Chapter 2

Database Environment

Chapter 2 - Objectives

- **Purpose of three-level database architecture**
- **Contents of external, conceptual, and internal levels**
- **Purpose of external/conceptual and conceptual/internal mappings**
- **Meaning of logical and physical data independence**
- **Distinction between DDL and DML**

Chapter 2 - Objectives

- **A classification of data models.**
- **Purpose/importance of conceptual modeling.**
- **Typical functions and services a DBMS should provide.**
- **Function and importance of system catalog.**

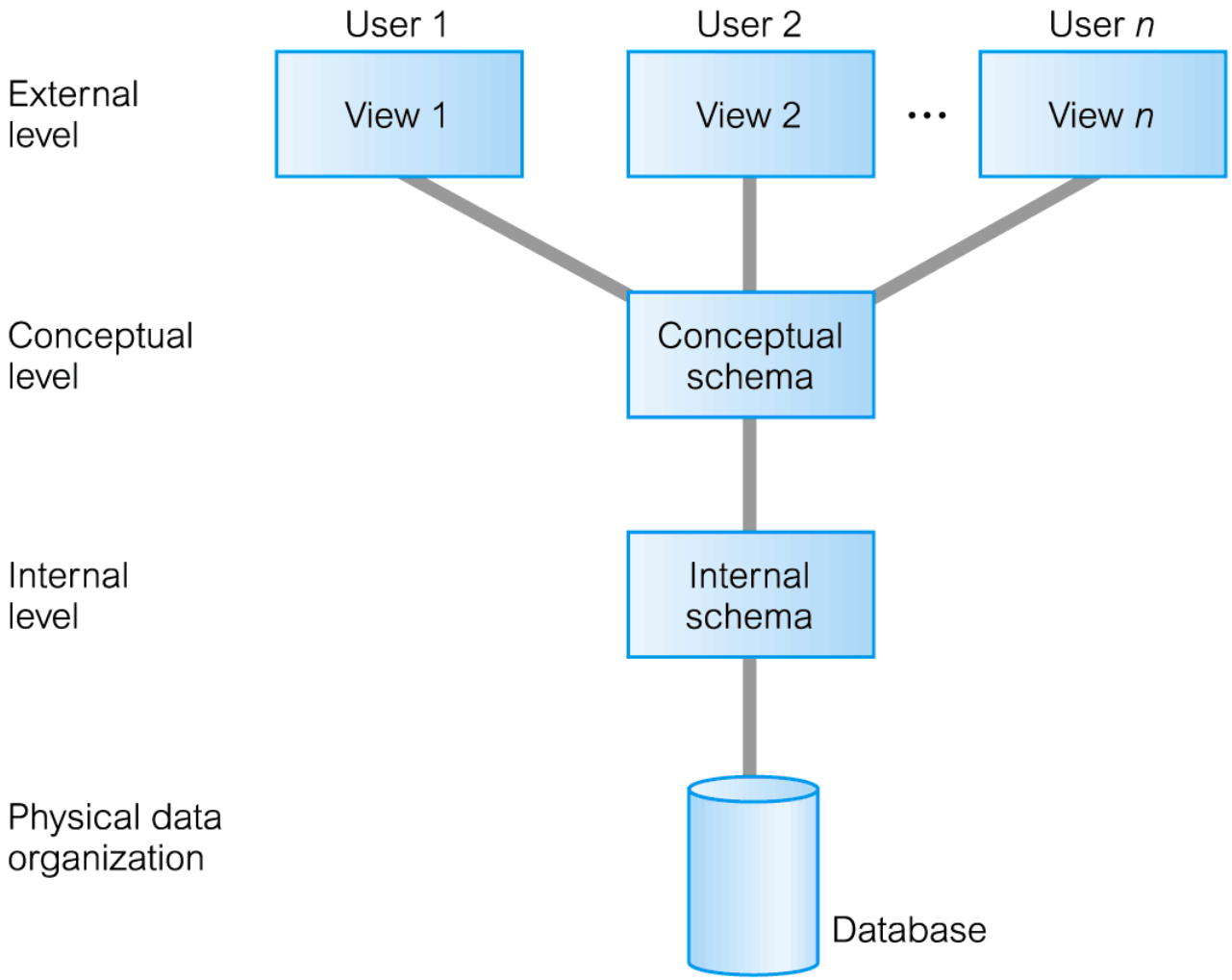
Objectives of Three-Level Architecture

- **All users should be able to access same data.**
- **A user's view is immune to changes made in other views.**
- **Users should not need to know physical database storage details.**

Objectives of Three-Level Architecture

- **DBA should be able to change database storage structures without affecting the users' views**
- **Internal structure of database should be unaffected by changes to physical aspects of storage**
- **DBA should be able to change conceptual structure of database without affecting all users**

ANSI-SPARC Three-Level Architecture



ANSI-SPARC Three-Level Architecture

- **External Level**
  - **Users' view of the database**
  - **Describes that part of database that is relevant to a particular user**
  - **Hides irrelevant data and data the user is not authorized to see**
    - **Adds security – user is not aware of what is missing**

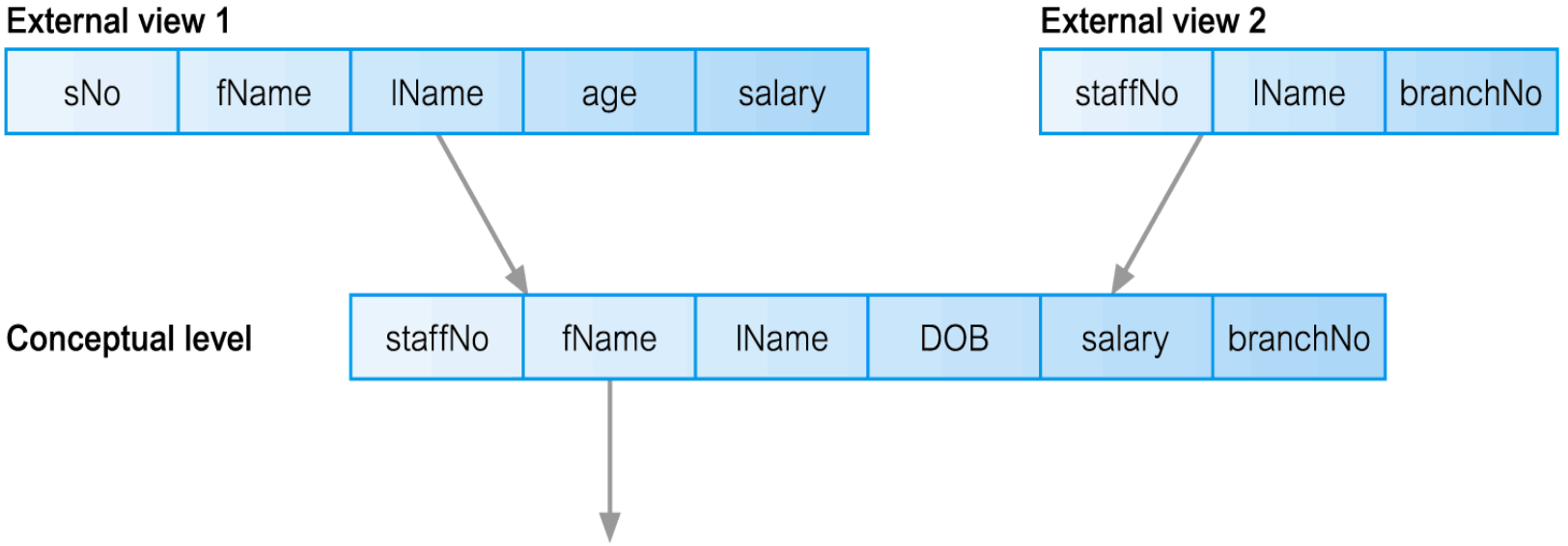
ANSI-SPARC Three-Level Architecture

- **Conceptual Level**
  - An overall view of the database
  - Deals with the logical structure of the entire database as seen by the DBA
    - Also known as the logical level
  - Describes what data is stored in database and relationships among the data
    - Entities, attributes, relationships
    - Constraints on the data
    - Semantic information about the data
    - Security and integrity information
  - Independent of the physical implementation

ANSI-SPARC Three-Level Architecture

- **Internal Level**
  - **Physical representation of the database on the computer**
  - **Describes how the data is stored and organized in the storage medium**
  - **Considerations include**
    - **Storage space allocation for data and indexes**
    - **Record descriptions for storage (with stored sizes for data items)**
    - **Data compression and encryption techniques**
  - **Managed by the operating system under the direction of the DBMS**

Differences between Three Levels of ANSI-SPARC Architecture



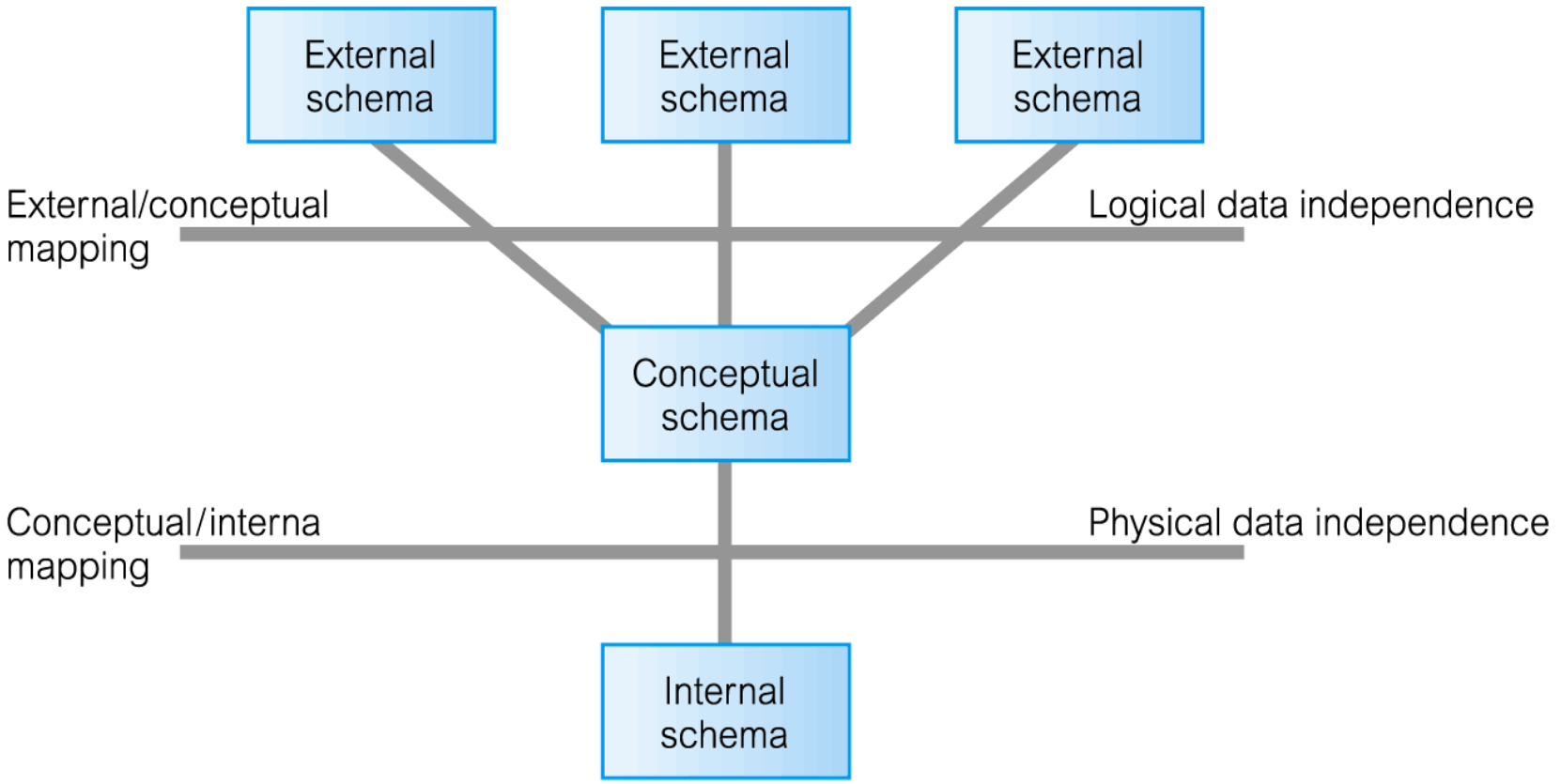
Internal level

```
struct STAFF {
        int staffNo;
        int branchNo;
        char fName [15];
        char lName [15];
        struct date dateOfBirth;
        float salary;
        struct STAFF *next;                 /* pointer to next Staff record */
};
index staffNo; index branchNo;              /* define indexes for staff */
```

Three Database Schemas

- Three-level architecture describes the database schema at different levels
  - External level has external schemas (or subschemas)
    - User Views
    - Highest level
  - Conceptual level has a conceptual schema
    - Describes all of the entities, attributes, and relationships along with the integrity constraints
  - Internal level has an internal schema
    - Complete description of the internal model
    - Data fields, indexes, storage structures used, etc.

Data Independence and the ANSI-SPARC Three-Level Architecture



Data Independence

- **Logical Data Independence**
  - **Refers to immunity of external schemas to changes in conceptual schema.**
  - **Conceptual schema changes (e.g. addition/removal of entities, attributes, or relationships).**
    - **Should not require changes to external schema or rewrites of application programs.**
    - **Users for whom the changes have been made need to be aware of them**
    - **Should not affect other users**

Data Independence

- **Physical Data Independence**
  - **Refers to immunity of conceptual schema to changes in the internal schema.**
  - **Internal schema changes (e.g. using different file organizations, storage structures/devices, indexing scheme, etc.).**
  - **Should not require change to conceptual or external schemas.**
  - **Users may notice change in performance**
    - **Deterioration in performance – most common reason for internal changes**

Data Sublanguages

- **Two DBMS sublanguages: DDL and DML**
- **A data *sublanguage* does not include constructs for all computing needs, such as conditional or iterative statements, which are provided by the high-level programming languages**
  - **Can be embedded in a high-level programming language (host language)**
  - **Many can be used in a non-embedded manner as well**
    - **Interactive commands that can be input directly from a terminal (or a tool)**

Database Languages

- **Data Definition Language (DDL)**
  - **Used to specify or modify the database schema**
  - **Not used to manipulate data**
  - **Allows the DBA or user to**
    - **describe and name entities, attributes, and relationships, required for the application.**
    - **define any associated integrity and security constraints.**
  - **Result of the compilation of DDL statements is a set of tables stored in a *system catalog*.**

System Catalog

- **Repository of information (metadata) describing the data in the database.**
- **One of the fundamental components of DBMS.**
- **Typically stores:**
  - **names, types, and sizes of data items;**
  - **constraints on the data;**
  - **names of authorized users;**
  - **data items accessible by a user and the type of access;**
  - **usage statistics.**

Database Languages

- **Data Manipulation Language (DML)**
  - **Provides basic data manipulation operations on data held in the database.**
    - **Insertion of new data into the database**
    - **Modification of data stored in the database**
    - **Retrieval of data contained in the database**
    - **Deletion of data from the database**
  - **The data retrieval part of DML is called a query language**
    - **A high-level special-purpose language used to satisfy diverse requests for the retrieval of data held in the database**

Database Languages

- **Two types of DMLs**
  - **Procedural DML**
    - **allows user to tell system exactly how to manipulate data**
  - **Non-Procedural DML**
    - **allows user to state what data is needed rather than how it is to be retrieved.**
    - **SQL and QBE**

Data Model

- **DDL is tied to a specific DBMS**
  - **We would like to have a higher-level description that is independent of the DBMS**

- **Data Model**
  - **Integrated collection of concepts for describing data, relationships between data, and constraints on the data in an organization.**
  - **Purpose**
    - **To represent data in an understandable way**
    - **To facilitate discussion between db designers and end-users/stakeholders**

Data Model

- **Data Model comprises:**
  - **a structural part**
    - **Consists of a set of rules according to which databases can be constructed**
  - **a manipulative part**
    - **Defines type of operations that are allowed on the data (including retrieving and updating data and change db structure)**
  - **possibly a set of integrity rules**
    - **Ensures the data is accurate**

Data Model

- **Categories of data models include:**
  - **Object-based**
  - **Record-based**
  - **Physical**

Data Models

- **Object-Based Data Models**
  - **Entity-Relationship**
  - **Semantic**
  - **Functional**
  - **Object-Oriented.**
- **Record-Based Data Models**
  - **Relational Data Model**
  - **Network Data Model**
  - **Hierarchical Data Model.**
- **Physical Data Models**

Data Models

- **Object-Based Data Models**
  - **Entity-Relationship**
  - **Semantic**
  - **Functional**
  - **Object-Oriented.**
- **Record-Based Data Models**
  - **Relational Data Model**
  - **Network Data Model**
  - **Hierarchical Data Model.**
- **Physical Data Models**

Relational Data Model

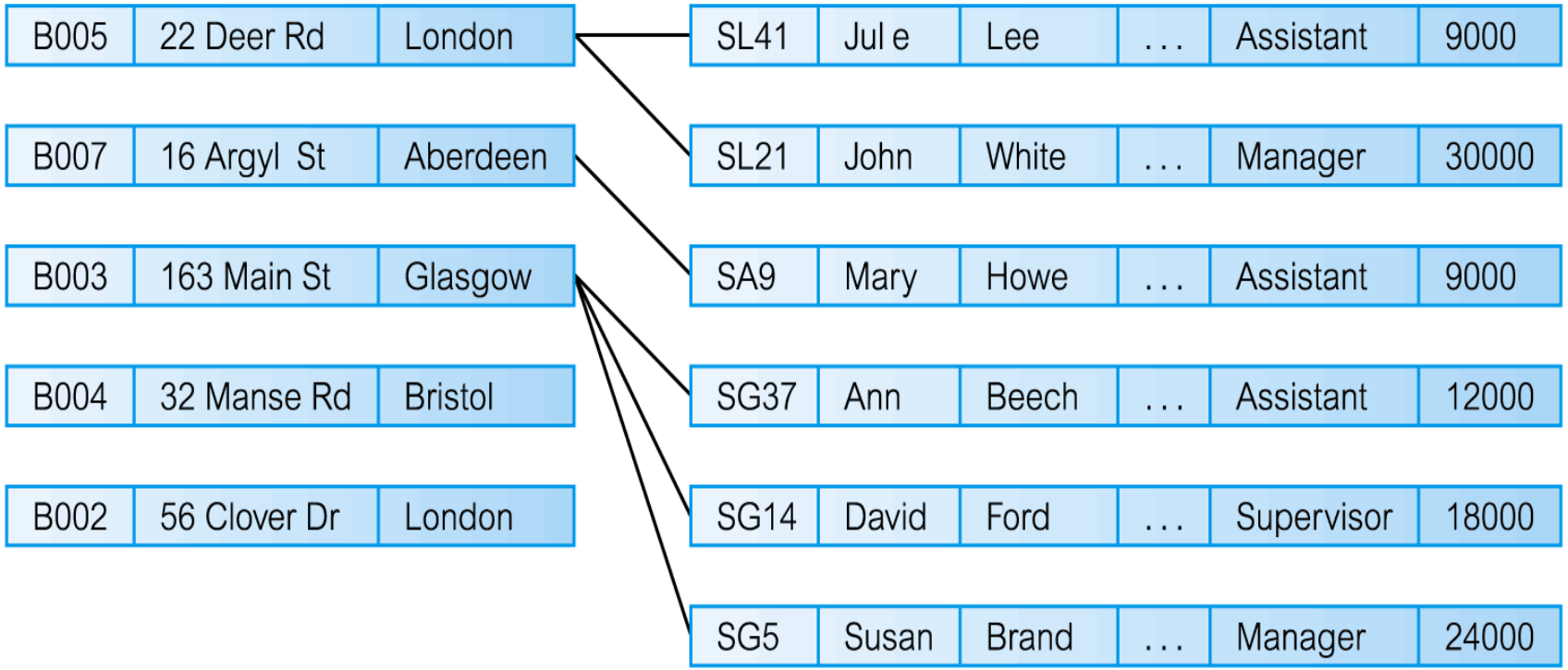
Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

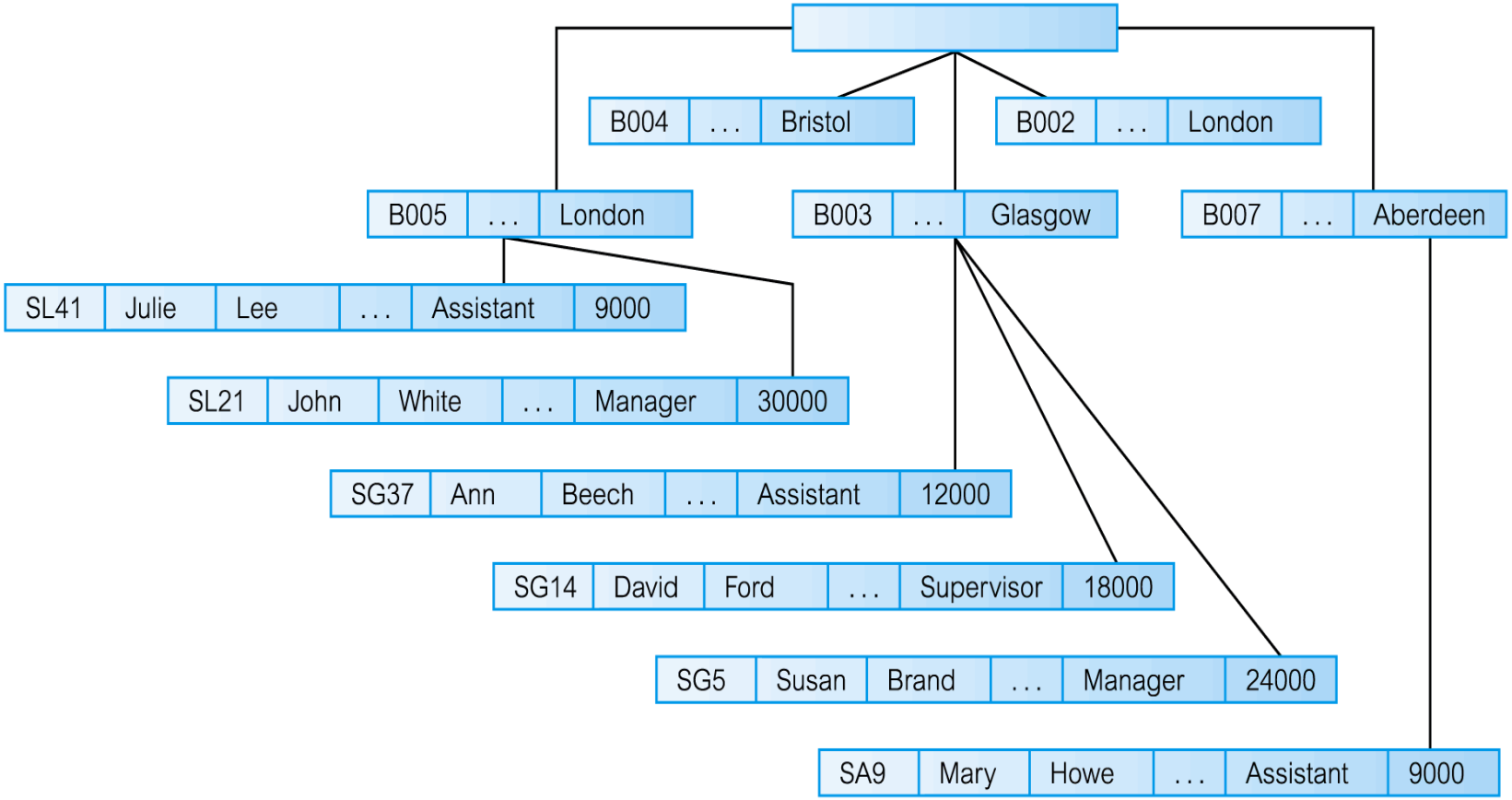
Staff

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

Network Data Model



Hierarchical Data Model



Conceptual Modeling

- **Conceptual schema is the core of a system supporting all user views.**
  - **Should be complete and accurate representation of an organization's data requirements.**
- **Conceptual modeling is process of developing a model of information use that is independent of implementation details.**
  - **Result is a conceptual data model.**

Conceptual Modeling

- **Conceptual model is sometimes called a logical model, but the two differ**
  - **Conceptual model**
    - **Independent of *all* implementation details**
  - **Logical model**
    - **Assumes knowledge of the underlying data model of the target DBMS**

Functions of a DBMS

- **Data Storage, Retrieval, and Update**
  - **Fundamental function**
- **A User-Accessible Catalog**
  - **Users of the DBMS, not end-users**
- **Transaction Support**
  - **All transaction updates are made or none of them are**
- **Concurrency Control Services**
  - **No interference with multiple users accessing**
- **Recovery Service**
  - **Must be able to recover if database is damaged in any way**

Functions of a DBMS

- **Authorization Services**
  - **Only authorized users have access (db and data w/in db)**
- **Support for Data Communication**
  - **Access over a network**
- **Integrity Services**
  - **Data is correct and consistent**
- **Services to Promote Data Independence**
  - **Physical and logical**
- **Utility Services**
  - **For administering the database**