FILE TYPES

Types of files - Confusion: Text files vs. Binary Files

■ Text files (ASCII / UNICODE)

– Bytes of data are organised as characters from respective character sets

■ Binary files

– Data in a specific format that requires interpretation.

■ Text files vs. Binary Files

– All files are in Binary

– Text Files are formatted in chunks of 8 bits or 16 bits

– Files in any other format are Binary Files

File Types

■ Most files contain a specific types of information

– A Java program

– A JPEG image

– A BITMAP image

– An MP3 clips

■ The kind of information is the file type

– So, the File System knows which operations it can do

– Most OS have associations between file types and applications

File Types Extensions

■ File names are often separated by a full-stop into 2 parts

– Main name

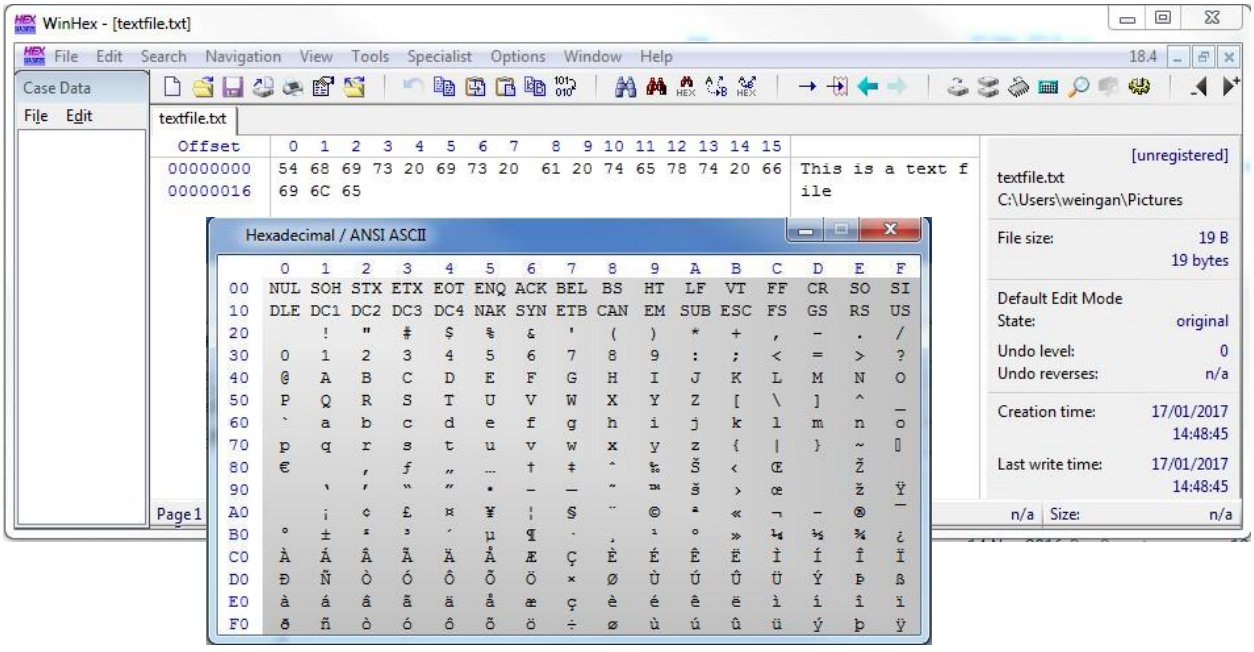
– File extension

■ The file extension was used by the OS to identify the type of file – But is not necessarily the actual file type

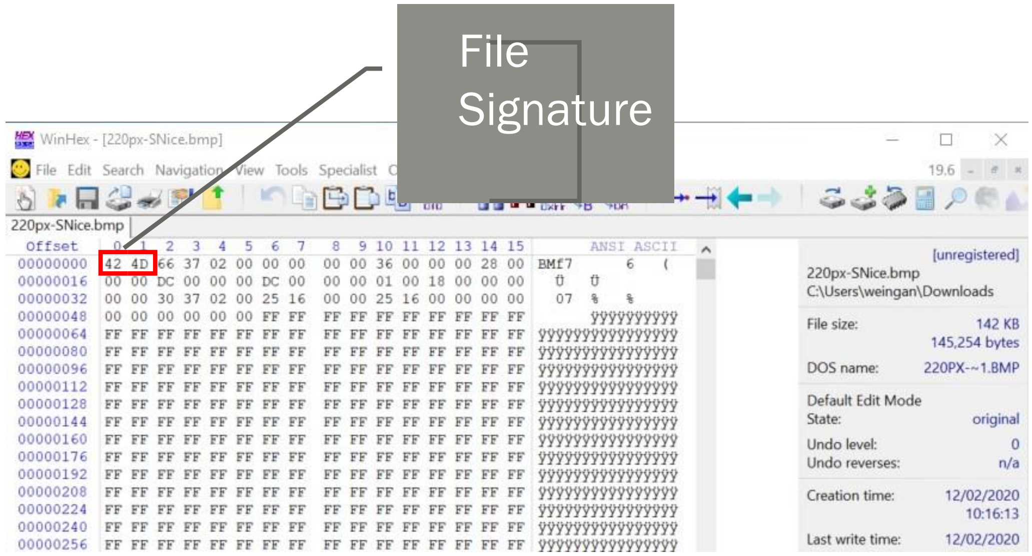
■ Windows 10 will inspect the file to ascertain the actual file type – Looking at the file header

Extension	File Type
.txt	Text data file
.mp3, .au, .wav	Audio file
.gif , .tiff , .jpg	Image file
.doc , .odt	Word processing files
.java , .sql	Programming source file

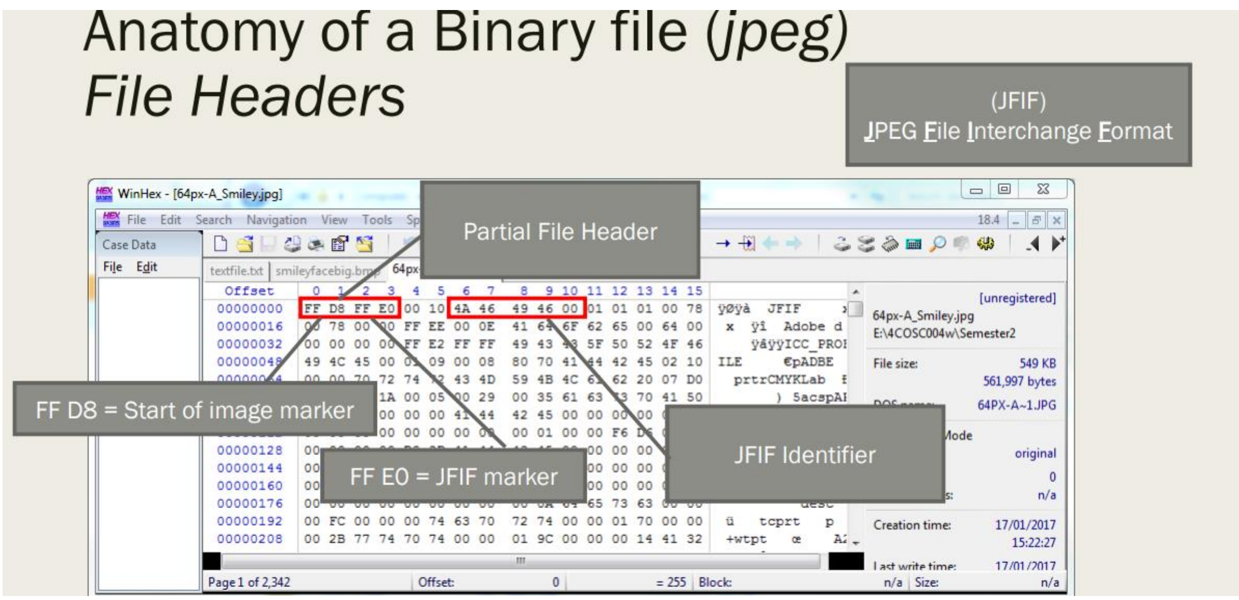
Anatomy of an ASCII File



Revisiting BMP files



Anatomy of a Binary file (jpeg) File Headers



File Signatures

■ There file signature databases – Filesignatures.net

■ Wikipedia often has high quality listings of the entire file header

File Operations

■ Create a file ■ Delete a file ■ Open a file ■ Close a file ■ Read data from a file

■ Write data to a file

■ Reposition the current file pointer in a file

■ Append data to the end of a file

■ Truncate a file – ie. delete all or part of it

■ Rename a file

■ Copy a file

File Protection

■ Multi-user Systems

■ Access control – Controls who can access files

■ Who can read

■ Who can write

■ Who can execute

Cyber-Security Triad

■ Three dimensions of cyber-security:

1. Confidentiality:

• Preventing access

• Keep the bad-guy out

2. Accessibility:

• Ensure access

• Make sure the good-guy can access the data

3. Integrity:

• Keep control of any changes made to the data

• Who can change it

• Keep track of any changes

File Permissions (Windows)

■ NTFS:

– Access Control Lists (ACL's)

■ Each file has list of user identities with permissions

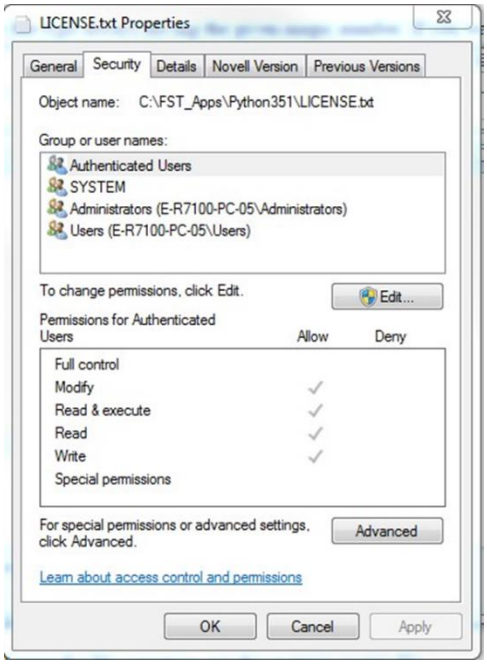
– Explorer

■ File Permissions

■ Security

– Different user, different permissions

■ No multi-user security for FAT32



Windows Permissions Classifications

■ Full control

– File can be written to/read from

– Permissions can be read and modified

– Ownership can be changed

– Folder can be listed and entries deleted

■ Modify

– Same as Full control

– But cannot change permissions or ownership

Windows Permissions Classifications

■ Read/Execute

– File can be read or executed as a program

– Folder can be listed and traversed

■ Read

– File can be read

■ But not executed

– Folder can be listed

■ But not traversed

■ Write

– File can be modified

– Files/subfolders can be created in a folder

■ But NOT deleted

■ List folder contents (for folders only)

– Same as Read/Execute, but not available for files, and only inherited by folders

Security Inheritance

■ Windows:

– New file or subfolder created, will inherit it's parent's permissions by default

– You can override

■ Unix:

– Permissions are not inherited for newly created files

– Based on user's umask

– Mask of permissions specific to that user – octal absolute format

DISK SCHEDULING

■ Must be efficient

■ Multiple processes, multiple requests to access disk

■ Disk scheduling techniques to manage request:

– First-come, first-served (FCFS): Requests are serviced in the order they arrive, irrespective of positions of heads

– Shortest-Seek-Time-first (SSTF): Minimise movement of disk heads

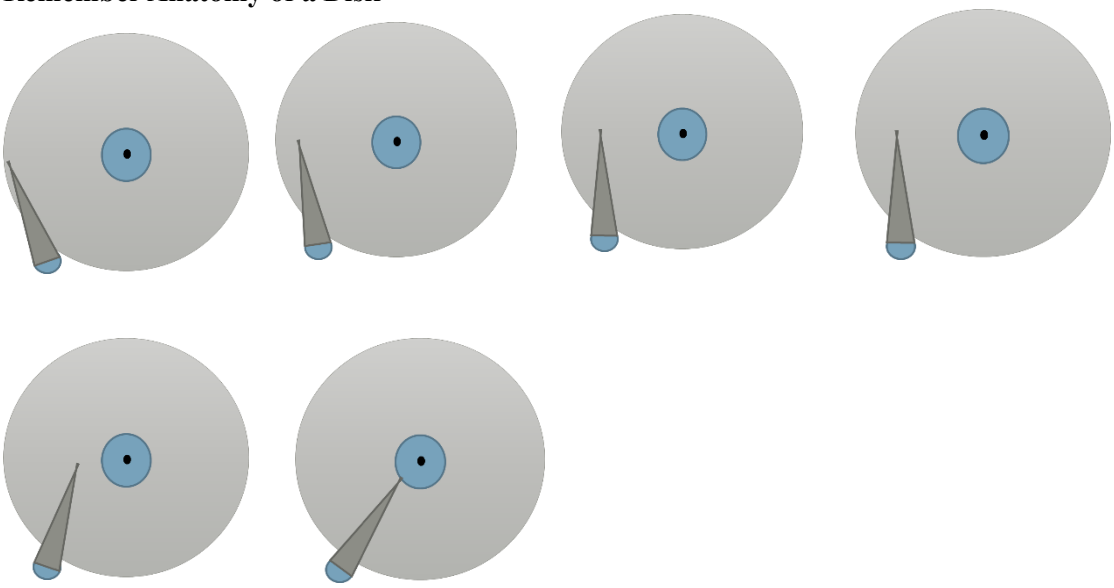
– SCAN: Disk heads continuously move in and out, servicing requests as the locations are found.

– C-SCAN: Circular scan

– Look: Like SCAN, but does not scan all the way to edge

– C-Look: Like C-SCAN

Remember Anatomy of a Disk



Example:

■ Work queue: 23, 89, 132, 42, 187

■ There are 200 cylinders 0-199

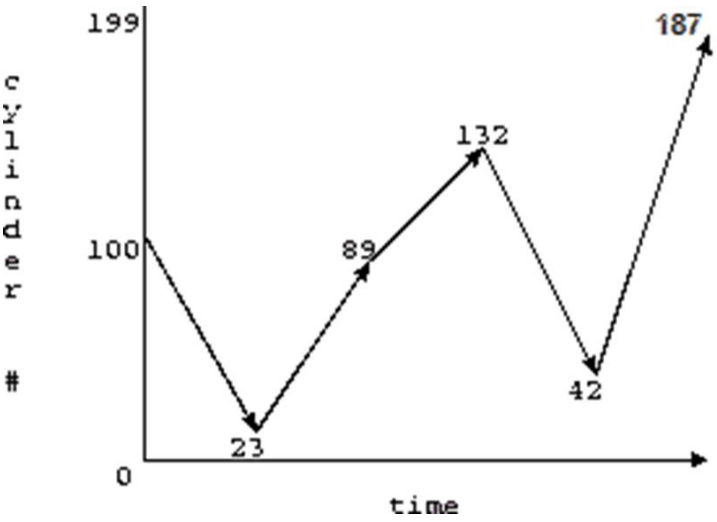
■ The disk head starts at number 100

First-Come-First-Served:

■ Work queue: 23, 89, 132, 42, 187

■ Total seek length:

23 − 100 + 89 − 23 + 132 − 89 + 132 − 89 + 42 − 132 + 187 − 42 = 421



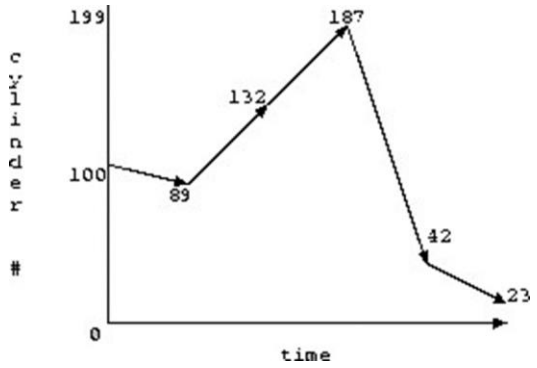
Shortest-Seek-Time-First:

■ Work queue: 23, 89, 132, 42, 187

■ Can be inefficient

– Multiple changing directions

– Starvation



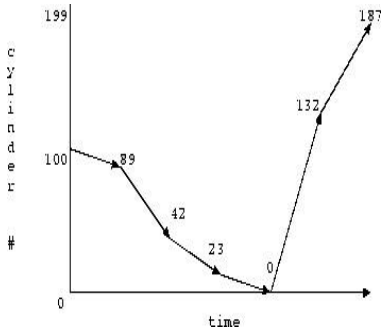
SCAN:

■ Work queue: 23, 89, 132, 42, 187

■ Elevator

■ Sweeps the disk, to-and-fro

■ LOOK is similar

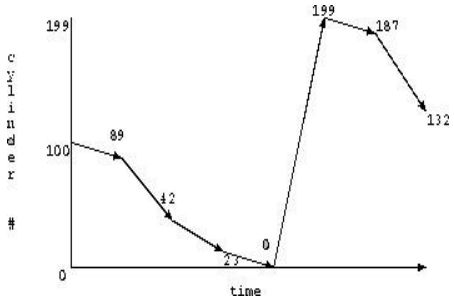


C-SCAN:

■ Work queue: 23, 89, 132, 42, 187

■ Elevator

■ Sweeps the disk, but one-direction



Performance:

■ Depends on number of requests

■ SCAN & C-SCAN are good for systems that place a heavy load on the disk, less likely to cause starvation

■ Default: SSTF or LOOK

– PRIORITY