programs

Operation
- Produce the required information flow

Maintenance and evaluation
- Introduce changes
- Make enhancements

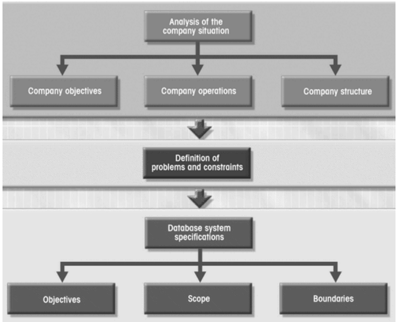
DBLC

- Initial study
- Database Design
  - Conceptual Design
    - Data analysis
    - ERD construction and data normalisation
    - Data model verification
    - ER Verification
  - DBMS software selection
  - Logical Design
  - Physical Design
- Implementation and Loading
- Testing and evaluation
- Operation
- Maintenance and Evaluation

DBLC – Initial Design

- Analyse company situation
  - Current Operating environment
  - Organisational Structure
  - Information Flow
- Define the problem
- Define scope and constraints of project
- Define design objectives

DBLC Initial Design



Analysis of the company situation

Company objectives
Company operations
Company structure

Definition of problems and constraints

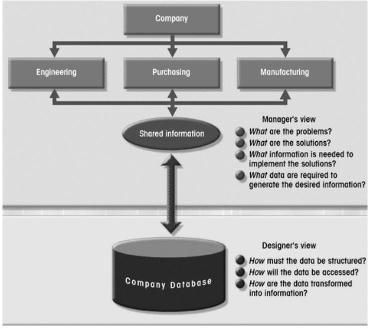
Database system specifications

Objectives
Scope
Boundaries

DBLC – DB DESIGN

- Database Design
  - Conceptual Design
  - DBMS software selection
  - Logical Design
  - Physical Design
- Critical part of whole process
- Focus on the Data requirements
- Final design should satisfy the requirements

DBLC – DB DESIGN


Company
Engineering
Purchasing
Manufacturing
Shared information
Manager's view
What are the problems?
What are the solutions?
What information is needed to implement the solutions?
What data are required to generate the desired information?
Company Database
Designer's view
How must the data be structured?
How will the data be accessed?
How are the data transformed into information?

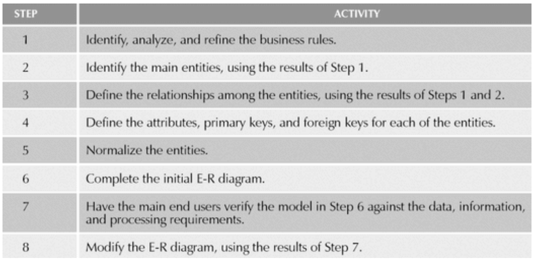
DBLC – DB DESIGN Conceptual

- Develop a high-level abstraction of the data model
- Conceptual Design consists of four distinct parts
  - Data analysis
  - ERD construction and data normalisation
  - Data model verification
  - ER Verification

Data Analysis and Requirements

- Focus on:
  - Information needs
  - Information users
  - Information sources
  - Information constitution
- Data sources
  - Developing and gathering end-user data views
  - Direct observation of current system
  - Interfacing with systems design group
- Business rules

Entity Relationship Modeling and Normalization



STEP	ACTIVITY
1	Identify, analyze, and refine the business rules.
2	Identify the main entities, using the results of Step 1.
3	Define the relationships among the entities, using the results of Steps 1 and 2.
4	Define the attributes, primary keys, and foreign keys for each of the entities.
5	Normalize the entities.
6	Complete the initial E-R diagram.
7	Have the main end users verify the model in Step 6 against the data, information, and processing requirements.
8	Modify the E-R diagram, using the results of Step 7.

E-R Modeling is Iterative

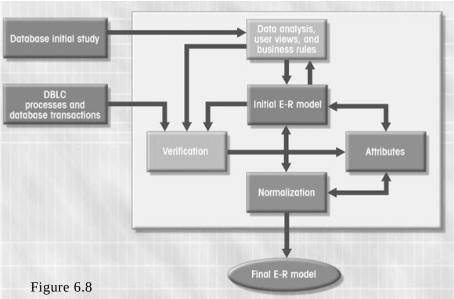
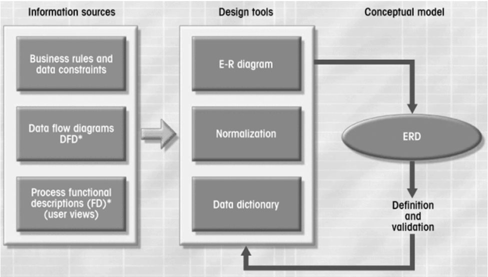


Figure 6.8

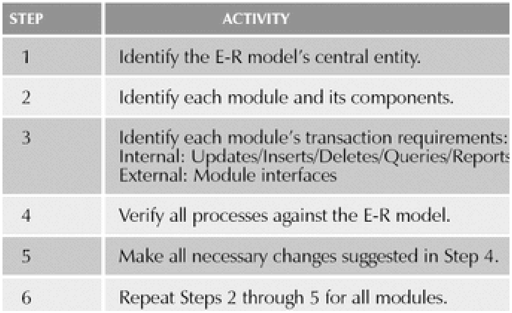
Concept Design: Tools and Sources



Data Model Verification

- E-R model is verified against proposed system processes
  - End user views and required transactions
  - Access paths, security, concurrency control
  - Business-imposed data requirements and constraints
- Reveals additional entity and attribute details
- Define major components as modules
  - Cohesively
  - Coupling

E-R Model Verification Process



STEP	ACTIVITY
1	Identify the E-R model's central entity.
2	Identify each module and its components.
3	Identify each module's transaction requirements: Internal: Updates/Inserts/Deletes/Queries/Reports External: Module interfaces
4	Verify all processes against the E-R model.
5	Make all necessary changes suggested in Step 4.
6	Repeat Steps 2 through 5 for all modules.

Iterative Process of Verification

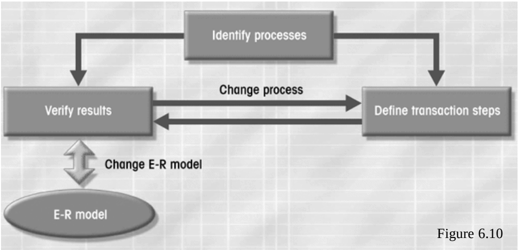


Figure 6.10

Distributed Database Design

- Design portions in different physical locations
- Development of data distribution and allocation strategies

DBLC – DB DESIGN Software Selection

- Selection of the correct software is critical
- For possible candidates, both advantages and disadvantages need to be compared
- Common influencing factors:
  - Cost – purchase and ongoing support
  - DBMS features and management tools
  - Hardware/Platform requirements
  - Underlying Model (e.g.. OOD , RD)

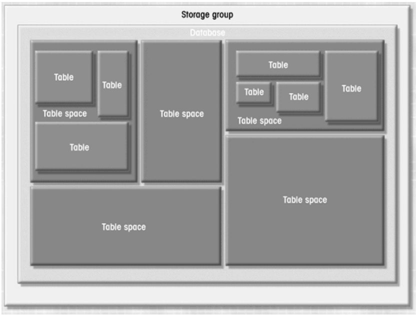
DBLC – DB DESIGN Logical Design

- Process of converting the conceptual model into an appropriate form for the DBMS chosen
- Maps conceptual objects to logical DB constructs (tables and relations)
- Common components:
  - Tables
  - Indexes
  - Views
  - Access Control

DBLC – DB DESIGN Physical Design

- Physical layout of data on partitions
- Highly technical
- Can have an big impact on performance
- Most designers favour transparency of physical layout
- Complicated on distributed systems

Physical Design



DBLC - Implementation and Loading

- Creation of special storage-related constructs to house end-user tables
- Data loaded/imported into tables
- Other issues to be considered
  - Performance
  - Security
  - Backup and recovery
  - Integrity
  - Company standards
  - Concurrency controls

DBLC - Testing and evaluation

- Database is tested and fine-tuned against a number of criteria, including
  - performance, integrity, concurrent access security constraints
- Done in parallel with application development
- Corrective actions taken to correct problems:
  - Fine-tuning based on reference manuals
  - Modification of physical design
  - Modification of logical design
  - Upgrade or change DBMS software or hardware

DBLC - Operation

- Database considered operational
- 'Go-live' procedures
- Possible running in parallel with old systems
- Starts process of system evaluation
- Unforeseen problems may surface
- Demand for change is constant and must be managed
- Importance of adherence to the original specification.

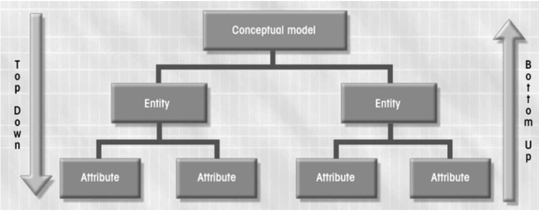
DBLC - Maintenance and Evaluation

- Preventative maintenance
- Corrective maintenance
- Adaptive maintenance
- Assignment of access permissions
- Generation of database access statistics to monitor performance
- Periodic security audits based on system-generated statistics
- Periodic system usage-summaries

Design Strategies

- Top-down
  - Identify the data sets
  - Define the data elements
- Bottom-up
  - Identify data elements
  - Group elements into data sets

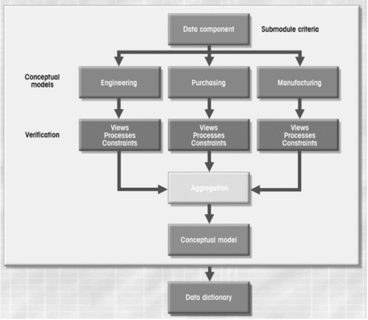
Design Strategies



Centralised vs. Decentralised Design

- Centralized design
  - Typical of simple databases
  - Conducted by single person or small team
- Decentralized design
  - Larger numbers of entities and complex relations
  - Spread across multiple sites
  - Developed by teams

Decentralised Design


Data component
Submodule criteria
Conceptual models
Engineering
Purchasing
Manufacturing
Verification
Views Processes Constraints
Views Processes Constraints
Views Processes Constraints
Aggregation
Conceptual model
Data dictionary