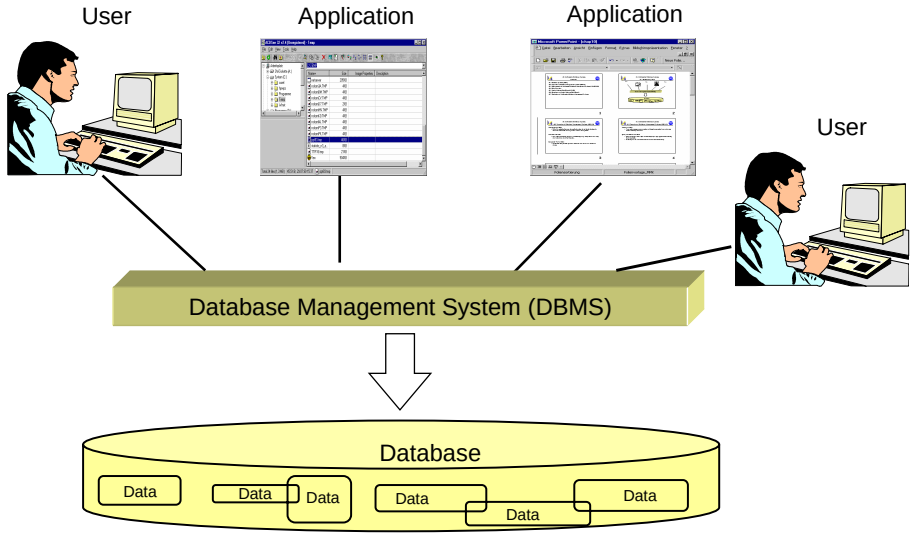
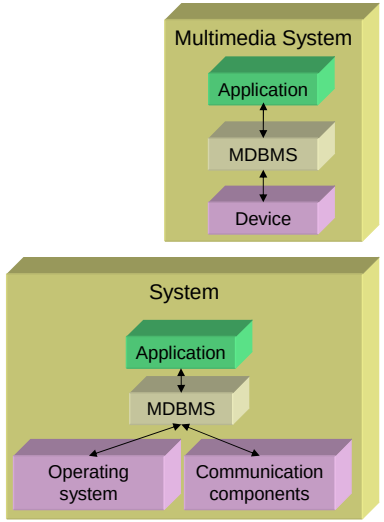
Database Systems



Multimedia Database Management System

- Main task of a Database Management System (DBMS) is to abstract from the details of:
  - Storage access
  - Storage management
- Location of the MDBMS:
  - Embedded between the application domain and the device domain
- *Integration into the system*:
  - Through operating system
  - Communication components



Multimedia Database Management System

- *Persistence of data*:
  - Data outlive processing programs and technologies, e.g. companies have to keep data in databases for several decades
- *Consistent view of data*:
  - Synchronization protocols provide a consistent view of data in a multi-user system
- *Security of data*:
  - Transaction concepts ensure security and integrity protection in case of system failure. Recovery of lost data.
- *Query and retrieval of data*:
  - Query languages such as SQL (Structured Query Language) enable formulating database queries
  - Each entry has its state information that can be retrieved correctly

Characteristics of MDBMS

- *Corresponding storage media*
  - Multimedia data must be stored and managed according to the specific characteristics of the available storage media
- *Descriptive search methods*
  - Query of multimedia data should base on a descriptive and content-oriented search, e.g. “Picture of a woman with a red scarf”
- *Device-independent interface*
  - Hide details of device control, but offer information on specific characteristics of available storage media (read-only, write-once, write-many)
- *Format-independent interface*
  - DBMS must hide internal storage format and offer conversions to formats requested by the applications (GIF, TIFF, JPEG, ...)
  - This allows changing to new storage technologies without any impact on multimedia applications

Characteristics of MDBMS

- *View-specific and simultaneous data access*
  - Allows consistent, multiple and simultaneous data access through different queries of several applications (e.g. shared editing)
- *Management of large amounts of data*
  - DBMS must be capable of handling and managing large amounts of data. Need of appropriate referencing mechanisms.
- *Relational consistency of data management*
  - Relations among data of one or different media must stay consistent corresponding to their specification. MMDBMS manages the following relations:
    - *Attribute Relation*: supports different presentation (audio, video, image) of one object
    - *Component Relation*: includes all parts belonging to one data object
    - *Substitution Relation*: defines different kinds of presentation of the same information, e.g. equation as tables, graphs, animation
    - *Synchronization Relation*: describes temporal relations between data units, e.g. lip synchronization of audio and video

Characteristics of MDBMS

- *Real-time data transfer*
  - DBMS must perform read and write operations of continuous data in real-time
  - The data transfer of continuous data has a higher priority than other database management actions
  - Primitives of multimedia operating systems should be used to support the real-time transfer of continuous data
- *Long transactions*
  - The transfer of large amounts of data will take a long time and must be done in a reliable fashion

Characteristics of MDBMS

- *Relation between the operating system and MDBMS:*
  - The operating system provides the management interface for MDBMS to all local devices
  - The MDBMS provides an abstraction of the stored data and their equivalent devices, as is the case in DBMS without multimedia
  - The communication system provides for MDBMS abstractions for communication with entities at remote computers
  - Operating system and communication system can unify all the different abstractions and offer them

Data Structure

Data can be stored in databases as

- Unstructured (unformatted): data are presented in a unit where content cannot be retrieved by accessing any structural detail

  Example: “Mr. Penguin is a student in the seventh term.”
- Structured form (formatted): data are stored in variables, field or attributes with corresponding values

  Example: o.student.surname = “mustername”
  o.student.name = “hermann”
  o.student.age = 41

Data Types

Multimedia data can be stored in databases as raw, registering and descriptive data types:

- *Raw Data*: represent the unformatted information content, e.g. letters, pixel, values
- *Registering data*: necessary for correct interpretation and identification of the data; usually concealed in the header. For example: format-description (GIF, TIFF, JPEG, ASCII, EBCDIC, ...), compressed/uncompressed data, etc.
- *Descriptive data*: information about content and structure of the multimedia data to make use easier and faster, e.g. semantic search

Data Types - Examples

Text

- Characters represent raw data
- Registering data describe the coding (e.g., ASCII)
- Descriptive data may include information for layout and logical structuring of the text, or keywords

Image

- Pixels represent raw data
- Registering data include the height and width of the picture
- Descriptive data are individual lines, surfaces and subjects

Data Types - Examples

Video sequence

- Pixel matrices represent the raw data
- Registering data provides, in addition to other information, the number of images per second
- Descriptive data provide a scene description, e.g. ”Jan‘s birthday party”

Audio sequence

- The digital sample values created by a simple PCM coding represent the raw data
- Registering data represent the properties of the audio coding
- Descriptive data represent the content of the audio

Operations on Data

An MDBMS must offer, for all data types, corresponding operations for:

- archival and
- retrieval

The media related operations will be handled as part or extension of query languages, e.g. SQL

Different classes of operations are needed:

- input
- output
- modification
- deletion
- comparison
- evaluation

Operations on Data

- *Input (insert / record) operation*:
  - Data will be written to the database
  - The raw and registering data are always needed, descriptive data can be attached later
- *Output (play) operation*:
  - Read the raw data from the database according to the registering data
- *Modification*:
  - Changing of raw, registering and descriptive data
  - Modification can also be understood as a data conversion from one format to another
- *Deletion operation*:
  - Remove an entry from the database
  - The consistency of the data must be preserved

Operations on Data

- *Comparison*:
  - Many queries to the MDBMS consist of a search and retrieval of the stored data
  - Queries are based on comparison information
  - Individual patterns in the particular medium are compared with the stored raw data → not successful enough
    → Pattern matching, search in descriptive data, etc.
- *Evaluation*:
  - Generation of the corresponding descriptive data from the raw and registering data

Integration in a Database Model

Design of multimedia database systems base on two different kinds of DBMS:

- *ERDBMS (Extensible Relational Database Management System):*
  - Definition of additional, application-dependent data types as domains for attributes
  - Definition of new functions to control behavior of and access to the data
  - Embedding new types and functions into existing RDMBS
- *OODBMS (Object-Oriented Database Management System):*
  - Different media are represented by classes, whose instance variables include the data as internal state
  - Class hierarchy allows object relations, offers well information navigation and flexible presentation possibilities

Relational Database Model

Simplest possibility to implement a multimedia database is to use the *relational database model*

- The attributes of different media in relational databases are defined
- Attributes can specify
  - text
  - audio
  - video
- Advantage
  - compatibility with existent database applications

Relational Database Model - Example

A relation “student” is given:

```
Student (
  Admission_Number      Integer,
  Name                  String,
  Picture               Image,
  Exercise_Device_1     Video,
  Exercise_Device_2     Video
)
```

- A relation’s attributes can be specified through different media types: image, exercise, video
- Other entries are ”athletics”, ”swimming” and ”analysis”

```
Athletics (
  Admission_Number      Integer,
  Qualification         Integer,
  The_High_Jump         Video,
  The_Mile_Run          Video
)

Swimming(
  Admission_Number      Integer,
  Crawl                 Video
)

Analysis (
  Qualification         Integer,
  Error_Pattern         String,
  Comment               Audio
)
```


Relational Database Model

- *Type 1 Relational Model*
  - Value of a certain attribute can be fixed over the particular set of the corresponding attribute types, e.g. the frame rate of the video can be fixed
  - In the example, the videos from the exercise devices 1 and 2 will play at the fixed rate defined by the type 1 specification
- *Type 2 Relational Model*
  - A variable number of entries can be defined through the type 2 relational model
  - In the example, the individual disciplines of each admitted student are identified through their admission numbers
- *Type 3 Relational Model*
  - Additionally, an entry can simultaneously belong to several relations
  - In the example, a video entry of a student can be assigned to the relation ”athletics” as well as to the relation ”analysis”

Object-oriented Database Model

In *object-oriented databases*...

- classes with objects are defined
- objects can be put in relations via a class hierarchy
- a semantic specialization of classes and objects can follow

Example

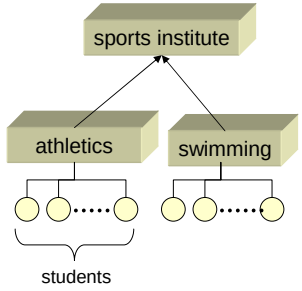
- Main class: sports institute
- Subclass: athletics, swimming
- Objects: students

Advantage:

- These system offer good information navigation and flexible presentation possibility

Disadvantage

- Query operations are incompletely supported



Conclusions

Media can be stored in many ways, only a few were talked about::

Optical Storage

- Compact Disc (CD) for multimedia data like images
- Digital Versatile Disc (DVD) for videos as main medium
- Important: new error protection mechanisms like Reed-Solomon code

Hard disk / file systems

- Different requirements to between “normal” data and multimedia data
- Suitable data structuring needed
- New strategies for disk scheduling needed

Databases

- Can base on relational or object-oriented models
- Structure and behavior of data has to be stored
- New language primitives are needed
- Needed here: more work on transaction management and content-based retrieval of data