

File Management and Data Processing

Sisira Kumar Kapat

Department of Computer Science & Engineering,
UCP Engineering School,
Berhampur, Orissa

Contents

- Concept of File and Folder
- File Access and Storage methods.
- Sequential, Direct, ISAM
- Data Capture, Data storage
- Data Processing and Retrieval

Concept of File and Folder

- File: File is a set of related data or information that is being stored in memory.
 - File is the common storage unit in a computer.
 - All programs and data are contained.
 - Computer reads and writes files.
 - Files that contain text is often referred as documents.
 - Files can have extensions like .txt, .jpg, .exe etc.
 - Data file contains data and information.
 - Program file contains program code and can be executed.
- Folder: A folder is used to contain many other files and folders.
 - It holds one or more files.
 - It can be empty with just a name.
 - A folder can also contain other folder and subfolders.
 - It is a method for organizing files.
 - Folder are also called as “directories”.
 - Any number of files and folders are possible.
 - Each folder can have different/numerous entries depending on the files created where each file has a position in a parent folder.

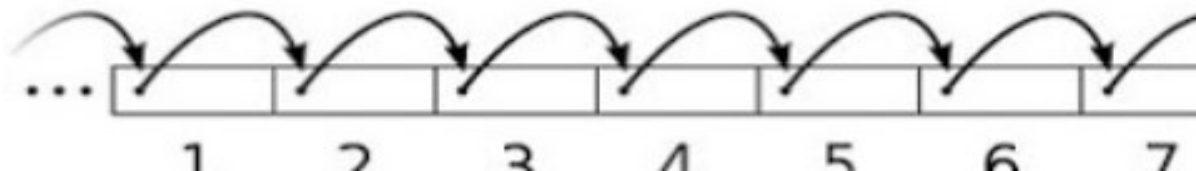
File Access and Storage Methods

- File access and storage methods.
 - Sequential
 - Direct
 - Indexed Sequential Access Method
- File Storage methods

Sequential

- Simplest type of access method.
- Information in the file is processed in order, one record after the other.
- *read next*- read the next position of the file and automatically advance a file pointer.
- *write next* append to the end of the file and advance to the newly written material.
- Such a method is reasonable for tape.

Sequential access



Direct

- *relative access method*
- A file-length logical record that allows the program to read and write record rapidly in no particular order.
- The direct access is based on the disk model of a file since disk allows random access to any file block.
- For direct access, the file is viewed as a numbered sequence of block or record.
- There is no restriction on the order of reading and writing for a direct access file.
- A block number provided by the user to the operating system is normally a *relative block number*, the first relative block of the file is 0 and then 1 and so on.

Indexed Sequential Access Method

- It control the pointer by using index.
- The index, like an index in the back of a book, contains the pointer to the various blocks.
- To find a record in the file, we first search the index and then by the help of pointer we access the file directly.

Data Capture

- Audio
 - convert speech to text such as google assistant, amazon alexa.
- Image
 - OCR (Optical Character Recognition)
 - Bar code recognition
- Text
 - identify post codes, logos, key words

Data storage

- Magnetic Disk
- Optical Storage
- Hard Disk Drive
- Solid State Drive

Data Processing and Retrieval

- Hard Disk
 - Platter
 - Track
 - Sector
 - Cluster
 - R/W Head
 - Spindle motor
 - Actuator

