

CONCEPT OF COMPUTER FILE

Definition of terms

- 1. **Data:** a raw fact that has not been processed. It is the smallest unit (an item) of information e.g. James, 45, Male etc.
- 2. **Field:** it is a space/named area of a record allocated to store an item of information, e.g. Name, Age, Gender etc.
- 3. **Record:** It is a collection of related data items or fields about an entity i.e. a person, thing or place. For example, a student record may consist of his/her gender, name, age etc.
- 4. (Computer) **file:** A file can be defined as a collection of related records that give a complete set of information about a certain item or entity.

A file may also refer to an object on a computer that stores data, information, settings, or commands used with a computer program. In a graphical user interface (GUI) such as Microsoft Windows, files display as icons that relate to the program that opens the file.

A file is created using a software program on the computer. For example, to create a text file you would use a text editor (notepad), to create an image file you would use an image editor (paint, corel draw etc.), and to create a document you would use a word processor (Microsoft Word, Word Perfect etc).

Computer files are stored on a drive (e.g., the hard drive), disc (e.g., DVD), and a diskette (e.g., floppy disk) and may also be contained in a folder (directory) on that medium. Every created file has an extension. A **file extension** or file name extension is the ending of a file that helps identify the type of file in operating systems, such as Microsoft Windows. In Microsoft Windows, the file name extension is a period that is often followed by three characters, but may also be two or four characters long. As an example the file name "myfile.txt" has a file extension of ".txt", which is a file name extension associated with text files.

Assignment:

List the type of file associated with each of the following file extensions:

1. ".3gp"
2. ".exe"
3. ".dll"
4. ".bat"
5. ".html"
6. ".cmd"
7. ".rtf"
8. ".accdb"
9. ".pptx"
10. ".cdr"

Types of data items

The following are **types of data** (item):

- **Numeric data:** these consists of digits (0-9) e.g. 1, 43, 0.56 etc.
- **Alphabetic data:** these consists of alphabetic characters (A-Z or a-z) only e.g. School, Bond, Name etc.
- **Alphanumeric data:** these are data made up of combination or alphabets and numbers.

Types of File Organisation Method

File organization is a way of organizing the data or records in a file. It refers to how the contents of a file are added and accessed, but not how files are organized in folders. The four file organisation methods are:

- 1. **Serial file organisation:** records are stored in the order they occur. They have not been sorted in any particular order.
- 2. **Sequential file organisation:** records are stored in a sorted order of a particular field(s), usually the key field(s).
- 3. **Indexed file organisation:** An indexed file contains records ordered by a *record key*. A record key uniquely identifies a record and determines the sequence in which it is accessed with respect to other records.
- 4. **Random or direct file organisation:** records are stored randomly in no particular order i.e. in any sequence

Methods of Accessing Files

Access method is a mechanism or manner in which the records in a file may be accessed. It defines the way the *read* and *write* operations are done. The methods of accessing files include:

- 1. **Sequential access:** a sequential file access is that in which the records are accessed in some sequence i.e., the information in the file is processed in order, one record after the other. It requires the program to start writing or reading at the beginning and continues until it finds the desired data. Device like magnetic tape enforces sequential access method.
- 2. **Direct/random access:** the records on the storage location can be accessed (read or written to) in any order i.e. randomly/directly. Devices such as magnetic disk storage and the main storage i.e. RAM and ROM are based on this method.
- 3. **Indexed-sequential access:** this mechanism is built on the basis of sequential access. An index is created for each file which contains pointer to various records (blocks). Index is searched *sequentially* and its pointer is used to access the file *directly*.

Computer File Classifications

Computer files can be classified as follows:

- 1. **Master file:** This is a computer file that is used as the authority in a given job that is relatively permanent. It is a permanent file, periodically updated, that serves as an authoritative source of data. It is an original file from which duplicates are made.
- 2. **Transaction file:** It is a computer file containing relatively transient data about a particular data processing task. It is a file, especially a data file, containing transaction records, used to update the master file.
- 3. **Reference file:** This is a computer file containing data, which is kept so that it can be referenced for future use. It is stable and permanent in nature.

Criteria for classifying computer file

Computer files can be classified according to the following criteria:

1. **Nature of content:** files of similar contents are classified together. Examples are database file, word processed file etc.
2. **Organisation method** i.e. whether sequential, direct etc.
3. **Storage medium:** whether they are stored in tapes, disks or any other storage devices.