

# File Systems and Directories

Instructor: Dmitri A. Gusev

Spring 2007

CSC 120.02: Introduction to Computer Science

Lecture 14, March 29, 2007

# File Systems

Main memory is *volatile*. Secondary memory is *nonvolatile*.

Secondary storage devices: magnetic disk drives, tape drives.

A *file* is a named collection of related data. A *file system* is the operating system's logical view of the files it manages.

A *directory*, or *folder*, is a named group of files.

A file may contain a program or data. A *file format* describes how the data is represented and structured inside the file.

A *text file* is a file that contains characters in some format, such as ASCII or Unicode. A *binary file* contains data in a specific format, requiring a special interpretation of its bits.

A readable text file storing a program is called a *source file*.

# File Types

A *file type* is the specific kind of information contained in a file, such as a Java program or a Microsoft Word document. A file type may be associated with one or more application programs capable of handling the corresponding files.

A *file extension* is the part of the file name that indicates which file type the file belongs to. More than one extension may be suitable for files of a given file type (.htm and .html; .icc and .icm).

The main file names and extensions can be changed, so the files can be *renamed*.

# File Operations

- create a file
- delete a file
- copy a file
- move a file to a different location
- rename 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 the file
- truncate a file (delete its contents, but not the file itself)
- create a shortcut to a file
- delete a shortcut
- set access permissions

The operating system maintains a table indicating which blocks of memory are free. For each directory, it maintains a table of its files. It also maintains a table of open files.

# File Access

For an open file, a *current file pointer* is indicating the place where the next read or write operation should occur.

*Sequential access*: The data in a file is accessed in a linear fashion

*Direct file access*: The data in a file is accessed directly, by specifying *logical record numbers*

*Buffering*: The data is read/written in chunks of a certain minimum size

Different *permissions (privileges)* can be set to control file access by different users

# Directories, a.k.a. Folders

A *directory tree* is a structure showing the nested directory organization of the file system

The topmost directory, in which all others are contained, is called the *root directory*. It is said to be located at the *highest level*.

A directory containing another is called its *parent directory*, and the one inside it is its *subdirectory*.

The operating system remembers each user's *current working directory*.

An *icon* is a small image representing a file or folder.

A *path* is a text address of a file or subdirectory in a file system. An *absolute path* name begins at the root. A *relative path* begins from the current working directory.

# Interfaces

The operating system may allow users to access the file system via

- a Graphical User Interface (GUI) and/or
- a non-graphical (text-based) interface, such as a *command-line interface*

# Disk Scheduling

- First-Come, First Served (FCFS)
- Shortest-Seek-Time-First (SSTF)
- SCAN (toward the spindle, out toward the platter edge, back toward the spindle, and so on); Circular SCAN is a modification of SCAN that treats the disk as if it were a ring (returns from one extreme to the other extreme without processing requests)
- LOOK: Only go as far as the innermost (or outermost) request