Navigating Linux System Commands

A guide for beginners to the Linux Shell and GNU coreutils

Sayan Ghosh

May 22, 2024

IIT Madras BS Data Science and Applications

Disclaimer

This document is a companion activity book for the System Commands (BSSE2001) course taught by Prof. Gandham Phanikumar at IIT Madras BS Program. This book contains resources, references, questions and solutions to some common questions on Linux commands, shell scripting, grep, sed, awk, and other system commands.

This was prepared with the help and guidance of the course instructors:

Santhana Krishnan and Sushil Pachpinde

Copyright

© This book is released under the public domain, meaning it is freely available for use and distribution without restriction. However, while the content itself is not subject to copyright, it is requested that proper attribution be given if any part of this book is quoted or referenced. This ensures recognition of the original authorship and helps maintain transparency in the dissemination of information.

Colophon

This document was typeset with the help of KOMA-Script and LATEX using the kaobook class.

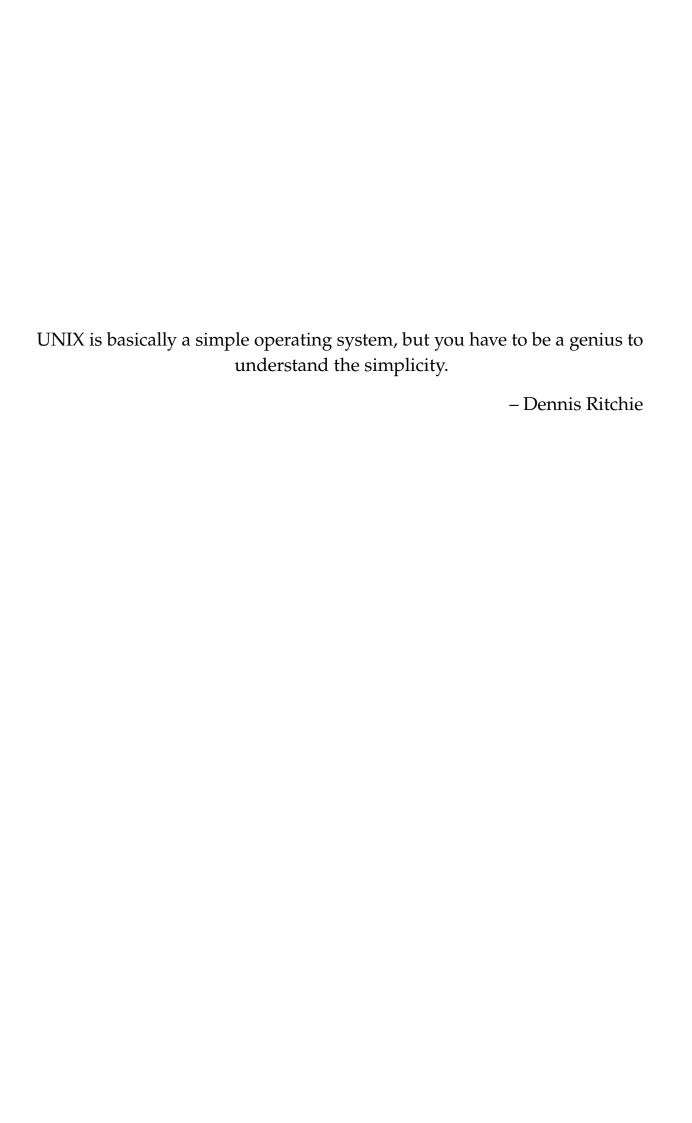
The source code of this book is available at:

https://github.com/sayan01/se2001-book

(You are welcome to contribute!)

Edition

Compiled on May 22, 2024



Preface

Through this work I have tried to make learning and understanding the basics of Linux fun and easy. I have tried to make the book as practical as possible, with many examples and exercises. The structure of the book follows the structure of the course *BSSE2001 - System Commands*, taught by **Prof. Gandham Phanikumar** at **IIT Madras BS Program.** .

The book takes inspiration from the previous works done for the course,

- ► Sanjay Kumar's Github Repository
- ► Cherian George's Github Repository
- ► Prabuddh Mathur's TA Sessions

as well as external resources like:

- ► Robert Elder's Blogs and Videos
- ► Aalto University, Finland's Scientific Computing Linux Shell Crash Course

The book covers basic commands, their motivation, use cases, and examples. The book also covers some advanced topics like shell scripting, regular expressions, and text processing using sed and awk.

This is not a substitute for the course, but a companion to it. The book is a work in progress and any contribution is welcome at https://github.com/sayan01/se2001-book

Sayan Ghosh

Contents

Pr	etace	!		V
Co	nten	its		vii
1	Esse	entials	of Linux	1
	1.1	Introd	duction	1
		1.1.1	What is Linux?	1
		1.1.2	Desktop Environments	2
		1.1.3	Window Managers	3
		1.1.4	Why Linux?	4
		1.1.5	What is Shell?	4
		1.1.6	Shell vs Terminal	4
		1.1.7	Why the Command Line?	5
		1.1.8	Command Prompt	6
	1.2	Simpl	le Commands in GNU Core Utils	7
		1.2.1	File System Navigation	8
		1.2.2	Manuals	9
		1.2.3	System Information	11
		1.2.4	File Management	13
		1.2.5	Text Processing and Pagers	16
		1.2.6	Aliases and Types of Commands	20
		1.2.7	User Management	21
		1.2.8	Date and Time	22
	1.3		gating the File System	25
	1.0	1.3.1	What is a File System?	25
		1.3.2	Why do we have proc and sys directories?	26
		1.3.3	Paths	27
		1.3.4	Basic Commands for Navigation	27
	1.4		ermissions	29
	1.1	1.4.1	Read	30
		1.4.2	Write	30
		1.4.3	Execute	30
		1.4.4	Interesting Caveats	30
		1.4.5	Changing Permissions	32
		1.4.6	Special Permissions	32
		1.4.7	Octal Representation of Permissions	34
	1.5		s of Files	37
	1.5	1.5.1	Regular Files	37
		1.5.2	Directories	37
		1.5.3	Symbolic Links	37
		1.5.4	Character Devices	37
		1.5.4	Block Devices	38
		1.5.6		39
		1.5.7	Named Pipes	39 39
		1.5.7	Sockets	39 40
	1 4		Types of Regular Files	40

	1.7	Types of Commands	42
2	She	ll Variables	45
	2.1	Creating Variables	45
	2.2	Printing Variables to the Terminal	46
	2.3	Bash Flags	47
	2.4	Signals	47
	2.5	Variable Manipulation	48
	2.6	Arrays	50
3	Proc	cess Management	53
		What are processes?	53
A	PPEN	IDIX	55
No	otatio	on	57
Αl	phab	petical Index	59

List of Figures

1.1	Linux Distributions Usage	2
1.2	Desktop Environment Usage	3
1.3	Operating System Onion Rings	4
1.4	GNU Core Utils Logo	7
1.5	ls -l Output	8
1.6	Linux Filesystem Hierarchy	25
1.7	Relative Path	27
1.8	File Permissions	29
1.9	Octal Permissions	35
T .	ist of Tables	
L	ist of Tables	
L :	ist of Tables Basic Shortcuts in Terminal	5
		5 7
1.1	Basic Shortcuts in Terminal	
1.1 1.2	Basic Shortcuts in Terminal	7
1.1 1.2 1.3	Basic Shortcuts in Terminal	7 10
1.1 1.2 1.3 1.4	Basic Shortcuts in Terminal	7 10 17
1.1 1.2 1.3 1.4 1.5	Basic Shortcuts in Terminal Basic Commands in GNU Core Utils Manual Page Sections Escape Characters in echo Date Format Specifiers	7 10 17 23

Essentials of Linux | 1

1.1 Introduction

1.1.1 What is Linux?

Definition 1.1.1 (Linux) Linux is a **kernel** that is used in many operating systems. It is open source and free to use. Linux is not an operating system unto itself, but the core component of it.

So what is **Ubuntu**? **Ubuntu** is one of the many *distributions* that use the Linux kernel. It is a complete operating system that is free to use and open source. It is based on the **Debian** distribution of Linux. There are many other *distributions* of Linux, such as:

- ▶ **Debian** Used primarily on servers, it is known for its stability.
 - Ubuntu A commercial distribution based on Debian which is popular among new users.
 - Linux Mint A distribution based on Ubuntu which is known for its ease of use. It is one of the distributions recommended to new users.
 - **Pop OS** A distribution based on Ubuntu which is known for its focus on developers, creators, and gamers.
 - and many more
- ▶ Red Hat Enterprise Linux (RHEL) A commercial distribution used primarily in enterprises. It is owned by Red Hat and is targeted primarily to companies with their free OS paid support model.
 - Fedora A community-driven distribution sponsored by Red Hat. It is known for its cutting-edge features and is used by developers. It remains on the upstream of RHEL, receiving new features before RHEL.
 - CentOS A discontinued distribution based on RHEL. It
 was known for its stability and was used in servers. It was
 downstream from RHEL.
 - CentOS Stream It is a midstream between the upstream development in Fedora Linux and the downstream development for Red Hat Enterprise Linux.
 - Rocky Linux A distribution created by the Rocky Enterprise Software Foundation after the announcement of discontinuation of CentOS. It is a downstream of RHEL that provides feature parity and binary compatibility with RHEL.
 - Alma Linux A distribution created by the CloudLinux team
 after the announcement of discontinuation of CentOS. It is a
 downstream of RHEL that provides feature parity and binary
 compatibility with RHEL.
- ► Arch Linux A community-driven distribution known for its simplicity and customizability. It is a *rolling release* distribution, which means that it is continuously updated. It is a bare-bones

1.1	Introduction	1
	What is Linux?	1
1.1.2	Desktop Environments .	2
	Window Managers	3
1.1.4	Why Linux?	4
1.1.5	What is Shell?	4
1.1.6	Shell vs Terminal	4
	Why the Command Line?	5
1.1.8	Command Prompt	6
1.2	Simple Commands in	
	GNU Core Utils	7
1.2.1	File System Navigation .	8
	Manuals	9
	System Information	11
	File Management	13
	Text Processing and	
	Pagers	16
1.2.6	Aliases and Types of	
		20
1.2.7	Commands	21
1.2.8	Date and Time	22
1.3	Navigating the File Sys-	
1.0	tem	25
131	What is a File System?	
	Why do we have proc and	23
1.0.2	sys directories?	26
133	Paths	
	Basic Commands for	۷,
1.0.1	Navigation	27
1 1		
1.4	File Permissions	29
	Read	30 30
	Execute	
	Interesting Caveats	
	Changing Permissions	
	Special Permissions	
	Octal Representation of	32
1.1./	Permissions	34
4 -		
1.5	Types of Files	37
	Regular Files	37
	Directories	37
	Symbolic Links	37
	Character Devices	37
	Block Devices	38
	Named Pipes	39
	Sockets	39
	Types of Regular Files	40
1.6	Inodes and Links	41
1.7	Types of Commands	42

Manjaro - A distribution based on Arch Linux which is known for its user-friendliness. It is a rolling release distribution that is easier to install for new users. It uses a different repository

distribution that lets the user decide which packages they want to

- for packages with additional testing.
 EndeavourOS A distribution based on Arch Linux which is known for its simplicity and minimalism. It is a rolling release distribution that is easier to install for new users. It uses the
- Artix Linux It uses the OpenRC init system instead of systemd. It also offers other *init systems* like runit, s6, dinit.

same repository for packages as **Arch Linux**.

- openSUSE It is a free and open-source Linux distribution developed by the openSUSE project. It is offered in two main variations: Tumbleweed, an upstream rolling release distribution, and Leap, a stable release distribution which is sourced from SUSE Linux Enterprise.
 - Tumbleweed Rolling Release upstream.
 - Leap Stable Release downstream.
- ► Gentoo A distribution known for its customizability and performance. It is a source-based distribution, which means that the user compiles the software from source code. It is known for its performance optimizations for the user's hardware.
- ► Void It is an independent rolling-release Linux distribution that uses the X Binary Package System package manager, which was designed and implemented from scratch, and the runit init system. Excluding binary kernel blobs, a base install is composed entirely of free¹ software.

1: "Free software" means software that respects users' freedom and community. Roughly, it means that the users have the freedom to run, copy, distribute, study, change and improve the software. Thus, "free software" is a matter of liberty, not price. To understand the concept, you should think of "free" as in "free speech," not as in "free beer." We sometimes call it "libre software," borrowing the French or Spanish word for "free" as in freedom, to show we do not mean the software is gratis.

You may have paid money to get copies of a free program, or you may have obtained copies at no charge. But regardless of how you got your copies, you always have the freedom to copy and change the software, even to sell copies.

- GNU on Free Software

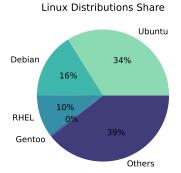


Figure 1.1: Linux Distributions Usage in 2024

- 2: GTK is a free software cross-platform widget toolkit for creating graphical user interfaces.
- 3: Ubuntu used to ship with Unity as the default desktop environment, but switched to GNOME in 2017.
- 4: Qt is cross-platform application development framework for creating graphical user interfaces.

1.1.2 Desktop Environments

Definition 1.1.2 (Desktop Environment) A desktop environment is a collection of software designed to give functionality and a certain look and feel to a desktop operating system. It is a combination of a window manager, a file manager, a panel, and other software that provides a graphical user interface and utilities to a regular desktop user, such as volume and brightness control, multimedia applications, settings panels, etc. This is only required by desktop (and laptop) uses and are not present on server instances.

There are many desktop environments available for Linux, but the important ones are:

- ► GNOME One of the most popular desktop environments for Linux. It is known for its simplicity and ease of use. It is the default desktop environment for many distributions, including Ubuntu. It is based on the GTK Toolkit. ² Popular distros shipping by default with GNOME are Fedora, RHEL, CentOS, Debian, Zorin, and Ubuntu. ³
- ► KDE Plasma A highly customizable desktop environment based on the Qt Tookit. ⁴ Many distributions like Slackware and Open-

SUSE ship with KDE Plasma as the default desktop environment, and most others have the option to install with KDE Plasma. Ubuntu's KDE Plasma variant is called **Kubuntu**.

- ➤ Xfce A lightweight desktop environment known for its speed and simplicity. It is based on the GTK Toolkit. It is used in many distributions like Xubuntu, Manjaro, and Fedora.
- ► LXQt A lightweight desktop environment known for its speed and simplicity. It is based on the Qt Toolkit. It is used in many distributions like Lubuntu.
- ► Cinnamon
- ► MATE

It is important to note that although some distributions come pre-bundled with certain Desktop Environments, it doesn't mean that you cannot use another DE with it. DE are simply packages installed on your distribution, and almost all the popular DEs can be installed on all distributions. Many distributions also come with multiple pre-bundled desktop environments due to user preferences. Most server distributions and some enthusiast distributions come with no pre-bundled desktop environment, and let the user determine which one is required, or if one is required.

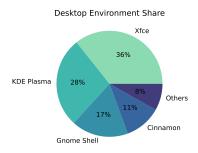


Figure 1.2: Desktop Environment Usage in 2022

1.1.3 Window Managers

Definition 1.1.3 (Window Manager) A window manager is system software that controls the placement and appearance of windows within a windowing system in a graphical user interface. It is a part of the desktop environment, but can also be used standalone. It is responsible for the appearance and placement of windows, and can also provide additional functionality like virtual desktops, window decorations, window title bars, and tiling.

Although usually bundled with a desktop environment, many window managers are also standalone and installed separately by the user if they don't want to use all the application from a single desktop environment.

Some popular window managers are:

- ► Openbox A lightweight window manager known for its speed and simplicity. It is used in many distributions like Lubuntu.
- ▶ i3 It is a tiling window manager ⁵ which is usually one of the first window managers that users try when they want to move away from a desktop environment and to a tiling window manager.
- ➤ awesome A tiling window manager that is highly configurable and extensible. It is written in Lua and is known for its beautiful configurations.
- ▶ **bspwm** A tiling window manager. It is based on binary space partitioning.
- dwm A dynamic tiling window manager that is known for its simplicity and minimalism. It is written in C and is highly configurable.

5: A tiling window manager is a window manager that automatically splits the screen into non-overlapping frames, which are used to display windows. Most desktop environments ship with a **floating** window manager instead, which users of other operating systems are more familiar with.

6: Although Linux is just a kernel and not an entire operating system, throughout this book I would be referring to GNU/Linux, the combination of GNU core utilities and the Linux kernel, as Linux in short.

other application programs layer shell utility programs | shell utility

Figure 1.3: Operating System Onion Rings - The layers of an operating system

- 7: POSIX, or Portable Operating System Interface, is a set of standards that define the interfaces and environment that operating systems use to access POSIX-compliant applications. POSIX standards are based on the Unix operating system and were released in the late 1980s.
- 8: Fish is a non-POSIX compliant shell that is known for its features like autosuggestion, syntax highlighting, and tab completions. Although a useful alternative to other shells for scripting, it should not be set as the default shell.

1.1.4 Why Linux?

You might be wondering "Why should I use Linux?" Most people use either **Windows** or **Mac** on their personal computers. Although these consumer operating systems get the job done, they don't let the user completely control their own hardware and software. Linux ⁶ is a free and open-source operating system that gives the user complete control over their system. It is highly customizable and can be tailored to the user's needs. It is also known for its stability and security. It is used in almost all servers, supercomputers, and embedded systems. It is also used in many consumer devices like Android phones, smart TVs, and smartwatches.

In this course we will be covering how to navigate the linux file system, how to manage files, how to manage the system, and how to write scripts to automate tasks. In the later part of the course we go over concepts such as pattern matching and text processing.

This course does not go into details of the linux kernel, but rather attempts to make the reader familiar with the *GNU core utils* and able to navigate around a linux server easily.

1.1.5 What is Shell?

The **kernel** is the core of the operating system. It is responsible for managing the hardware and providing services to the user programs. The **shell** is the interface between the user and the kernel (Figure 1.3). Through the **shell** we can run many commands and utilities, as well as some inbuilt features of the shell.

Definition 1.1.4 (Shell) A shell is a command-line interpreter that provides a way for the user to interact with the operating system. It takes commands from the user and executes them. It is a program that provides the user with a way to interact with the operating system.

The most popular shell in Linux is the **bash** shell. It is the default shell in most distributions. It is a POSIX-compliant ⁷ shell. There are many other shells available, such as **zsh**, **fish** ⁸ , **dash**, **csh**, **ksh**, and **tcsh**. Each shell has its own features and syntax, but most of the keywords and syntax are the same. In this course we will be covering only the **bash** shell and its syntax, but most of what we learn here is also applicable on other shells as well.

1.1.6 Shell vs Terminal

Definition 1.1.5 (Terminal) A terminal is a program that provides a way to interact with the shell. It is a program that provides a text-based interface to the shell. It is also known as a terminal emulator.

The terminal is the window that you see when you open a terminal program. It provides a way to interact with the shell. The shell is the program that interprets the commands that you type in the terminal.

The terminal is the window that you see, and the shell is the program that runs in that window. Whereas the shell is the application that is parsing your input and running the commands and keywords, the terminal is the application that lets you see the shell graphically. There are multiple different terminal emulators, providing a lot of customization and beautification to the terminal, as well as providing useful features such as *scroll back*, copying and pasting, and so on.

Some popular terminal emulators are:

- ▶ **gnome-terminal** The default terminal emulator for the GNOME desktop environment.
- ▶ **konsole** The default terminal emulator for the KDE desktop environment.
- ► xfce4-terminal The default terminal emulator for the Xfce desktop environment.
- ▶ alacritty A terminal emulator known for its speed and simplicity.
- terminator A terminal emulator known for its features like splitting the terminal into multiple panes.
- ▶ tilix A terminal emulator known for its features like splitting the terminal into multiple panes.
- ▶ st A simple terminal emulator known for its simplicity and minimalism.
- ▶ urxvt
- ► kitty
- **▶** terminology

In most terminal emulators, there are some basic shortcuts that can be used to make the terminal experience more efficient. Some of the basic shortcuts are listed in Table Table 1.1.

Shortcut	Description
	Description
Ctrl + C	Terminate the current process
Ctrl + D	Exit the shell
Ctrl + L	Clear the terminal screen
Ctrl + A	Move the cursor to the beginning of the line
Ctrl + E	Move the cursor to the end of the line
Ctrl + U	Delete from the cursor to the beginning of the line
Ctrl + K	Delete from the cursor to the end of the line
Ctrl + W	Delete the word before the cursor
Ctrl + Y	Paste the last deleted text
Ctrl + R	Search the command history
Ctrl + Z	Suspend the current process
Ctrl + \	Terminate the current process
Ctrl + S	Pause the terminal output
Ctrl + Q	Resume the terminal output

Table 1.1: Basic Shortcuts in Terminal

1.1.7 Why the Command Line?

Both the command line interface (CLI) and the graphical user interface (GUI) are simply shells over the operating system's kernel. They let you interact with the kernel, perform actions and run applications.

GUI:

The GUI requires a mouse and a keyboard, and is more intuitive and easier to use for beginners. But it is also slower and less efficient than the CLI. The GUI severely limits the user's ability to automate tasks and perform complex operations. The user can only perform those operations that the developers of the GUI have thought of and implemented.

CLI:

The CLI is faster and more efficient than the GUI as it lets the user use the keyboard to perform actions. Instead of clicking on pre-defined buttons, the CLI lets you construct your instruction to the computer using syntax and semantics. The CLI lets you combine simple commands that do one thing well to perform complex operations. The biggest advantage of the CLI is that it lets you automate tasks. It might be faster for some users to rename a file from **file1** to **file01** using the GUI, but it will always be faster to automate this using the CLI if you want to do this for thousands of files in the folder.

In this course we will be learning how to use the CLI to navigate the file system, manage files, manage the system, process text, and write scripts to automate tasks.

1.1.8 Command Prompt

The command prompt is the text that is displayed in the terminal to indicate that the shell is ready to accept commands. It usually ends with a \$ or a # symbol. The \$ symbol indicates that the shell is running as a normal user, and the # symbol indicates that the shell is running as the root user. The root user has complete control over the system and can perform any operation on the system.

An example of a command prompt is:

username@hostname:~\$

Here, username is the name of the user, hostname is the name of the computer, and \$ indicates that the shell is running as a normal user. The \sim symbol indicates that the current working directory is the user's home directory. 9 This prompt can be changed and customized according to the user's preferences using the PS1 variable discussed in Chapter 2.

^{9:} The home directory is the directory where the user's files and settings are stored. It is usually located at /home/username. This can be shorted to ~ in the shell.

1.2 Simple Commands in GNU Core Utils

Definition 1.2.1 (GNU Core Utils) The GNU Core Utilities are the basic file, shell, and text manipulation utilities of the GNU operating system. These are the utilities that are used to interact with the operating system and perform basic operations. ^a

The shell lets you simply type in the name of the command and press enter to run it. You can also pass arguments to the command to modify its behavior. Although the commands are simple, they are powerful and can be combined to perform complex operations. ¹⁰

Some basic commands in the core-utils are listed in Table Table 1.2.

Command	Description
ls	List the contents of a directory
cd	Change the current working directory
pwd	Print the current working directory
mkdir	Create a new directory
rmdir	Remove a directory
touch	Create a new file
rm	Remove a file
ср	Copy a file
mv	Move a file
echo	Print a message
cat	Concatenate and display the contents of a file
less	Display the contents of a file one page at a time
head	Display the first few lines of a file
tail	Display the last few lines of a file
find	Find files and directories
locate	Find files and directories
which	Find the location of a command
uname	Print system information
ps	Display information about running processes
kill	Terminate a process
chmod	Change the permissions of a file
chown	Change the owner of a file
chgrp	Change the group of a file
date	Print the current date and time
cal, ncal	Print a calendar
df	Display disk space usage
du	Display disk usage
free	Display memory usage
top	Display system information
history	Display the command history
sleep	Pause the shell for a specified time
true	Do nothing, successfully
false	Do nothing, unsuccessfully
tee	Read from stdin and write to stdout and files
whoami	Print the current user
groups	Print the groups the user belongs to
clear	Clear the terminal screen
exit	Exit the shell



Figure 1.4: GNU Core Utils Logo

10: The combination of commands to perform complex operations is called *piping*. This will be covered later.

Table 1.2: Basic Commands in GNU Core Utils

a GNU Core Utils

1.2.1 File System Navigation

pwd:

The pwd command prints the current working directory. The current working directory is the directory that the shell is currently in. The shell starts in the user's home directory when it is opened. The pwd command prints the full path of the current working directory.

ls:

The ls command lists the contents of a directory. By default, it lists the contents of the current working directory. The ls command can take arguments to list the contents of a different directory.

```
$ $ ls
2 Desktop Documents Downloads Music Pictures Videos
```

We can also list hidden files ¹¹ using the -a flag.

```
1 $ ls -a
2 . ....bashrc Desktop Documents Downloads Music Pictures
Videos
```

Here, the . and . . directories are special directories. The . directory is the current directory, and the . . directory is the parent directory. The .bashrc file is a configuration file for the shell which is a hidden file.

ls can also list the details of the files using the -l flag.

```
$\frac{1}{2}$ total 24

$\frac{1}{3}$ drwxr-xr-x 2 username group 4096 Mar 1 12:00 Desktop

$\frac{4}{4}$ drwxr-xr-x 2 username group 4096 Mar 1 12:00 Documents

$\frac{5}{6}$ drwxr-xr-x 2 username group 4096 Mar 1 12:00 Downloads

$\frac{6}{6}$ drwxr-xr-x 2 username group 4096 Mar 1 12:00 Music

$\frac{7}{6}$ drwxr-xr-x 2 username group 4096 Mar 1 12:00 Pictures

$\frac{8}{6}$ drwxr-xr-x 2 username group 4096 Mar 1 12:00 Videos
```

As seen in Figure 1.5, the first column is the file type and permissions. The second column is the number of links to the file or directory. The third and fourth columns are the owner and group of the file or directory. The fifth column is the size of the file or directory. The sixth, seventh, and eighth columns are the last modified date and time of the file or directory. The ninth column is the name of the file or directory. ¹²

We can also list the inode numbers ¹³ using the -i flag.

```
1 $ ls -i
2 123456 Desktop 123457 Documents 123458 Downloads 123459 Music
123460 Pictures 123461 Videos
```

Inodes will be discussed in detail later in the course.

cd:

The cd command changes the current working directory. It takes the path to the directory as an argument.

11: Hidden files are files whose names start with a dot. They are hidden by default in the 1s command.

```
# of Maralinks owner of file Size in Buffes Directory or File Name

**The Name**

**Th
```

Figure 1.5: ls -l Output

12: More details about the file permissions and the file types will be covered later in the course.

13: An inode is a data structure on a filesystem on Linux and other Unix-like operating systems that stores all the information about a file except its name and its actual data. This includes the file type, the file's owner, the file's group, the file's permissions, the file's size, the file's last modified date and time, and the file's location on the disk. The inode is the reference pointer to the data in the disk.

The cd command can also take the ~ symbol as an argument to change to the user's home directory. This is the default behavior of the cd command when no arguments are passed.

```
1 $ cd
2 $ pwd
3 /home/username
```

If we want to go back to the previous directory, we can use the - symbol as an argument to the cd command. $^{\rm 14}$

```
1    $ cd Documents
2    $ pwd
3    /home/username/Documents
4    $ cd -
5    $ pwd
6    /home/username
```

14: This internally uses the OLDPWD environment variable to change the directory. More about variables will be covered later in the course.

Question 1.2.1 ls can only show files and directories in the **cwd**¹⁵, not subdirectories. True or False?

15: cwd means Current Working Directory

Answer 1.2.1 False. 1s can show files and directories in the cwd, and also in subdirectories. The -R flag can be used to show files and directories in subdirectories, recursively.

1.2.2 Manuals

man:

How to remember so many flags and options for each of the commands? The man command is used to display the manual pages for a command.

Definition 1.2.2 (Manual Pages) Manual pages are a type of software documentation that provides details about a command, utility, function, or file format. They are usually written in a simple and concise manner and provide information about the command's syntax, options, and usage.

```
1 s man ls
```

This will display the manual page for the ls command. The manual page is divided into sections, and you can navigate through the sections using the arrow keys. Press q to exit the manual page.

Example manual page:

```
\begin{array}{c|c} 1\\ LS(1) \end{array} User Commands \begin{array}{c|c} LS(1) \end{array}
```

```
4 NAME
5
          ls - list directory contents
6
  SYNOPSIS
          ls [OPTION]... [FILE]...
8
9
  DESCRIPTION
10
11
          List information about the FILEs (the current directory by
       default). Sort entries alphabetically if none of -cftuvSUX
       nor --sort is specified.
12
         Mandatory arguments to long options are mandatory for short
13
        options too.
14
15
          -a, --all
                 do not ignore entries starting with .
16
17
          -A, --almost-all
18
                 do not list implied . and ..
19
20
```

The manual page provides information about the command, its syntax, options, and usage. It is a good practice to refer to the manual page of a command before using it.

To exit the manual page, press q.

There are multiple sections in the manual page, man takes the section number as an argument to display the manual page from that section.

```
1 $ man 1 ls
```

This will display the manual page for the 1s command from section 1. The details of the sections can be seen in Table Table 1.3.

Table 1.3: Manual Page Sections

Section	Description
1	User Commands
2	System Calls
3	Library Functions
4	Special Files usually found in /dev
5	File Formats and conventions
6	Games
7	Miscellaneous
8	System Administration
9	Kernel Developer's Manual

Man pages only provide information about the commands and utilities that are installed on the system. They do not provide information about the shell builtins or the shell syntax. For that, you can refer to the shell's documentation or use the **help** command.

Some commands also have a --help flag that displays the usage and options of the command.

Some commands have their own **info** pages, which are more detailed than the **man** pages.

To be proficient with shell commands, one needs to read the man, info, and help pages. 16

^{16:} An useful video by Robert Elder about the differences between man, info, and help can be found on YouTube.

Exercise 1.2.1 Run man, info, and --help on all the commands discussed in this section. Note the differences in the information provided by each of them. Read the documentations carefully and try to understand how each command works, and the pattern in which the documentations are written.

1.2.3 System Information

uname:

The uname command prints system information. It can take flags to print specific information about the system. By default, it prints only the kernel name.

```
1 $ uname
2 Linux
```

The -a flag prints all the system information. ¹⁷

```
$ uname -a
Linux rex 6.8.2-arch2-1 #1 SMP PREEMPT_DYNAMIC Thu, 28 Mar 2024
17:06:35 +0000 x86_64 GNU/Linux
```

ps:

The ps command displays information about running processes. By default, it displays information about the processes run by the current user that are running from a terminal. 18

There are a lot of flags that can be passed to the ps command to display more information about the processes. These will be covered in Chapter 3.

Remark 1.2.1 ps has three types of options:

- UNIX options
- ► BSD options
- ▶ GNU options

The UNIX options are the preceded by a hyphen (-) and may be grouped. The BSD options can be grouped, but should not be preceded by a hyphen (-). The GNU options are preceded by two hyphens (-). These are also called long options.

The same action can be performed by using different options, for example, ps -ef and ps aux are equivalent, although first is using **UNIX** options, and the latter is using **BSD** options.

Another difference in **GNU core utils** and **BSD utils** is that the **GNU** utils have long options, whereas the **BSD** utils do not.

BSD utils also usually do not support having flags after the positional arguments, whereas most **GNU** utils are fine with this.

17: Here rex is the hostname of the system, 6.8.2-arch2-1 is the kernel version, x86_64 is the architecture, and GNU/Linux is the operating system.

18: The **PID** is the process ID, the **TTY** is the terminal the process is running from, the **TIME** is the time the process has been running, and the **CMD** is the command that is running.

kill:

The kill command is used to terminate a process. It takes the process ID as an argument.

1 \$ kill 12345

The kill command can also take the **signal** number as an argument to send a signal to the process. For example, the SIGKILL signal can be sent to the process to terminate it. ¹⁹

1 \$ **kill** -9 12345

19: The SIGKILL signal is used to terminate a process immediately. It cannot be caught or ignored by the process. It is 1 numbered as 9.

free:

The free command is used to display the amount of free and used memory in the system.

1	\$ free					
2		total	used	free	shared	buff/
	cache	available				
3	Mem:	8167840	1234560	4567890	123456	
	2367890	4567890				
4	Swap:	2097148	0	2097148		

The free command can take the -h flag to display the memory in human-readable format.

1	\$ free -h					
2		total	used	free	shared	buff/
	cache	available				
3	Mem:	7.8Gi	1.2Gi	4.3Gi	120Mi	
	2.3Gi	4.3Gi				
4	Swap:	2.0Gi	0B	2.0Gi		

df:

The df command is used to display the amount of disk space available on the filesystems.

```
1 $ df

2 Filesystem 1K-blocks Used Available Use% Mounted on

3 /dev/sda1 12345678 1234567 11111111 10% /

4 /dev/sda2 12345678 1234567 11111111 10% /home
```

The df command can take the -h flag to display the disk space in human-readable format.

```
1 $ df -h
2 Filesystem Size Used Avail Use% Mounted on
3 /dev/sda1 12G 1.2G 9.9G 11% /
4 /dev/sda2 12G 1.2G 9.9G 11% /home
```

du:

The du command is used to display the disk usage of directories and files. By default, it displays the disk usage of the current directory.

```
7 4 ./Videos
8 28
```

The du command can take the -h flag to display the disk usage in human-readable format. The -s flag displays the total disk usage of the directory.

Question 1.2.2 How to print the kernel version of your system?

Answer 1.2.2 uname -r will print the kernel version of your system. uname is a command to print system information. The -r flag is to print the kernel release. There are other flags to print other system information.

We can also run uname -a to get all fields and extract only the kernel info using commands taught in later weeks.

Question 1.2.3 How to see how long your system is running for? What about the time it was booted up?

Answer 1.2.3 uptime will show how long the system is running for. uptime -s will show the time the system was booted up. The -s flag is to show the time of last boot.

Question 1.2.4 How to see the amount of free memory? What about free hard disk space? If we are unable to understand the big numbers, how to convert them to human readable format? What is difference between MB and MiB?

Answer 1.2.4 free will show the amount of free memory. df will show the amount of free hard disk space.

df -h and free -h will convert the numbers to human readable format.

MB is Megabyte, and MiB is Mebibyte.

1 MB = 1000 KB, 1 GB = 1000 MB, 1 TB = 1000 GB, this is SI unit. 1 MiB = 1024 KiB, 1 GiB = 1024 MiB, 1 TiB = 1024 GiB, this is 2^{10} unit.

1.2.4 File Management

file:

The file command is used to determine the type of a file. It can take multiple file names as arguments.

```
$ file file1
file1: ASCII text
file /bin/bash
```

```
/bin/bash: ELF 64-bit LSB shared object, x86-64, version 1 (SYSV ), dynamically linked, interpreter /lib64/ld-linux-x86-64.so .2, for GNU/Linux 3.2.0, BuildID[sha1]=1234567890abcdef, stripped
```

mkdir:

The mkdir command is used to create new directories. It can take multiple directory names as arguments.

Exercise 1.2.2 Run man ls to find out what the -F flag does, and why we used it in the above example.

touch:

The touch command is used to create new files. It can take multiple file names as arguments. If a file is already present, the touch command updates the last modified date and time of the file, but does not modify the contents of the file.

```
1  $ touch file1 file2 file3
2  $ ls -l
3   -rw-r--r-- 1 username group 0 Mar  1 12:00 file1
4   -rw-r--r-- 1 username group 0 Mar  1 12:00 file2
5   -rw-r--r-- 1 username group 0 Mar  1 12:00 file3
6  $ sleep 60
7  $ touch file3
8  $ ls -l
9   -rw-r--r-- 1 username group 0 Mar  1 12:00 file1
10   -rw-r--r-- 1 username group 0 Mar  1 12:00 file2
11   -rw-r--r-- 1 username group 0 Mar  1 12:01 file3
```

Exercise 1.2.3 Notice the difference in the last modified date and time of the file3 file from the other files. Also notice the sleep command used to pause the shell for 60 seconds.

rmdir:

The rmdir command is used to remove directories. It can take multiple directory names as arguments.

Remark 1.2.2 The rmdir command can only remove empty directories. This is a safety feature so that users dont accidentally delete directories with files in them. To remove directories with files in them along with those files, use the rm command.

rm:

The rm command is used to remove files and directories. It can take multiple file and directory names as arguments.

```
1  $ touch file1 file2 file3
2  $ ls -F
3  file1 file2 file3
4  $ rm file1 file2
5  $ ls -F
6  file3
```

However, using rm to delete a directory will give an error.

```
1    $ mkdir a
2    $ rm a
3    rm: cannot remove 'a': Is a directory
```

This is because the rm command does not remove directories by default. This is a safety feature to prevent users from accidentally deleting directories with files in them.

To remove directories along with their files, use the -r flag.

```
1 | $ rm -r a
```

To force the removal of files and directories without a confirmation, use the -f flag.

Warning 1.2.1 The rm command is a dangerous command. It does not move the files to the trash, but permanently deletes them. Be extremely careful when using the rm command. Only use the -f flag if you are absolutely sure that you want to delete the files.

To force rm to always ask for confirmation before deleting files, use the -i flag.

cp:

The cp command is used to copy files. It takes the source file and the destination file as arguments.

```
1  $ touch file1
2  $ ls -F
3  file1
4  $ cp file1 file2
5  $ ls -F
6  file1 file2
```

The cp command can also take the -r flag to copy directories.

```
1    $ mkdir a
2    $ touch a/file1
3    $ cp -r a b
4    $ ls -R
5    .:
6    a/ b/
7    ./a:
```

```
9 file1
10
11 ./b:
12 file1
```

Exercise 1.2.4 Why did we use the -R flag in the above example? What does it do?

There are three ways to copy files using cp:

```
1 SYNOPSIS
2 cp [OPTION]... [-T] SOURCE DEST
3 cp [OPTION]... SOURCE... DIRECTORY
4 cp [OPTION]... -t DIRECTORY SOURCE...
```

Exercise 1.2.5 There are three ways of running the cp command to copy a file. Here we have demonstrated only one. Read the manual page of the cp command to find out the other two ways and try them out yourself.

mv:

20: This means that mv also has three ways of running it.

The mv command is used to move files. The syntax is similar to the cp command. ²⁰ It is used to move files from one location to another, or to rename files.

```
1  $ touch file1
2  $ ls -F
3  file1
4  $ mv file1 file2
5  $ ls -F
6  file2
```

Exercise 1.2.6 Create a directory dir1 using the mkdir command, then create a file file1 inside dir1. Now move (rename) the dir1 directory to dir2 using the mv command. The newly created directory should be named dir2 and should contain the file1 file. Were you require to use the -r flag with the mv command like you would have in cp command?

1.2.5 Text Processing and Pagers

echo:

The echo command is used to print a message to the terminal. It can take multiple arguments and print them to the terminal.

```
$ echo Hello, World!
Hello, World!
```

The echo command can also take the -e flag to interpret backslash escapes.

```
$ echo -e "Hello, \nWorld!"
Hello
World!
```

Some escape characters in **echo** are listed in Table Table 1.4.

Escape Description 11 backslash \a alert (BEL) backspace \b produce no further output \ c escape \e \f form feed \n new line carriage return ١r horizontal tab \t vertical tab \ONNN byte with octal value NNN (1 to 3 digits) byte with hexadecimal value HH (1 to 2 digits) \xHH

Exercise 1.2.7 Run the command echo -e "\x41=\0101" and try to understand the output and the escape characters used.

cat:

The cat command is used to concatenate and display the contents of files.

```
$ cat file1
Hello, World! from file1
```

The cat command can take multiple files as arguments and display their contents one after another.

```
$ cat file1 file2
Hello, World! from file1
Hello, World! from file2
```

less:

Sometimes the contents of a file are too large to be displayed at once. Nowadays modern terminal emulators can scroll up and down to view the contents of the file, but actual ttys cannot do that. To view the contents of a file one page at a time, use the less command. less is a pager program that displays the contents of a file one page at a time. ²¹

```
$ less file1
```

To scroll up and down, use the arrow keys, or the j and k keys. ²² Press q to exit the less command.

head:

The head command is used to display the first few lines of a file. By default, it displays the first 10 lines of a file.

```
$ head /etc/passwd
root:x:0:0:root:/root:/usr/bin/bash
bin:x:1:1::/:/usr/bin/nologin
```

Table 1.4: Escape Characters in echo

- 21: more is another pager program that displays the contents of a file one page at a time. It is older than less and has fewer features. less is an improved version of more and is more commonly used. Due to this, it is colloquially said that "less is more", as it has more features.
- 22: Using j and k to move the cursor up and down is a common keybinding in many terminal applications. This originates from the vi text editor which will be covered later in the course.

```
daemon:x:2:2::/:/usr/bin/nologin
4
    mail:x:8:12::/var/spool/mail:/usr/bin/nologin
5
    ftp:x:14:11::/srv/ftp:/usr/bin/nologin
6
    http:x:33:33::/srv/http:/usr/bin/nologin
    nobody:x:65534:65534:Kernel Overflow User:/:/usr/bin/nologin
8
    dbus:x:81:81:System Message Bus:/:/usr/bin/nologin
    systemd-coredump:x:984:984:systemd Core Dumper:/:/usr/bin/
10
       nologin
11
     systemd-network:x:982:982:systemd Network Management:/:/usr/bin/
       nologin
```

23: We can also directly run head -5 /etc/passwd to display the first 5 lines of the file.

The head command can take the -n flag to display the first n lines of a file. 23

Remark 1.2.3 Here we are listing the file /etc/passwd which contains information about the users on the system. The file will usually be present on all Unix-like systems and have a lot of system users. ²⁴

24: A system user is a user that is used by the system to run services and daemons. It does not belong to any human and usually is not logged into. System users have a user ID less than 1000.

tail:

The tail command is used to display the last few lines of a file. By default, it displays the last 10 lines of a file.

```
$ tail /etc/passwd

rtkit:x:133:133:RealtimeKit:/proc:/usr/bin/nologin

sddm:x:964:964:SDDM Greeter Account:/var/lib/sddm:/usr/bin/
    nologin

usbmux:x:140:140:usbmux user:/:/usr/bin/nologin

sayan:x:1000:1001:Sayan:/home/sayan:/bin/bash

qemu:x:962:962:QEMU user:/:/usr/bin/nologin

cups:x:209:209:cups helper user:/:/usr/bin/nologin

dhcpcd:x:959:959:dhcpcd privilege separation:/:/usr/bin/nologin

saned:x:957:957:SANE daemon user:/:/usr/bin/nologin
```

The tail command can take the -n flag to display the last n lines of a file.

```
1    $ tail -n 5 /etc/passwd
2    sayan:x:1000:1001:Sayan:/home/sayan:/bin/bash
3    qemu:x:962:962:QEMU user:/:/usr/bin/nologin
4    cups:x:209:209:cups helper user:/:/usr/bin/nologin
5    dhcpcd:x:959:959:dhcpcd privilege separation:/:/usr/bin/nologin
6    saned:x:957:957:SANE daemon user:/:/usr/bin/nologin
```

Exercise 1.2.8 Notice that the UID (3rd column) of the sayan user is 1000. The last column is /bin/bash instead of /usr/bin/nologin like others. This is because it is a regular user and not a system user.

The wc command is used to count the number of lines, words, and characters in a file. By default, it displays the number of lines, words, and characters in a file.

We can also use the -1, -w, and -c flags to display only the number of lines, words, and characters respectively.

```
1 $ wc -l /etc/passwd 2 43 /etc/passwd
```

Question 1.2.5 Can we print contents of multiple files using a single command?

Answer 1.2.5 cat file1 file2 file3 will print the contents of file1, file2, and file3 in the order given. The contents of the files will be printed one after the other.

Question 1.2.6 Can cat also be used to write to a file?

Answer 1.2.6 Yes, cat > file1 will write to file1. The input will be taken from the terminal and written to file1. The input will be written to file1 until the user presses Ctrl+D to indicate end of input. This is redirection, which we see in later weeks.

Question 1.2.7 How to list only first 10 lines of a file? How about first 5? Last 5? How about lines 105 to lines 152?

Answer 1.2.7 head filename will list the first 10 lines of filename.

head -n 5 filename will list the first 5 lines of filename.

tail -n 5 filename will list the last 5 lines of filename.

head -n 152 filename | tail -n 48 will list lines 105 to 152 of filename. This uses | which is a pipe, which we will see in later weeks.

Question 1.2.8 Do you know how many lines a file contains? How can we count it? What about words? Characters?

Answer 1.2.8 wc filename will count the number of lines, words, and characters in filename.

wc -l filename will count the number of lines in filename.

wc -w filename will count the number of words in filename.

wc -c filename will count the number of characters in filename.

Question 1.2.9 How to delete an empty directory? What about non-empty directory?

Answer 1.2.9 rmdir dirname will delete an empty directory. rm -r dirname will delete a non-empty directory.

Question 1.2.10 How to copy an entire folder to another name? What about moving?

Why the difference in flags?

Answer 1.2.10 cp -r sourcefolder targetfolder will copy an entire folder to another name.

mv sourcefolder targetfolder will move an entire folder to another name.

The difference in flags is because cp is used to copy, and mv is used to move or rename a file or folder. The -r flag is to copy recursively, and is not needed for mv as it is not recursive and simply changes the name of the folder (or the path).

1.2.6 Aliases and Types of Commands

alias:

The alias command is used to create an alias for a command. An alias is a custom name given to a command that can be used to run the command. ²⁵

25: Aliases are used to create shortcuts for long commands. They can also be used to create custom commands.

The alias ll is created for the ls -l command.

Warning 1.2.2 Be careful when creating aliases. Do not create aliases for existing commands. This can lead to confusion and errors.

But this alias is temporary and will be lost when the shell is closed. To make the alias permanent, add it to the shell configuration file. For **bash**, this is the .bashrc file in the home directory.

```
1 $ echo "alias ll='ls -l'" >> ~/.bashrc
```

This will add the alias to the .bashrc file. To make the changes take effect, either close the terminal and open a new one, or run the source command.

```
1 $ source ~/.bashrc
```

Warning 1.2.3 Be careful when editing the shell configuration files. A small mistake can lead to the shell not working properly. Always keep a backup of the configuration files before editing them.

We can see the aliases that are currently set using the alias command.

```
s alias alias ll='ls -l'
```

We can also see what a particular alias expands to using the alias command with the alias name as an argument.

```
s alias ll alias ll='ls -l'
```

To remove an alias, use the unalias command.

1 \$ unalias ll

Exercise 1.2.9 Create an alias la for the ls -a command. Make it permanent by adding it to the .bashrc file. Check if the alias is set using the alias command.

Exercise 1.2.10 Create an alias rm for the rm -i command. Make it permanent by adding it to the .bashrc file. Check if the alias is set using the alias command. Try to delete a file using the rm command. What happens?

1.2.7 User Management

whoami:

The whoami command is used to print the username of the current user.

```
sayan $ whoami
```

groups:

The groups command is used to display the groups that the current user belongs to.

passwd

The passwd command is used to change the password of the current user. The root user can also change the password of other users. ²⁶

who:

The who command is used to display the users who are currently logged in.

26: This executable is a special one, as it is a setuid program. This will be discussed in detail in Subsection 1.4.6.

```
1 $ who
  sayan
           tty2
                        2024-05-22 13:49
                        2024-05-22 15:58 (:0)
3
  sayan
           pts/0
  sayan
                        2024-05-22 15:58 (tmux(1082933).%2)
           pts/1
                        2024-05-22 15:58 (tmux(1082933).%1)
  sayan
           pts/2
           pts/3
                      2024-05-22 15:58 (tmux(1082933).%3)
  savan
6
                        2024-05-22 15:58 (tmux(1082933).%4)
           pts/4
  sayan
8
  sayan
           pts/5
                        2024-05-22 15:58 (tmux(1082933).%5)
  sayan
           pts/6
                        2024-05-22 15:58 (tmux(1082933).%6)
10
  sayan
           pts/7
                        2024-05-22 15:58 (tmux(1082933).%7)
  sayan
           pts/8
                        2024-05-22 15:58 (tmux(1082933).%8)
11
                        2024-05-22 15:58 (tmux(1082933).%9)
12 sayan
           pts/9
                        2024-05-22 17:58 (:0)
13 sayan
           pts/10
                        2024-05-22 18:24 (tmux(1082933).%10)
14 sayan
           pts/11
15 sayan
           pts/12
                        2024-05-22 18:24 (tmux(1082933).%11)
```

Exercise 1.2.11 Run the who command on the system commands VM. What is the output?

w:

The w command is used to display the users who are currently logged in and what they are doing.

```
1 $ W
2 19:47:07 up 5:57, 1 user, load average: 0.77, 0.80, 0.68
3 USER TTY LOGIN@ IDLE JCPU PCPU WHAT
4 sayan tty2 13:49 5:57m 19:10 21.82s dwm
```

This is different from the who command as it only considers the login shell. Here **dwm** is the window manager running on the tty2 terminal.

1.2.8 Date and Time

date:

The date command is used to print formatted date and time information. Without any arguments, it prints the current date and time.

```
1 $ date
2 Mon May 20 06:23:07 PM IST 2024
```

We can specify the date and time to be printed using the -d flag.

```
1 $ date -d "2020-05-20 00:30:45"
2 Wed May 20 12:30:45 AM IST 2020
3 $ date -d "2019-02-29"
4 date: invalid date '2019-02-29'
5 $ date -d "2020-02-29"
6 Sat Feb 29 12:00:00 AM IST 2020
```

Exercise 1.2.12 Why did we get an error when trying to print the date 2019-02-29?

We can also modify the format of the date and time using the + flag and different **format specifiers**. Some of the important format specifiers are

listed in Table Table 1.5. Rest of the format specifiers can be found in the date manual page.

Table 1.5: Date Format Specifiers

```
Specifier
          Description
   %Y
          Year
   %m
          Month
  %d
          Day
  %Н
          Hour
   %M
          Minute
  %S
          Second
  %A
          Full weekday name
   %В
          Full month name
          Abbreviated weekday name
   %a
          Abbreviated month name
   %b
```

We can even mention relative dates and times using the date command.

```
1    $ date -d "next year"
2    Tue May 19 06:23:07 PM IST 2025
3    $ date -d "next month"
4    Thu Jun 20 06:23:07 PM IST 2024
5    $ date -d "tomorrow"
6    Tue May 21 06:23:07 PM IST 2024
7    $ date -d "yesterday"
8    Sun May 19 06:23:07 PM IST 2024
```

cal:

The cal command is used to print a calendar. By default, it prints the calendar of the current month.

We can specify the month and year to print the calendar of that month and year.

There are multiple flags that can be passed to the cal command to display different types of calendars and of multiple months or of entire year.

Remark 1.2.4 In Linux, there are sometimes multiple implementations of the same command. For example, there are two implementations of the cal command, one in the **bsdmainutils** package, which is the **BSD** implementation and also includes another binary named **ncal** for printing the calendar in vertical format. The other implementation is in the **util-linux** package, which does not contain a **ncal** binary. The flags and the output of the cal command can differ between the two implementations.

Question 1.2.11 How to print the current date and time in some custom format?

Answer 1.2.11 date -d today +%Y-%m-%d will print the current date in the format YYYY-MM-DD. The format can be changed by changing the format specifiers. The format specifiers are given in the man date page. The -d today can be dropped, but is mentioned to show that the date can be changed to any date. It can be strings like '2024-01-01' or '5 days ago' or 'yesterday', etc.

These are some of the basic commands that are used in the terminal. Each of these commands has many more options and flags that can be used to customize their behavior. It is left as an exercise to the reader to explore the manual pages of these commands and try out the different options and flags.

Many of the commands that we have discussed here are also explained in the form of short videos on Robert Elder's Youtube Channel.

1.3 Navigating the File System

1.3.1 What is a File System?

Unlike Windows which has different drive letters for different partitions, Linux follows a unified file structure. The filesystem hierarchy is a tree of directories and files²⁷. The root²⁸ of the filesystem tree is the directory /. The basic filesystem hierarchy structure can be seen in Figure 1.6 and Table Table 1.6.

But what does so many directories mean? What do they do? What is the purpose of each directory?

Directory Path	Description	
/	Root directory	
/bin	Essential command binaries	
/boot	Static files of the bootloader	
/dev	Device files	
/etc	Host-specific system configuration	
/home	User home directories	
/lib	Essential shared libraries and kernel modules	
/media	Mount point for removable media	
/mnt	Mount point for mounting a filesystem temporarily	
/opt	Add-on application software packages	
/proc	Virtual filesystem providing process information	
/root	Home directory for the root user	
/run	Run-time variable data	
/sbin	Essential system binaries	
/srv	Data for services provided by the system	
/sys	Kernel and system information	
/tmp	Temporary files	
/usr	Secondary hierarchy	
/var	Variable data	

Some directories do not store data on the disk, but are used to store information about the system. These directories are called *virtual* directories. For example, the **/proc** directory is a virtual directory that provides information about the running processes. The **/sys** directory is another virtual directory that provides information about the system. The **/tmp** is a *volatile* directory whose data is deleted as soon as the system is turned off. The **/run** directory is another volatile directory that stores runtime data.

Rest directories are stored on the disk. The reason for having so many directories is to categorize the type of files they store. For example, all the executable binaries of different applications and utilities installed in the system is stored in **/bin** and **/sbin** directories. All the shared libraries installed on the system are stored in **/lib** directory. Sometimes some applications are installed in **/opt** directory which are not installed directly by the package manager. ²⁹

We also need to store the user's documents and files. This is done in the **/home** directory. Each user has their own directory in the **/home** directory. The root user's directory is **/root**. All the application's configuration files are stored in the user's home directory in the **/home** directory itself. This

27: A non-directory is a leaf node of a tree.

28: The root of a tree is the first node from which the tree originates. A tree can have only one root.

Table 1.6: Linux Filesystem Hierarchy

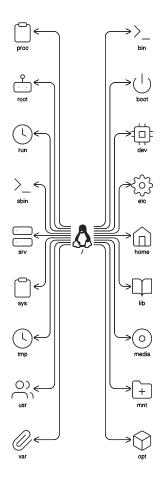


Figure 1.6: Linux Filesystem Hierarchy

29: In Linux, you do not install applications by downloading them from the internet and running an installer like in Windows. You use a package manager to install applications. The package manager downloads the application from the internet and installs it on your system automatically. This way the package manager can also keep track of the installed applications and their dependencies and whether they should be updated. This is similar to the *Play Store* on mobile phones.

separation of application binary and per-user application settings helps people to easily change systems but keep their own /home directory constant and in turn, also all their application settings.

Some settings however needs to be applied system-wide and for all users. These settings are stored in the **/etc** directory. This directory contains all the system-wide configuration files.

To boot up the system, the bootloader needs some files. These files are stored in the **/boot** directory. ³⁰ The bootloader is the first program that runs when the computer is turned on. It loads the operating system into memory and starts it.

Although the file system is a unified tree hierarchy, this doesn't mean that we cannot have multiple partitions on Linux: au contraire, it is easier to manage partitions on Unix. We simply need to mention which empty directory in the hierarchy should be used to mount a partition. As soon as that partition is mounted, it gets populated with the data stored on that disk with all the files and subdirectories, and when the device is unmounted the directory again becomes empty. Although a partition can be mounted on any directory, there are some dedicated folders in / as well for this purpose. For example, the /mnt directory is used to mount a filesystem temporarily, and the /media directory is used to mount removable media like USB drives, however it is not required to strictly follow this.

Finally, the biggest revelation in Linux is that, everything is a file. Not only are all the system configurations stored as **plain text** files which can be read by humans, but the processes running on your system are also stored as files in proc. Your kernel's interfaces to the applications or users are also simple files stored in sys. Biggest of all, even your hardware devices are stored as files in dev. ³¹

The /usr directory is a secondary hierarchy that contains subdirectories similar to those present in /. This was created as the olden system had started running out of disk space for the /bin and /lib directories. Thus another directory named usr was made, and subdirectiores like /usr/bin and /usr/lib were made to store half of the binaries and libraries. There wasn't however any rigid definition of which binary should go where. Modern day systems have more than enough disk space to store everything on one partition, thus the /bin and /lib dont really exist any more. If they do, they are simply shortcuts ³² to the /usr/bin and /usr/lib directories which are still kept for backwards compatibility.

These can also be loosely classified into *sharable* and *non-sharable* directories and *static* and *variable* directories as shown in Table Table 1.7.

	Sharable	Non-sharable
Static	/usr, /opt	/etc, /boot
Variable	/var/spool	/tmp, /var/log

30: Modern systems use **UEFI** instead of **BIOS** to boot up the system. The bootloader is stored in the **/boot/EFI** directory or in the **/efi** directory directly.

31: Device files are not stored as normal files on the disk, but are special files that the kernel uses to communicate with the hardware devices. These are either **block** or **character** devices. They are used to read and write data to and from the hardware devices.

32: Shortcuts in Linux are called *symbolic links* or *symlinks*.

Table 1.7: Linux Filesystem Directory Classification

1.3.2 Why do we have proc and sys directories?

33: Unix System V is one of the first commercial versions of the Unix operating system. It was originally developed by AT&T and first released in 1983.

The /proc tree originated in System V Unix 33 , where it only gave information about each running process, using a /proc/PID/ format.

Linux greatly extended that, adding all sorts of information about the running kernel's status. In addition to these read-only information files, Linux's /proc also has writable virtual files that can change the state of the running kernel. BSD ³⁴ type OSes generally do not have /proc at all, so much of what you find under proc is non-portable.

The intended solution for this mess in Linux's /proc is /sys. Ideally, all the non-process information that got crammed into the /proc tree should have moved to /sys by now, but historical inertia has kept a lot of stuff in /proc. Often there are two ways to effect a change in the running kernel: the old /proc way, kept for backwards compatibility, and the new /sys way that you're supposed to be using now.

34: BSD, or Berkeley Software Distribution, is a Unix-like operating system that was developed at the University of California, Berkeley. It was first released in 1977 and was based on the original Unix source code from AT&T. BSD is not linux, it is a totally different kernel, with similar core utils to GNU.

1.3.3 Paths

Whenever we open a terminal on a Linux system, we are placed in a directory. This is called the *current working directory*. All shells and applications have a current working directory from where they are launched.

To refer to and identify the directory you are talking about, we use a path.

Definition 1.3.1 (Path) Path is a traversal in the filesystem tree. It is a way to refer to a file or directory.

Absolute Path:

The traversal to the directory from the root directory is called the **absolute path**. For example, if we want to refer to the directory named **alex** inside the directory **home** in the root of the file system, then it is qualified as:

/home/alex

Relative Path:

The traversal to the directory from the current working directory is called the **relative path**. For example, if we want to refer to the directory named **alex** inside the directory **home** from the **/usr/share** directory, then it will be qualified as:

../../home/alex

Remark 1.3.1 The .. in the path refers to the parent directory. It is used in relative paths to refer to directories whose path requires travelling up the tree.

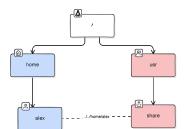


Figure 1.7: Relative Path

1.3.4 Basic Commands for Navigation

The file system can be navigated in the Linux command line using the following commands:

- ▶ pwd: Print the current working directory
- ▶ ls: List the contents of the current directory

in scripts.

35: **pushd** and **popd** are useful for quickly switching between directories

- ▶ cd: Change the current working directory
- ▶ mkdir: Create a new directory
- ► rmdir: Remove a directory
- ▶ touch: Create a new file
- ► rm: Remove a file
- ▶ pushd: Push the current directory to a stack
- ▶ **popd**: Pop the current directory from a stack³⁵

More details about these commands can be found in their respective man pages. For example, to find more about the **ls** command, you can type man ls.

Question 1.3.1 What is the command to list the contents of the current directory?

Answer 1.3.1 ls

Question 1.3.2 What is the command to list the contents of the current directory including hidden files?

Answer 1.3.2 ls -a

Question 1.3.3 What is the command to list the contents of the current directory in a long list format? (show permissions, owner, group, size, and time)

Answer 1.3.3 ls - l

Question 1.3.4 What is the command to list the contents of the current directory in a long list format and also show hidden files?

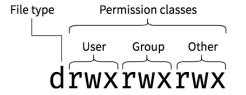
```
Answer 1.3.4 ls -al or ls -la or ls -l -a or ls -a -l
```

Question 1.3.5 The output of ls gives multiple files and directories in a single line. How can you make it print one file or directory per line?

Answer 1.3.5 ls -1

This can also be done by passing the output of ls to cat or storing the output of ls in a file and then using cat to print it. We will see these in later weeks.³⁶

1.4 File Permissions



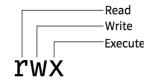


Figure 1.8: File Permissions

Definition 1.4.1 (File Permissions) File permissions define the access rights of a file or directory. There are three basic permissions: read, write, and execute. These permissions can be set for the owner of the file, the group of the file, and others.

We have already briefly seen how to see the permissions of a file using the ls -l command.

```
1  $ touch hello.txt
2  $ mkdir world
3  $ ls -l
4  total 4
5  -rw-r--r-- 1 sayan sayan 0 May 21 15:20 hello.txt
6  drwxr-xr-x 2 sayan sayan 4096 May 21 15:21 world
```

Here, the first column of the output of ls -l shows the permissions of the file or directory. As seen in Figure 1.8, the permissions are divided into four parts:

- ► The first character shows the type of the file. for a regular file and d for a directory and more. ³⁷
- ► The next three characters show the permissions for the owner of the file.
- ► The next three characters show the permissions for the group of the file
- ► The last three characters show the permissions for others.

Definition 1.4.2 (Owner) Although this can be changed, the owner of a file is usually the user who created it. All files in the filesystem have an owner. This is symbolically coded as \mathbf{u} .

Definition 1.4.3 (Group) The group of a file is usually the group of the user who created it. But it can also be changed to any other existing group in the system. All users in the group a have the same permissions on the file. This is symbolically coded as \mathbf{g} .

Definition 1.4.4 (Others) Others are all the users who are not the owner of the file and are not in the group of the file. This is symbolically coded as o.

37: There are other types of files as well, like \(\text{l for a symbolic link, c for a character device, b for a block device, s for a socket, and \(\text{p for a pipe}.\) These will be discussed later.

a except the owner of the file

There are three actions that can be performed on a file: read, write, and execute.

- ▶ **Read**: The read permission allows the file to be read. This is symbolically coded as **r**.
- ► Write: The write permission allows the file to be modified. This is symbolically coded as w.
- ► Execute: The execute permission allows the file to be executed. ³⁸ This is symbolically coded as **x**.

These however, have different meanings for files and directories.

1.4.1 Read

- ► For a file, the read permission allows the file to be read. You can use commands like cat or less to read the contents of the file if the user has **read** permissions.
- ► For a directory, the read permission allows the directory to be listed using ls.

1.4.2 Write

- ► For a file, the write permission allows the file to be modified. You can use commands like echo along with redirection ³⁹ or a text editor like vim or nano to write to the file if the user has write permissions.
- ► For a directory, the write permission allows the directory to be modified. You can create, delete, or rename files in the directory if the user has write permissions.

1.4.3 Execute

- ➤ For a file, the execute permission allows the file to be executed. This is usually only needed for special files like executables, scripts, or libraries. You can run the file as a program if the user has **execute** permissions.
- ► For a directory, the execute permission allows the directory to be traversed into. You can change to the directory if the user has **execute** permissions using cd. You can also only long-list the contents of the directory if the user has **execute** permissions on that directory.

1.4.4 Interesting Caveats

This causes some interesting edge-cases that one needs to be familiar with.

Cannot modify a file? Think again!

If you have **write** and **execute** permissions on a directory, even if you do not have **write** permission on a file inside the directory, you can **delete** the file due to your **write** permission on the directory, and then re-create

38: Executing a file means running the file as a program. For a directory, the execute permission allows the directory to be traversed into.

39: Redirection is a way to send the output of a command to a file.

the modified version of the file with the same name. But if you try to simply modify the file directly, you will get permission error.

```
$ mkdir test
    $ cd test
2
    $ echo "hello world" > file1
3
    $ chmod 400 file1  # 400 means read permission only
4
    $ cat file1
5
6
    hello world
    $ echo "hello universe" > file1 # unable to write
    -bash: file1: Permission denied
    $ rm file1 # can remove as we have write permission on folder
10
    rm: remove write-protected regular file 'file1'? y
    $ echo "hello universe" > file1 # can create new file
11
    $ cat file1
12
    hello universe
13
```

However, this only works on files. You cannot remove a directory if you do not have **write** permission on the directory, even if you have **write** permission on its parent directory.

Can list names but not metadata?

If you have **read** permission on a directory but not **execute** permission, you cannot traverse into the directory, but you can still use ls to list the contents of the directory. However, you cannot use ls -l to long-list the contents of the directory. That is, you only have access to the name of the files inside, not their metadata.

```
$ mkdir test
    $ touch test/1 test/2
    $ chmod 600 test # removing execute permission from folder
3
    $ ls test # we can still list the files due to read permission
5
    $ ls -l test # but cannot long-list the files
    ls: cannot access 'test/2': Permission denied
    ls: cannot access 'test/1': Permission denied
8
9
    total 0
    -????????? ? ? ? ?
                                   ? 1
10
    -????????? ? ? ? ?
                                   ? 2
```

Cannot list names but can traverse?

If you have **execute** permission on a directory but not **read** permission, you can traverse into the directory but you cannot list the contents of the directory.

Subdirectories with all permissions, still cannot access?

If you have all the permissions to a directory, but dont have **execute** permission on its parent directory, you cannot access the subdirectory, or even list its contents.

```
$ mkdir test
1
   $ mkdir test/test2 # subdirectory
   $ touch test/test2/1 # file inside subdirectory
3
   $ chmod 700 test/test2 # all permissions to subdirectory
    $ chmod 600 test # removing execute permission from parent
      directory
   $ ls test
6
    test2
8
    $ cd test/test2 # cannot access subdirectory
    -bash: cd: test/test2: Permission denied
    $ ls test/test2 # cannot even list contents of subdirectory
10
    ls: cannot access 'test/test2': Permission denied
11
```

1.4.5 Changing Permissions

The permissions of a file can be changed using the **chmod** command.

Synopsis:

```
chmod [OPTION]... MODE[,MODE]... FILE...
chmod [OPTION]... OCTAL-MODE FILE...
```

OCTAL-MODE is a 3 or 4 digit octal number where the first digit is for the owner, the second digit is for the group, and the third digit is for others. We will discuss how the octal representation of permissions is calculated in the next section.

The **MODE** can be in the form of **ugoa+-=rwxXst** where:

- ▶ **u** is the user who owns the file
- ▶ **g** is the group of the file
- ▶ **o** is others
- ► **a** is all
- ► + adds the permission
- ▶ removes the permission
- ► = sets the permission
- ▶ r is read
- **w** is write
- ▶ x is execute
- ► **X** is execute only if its a directory or already has execute permission.
- ► **s** is setuid/setgid
- ▶ t is restricted deletion flag or sticky bit

We are already familiar with what \mathbf{r} , \mathbf{w} , and \mathbf{x} permissions mean, but what are the other permissions?

1.4.6 Special Permissions

Definition 1.4.5 (SetUID/SetGID) The setuid and setgid bits are special permissions that can be set on executable files. When an executable file has the setuid bit set, the file will be executed with the privileges of the owner of the file. When an executable file has the setgid bit set, the file will be executed with the privileges of the group of the files.

SetUID:

This is useful for programs that need to access system resources that are only available to the owner or group of the file.

A very notable example is the passwd command. This command is used to set the password of an user. Although changing password of a user is a priviledged action that only the root user can do, the passwd command can be run by any user to change *their* password. This is possible due to the setuid bit set on the passwd command. When the passwd command is run, it is run with the privileges of the root user, and thus can change the password of that user.

You can check this out by running ls -l /usr/bin/passwd and seeing the s in the permissions.

SetGID:

The behaviour of **SetGID** is similar to **SetUID**, but the file is executed with the privileges of the group of the file.

However, **SetGID** can also be applied to a directory. When a directory has the **SetGID** bit set, all the files and directories created inside that directory will inherit the group of the directory, not the group of the user who created the file or directory. This is highly useful when you have a directory where multiple users need to work on the same files and directories, but you want to restrict the access to only a certain group of users. The primary group of each user is different from each other, but since they are also part of another group (which is the group owner of the directory) they are able to read and write the files present in the directory. However, if the user creates a file in the directory, the file will be owned by the user's primary group, not the group of the directory. So other users would not be able to access the file. This is fixed by the **SetGID** bit on the directory.

```
$ mkdir test
2
    $ ls -ld test # initially the folder is owned by the user's
      primary group
    drwxr-xr-x 2 sayan sayan 4096 May 22 16:27 test
3
    $ chgrp wheel test # we change the group of the folder to wheel,
4
       which is a group that the user is part of
    $ ls -ld test
    drwxr-xr-x 2 sayan wheel 4096 May 22 16:27 test
    $ whoami # this is the current user
8
    savan
    $ groups # this is the users groups, first one is its primary
9
       group
10
    sayan wheel
    $ touch test/file1 # before setting the SetGID bit, a new file
11
      will have group owner as the primary group of the user
       creating it
12
    $ ls -l test/file1 # notice the group owner is sayan
13
    -rw-r--r-- 1 sayan sayan 0 May 22 16:29 test/file1
14
    $ chmod g+s test # we set the SetGID bit on the directory
    $ ls -ld test # now the folder has a s in the group permissions
15
    drwxr-sr-x 2 sayan wheel 4096 May 22 16:29 test
```

```
$ touch test/file2 # now if we create another new file, it will
have the group owner as the group of the directory

$ ls -l test/file2 # notice the group owner is wheel
-rw-r--r-- 1 sayan wheel 0 May 22 16:29 test/file2
```

Restricted Deletion Flag or Sticky Bit:

The restricted deletion flag or sticky bit is a special permission that can be set on directories.

Historically, this bit was to be applied on executable files to keep the program in memory after it has finished executing. This was done to speed up the execution of the program as the program would not have to be loaded into memory again. This was called **sticky bit** because the program would stick in memory. 40

However, this is no longer how this bit is used.

When the sticky bit is set on a directory, only the owner of the file, the owner of the directory, or the root user can delete or rename files in the directory.

This is useful when you have a directory where multiple users need to write files, but you want to restrict the deletion of files to only the owner of the file or the owner of the directory.

The most common example of this is the /tmp directory. The /tmp directory is a directory where temporary files are stored. You want to let any user create files in the /tmp directory, but you do not want any user to delete files created by other users.

```
s ls -ld /tmp
drwxrwxrwt 20 root root 600 May 22 16:43 /tmp
```

Exercise 1.4.1 Log into the system commands VM and cd into the /tmp directory. Create a file in the /tmp directory. Try to find if there are files created by other users in the /tmp directory using ls-l command. If there are files created by other users, try to delete them. a

1.4.7 Octal Representation of Permissions

The permissions of a file for the file's owner, group, and others can be represented as a 3 or 4 digit octal number. ⁴¹ Each of the octal digits is the sum of the permissions for the owner, group, and others.

- ▶ **Read** permission is represented by 4
- ► Write permission is represented by 2
- ▶ **Execute** permission is represented by 1

Thus if a file has read, write, and execute permissions for the owner, read and execute permissions for the group, and only read permission for others, the octal representation of the permissions would be 751.

The octal format is usually used more than the symbolic format as it is easier to understand and remember and it is more concise.

40: The part of memory where the program's text segment is stored is called the *swap*.

41: If the octal is 4 digits, the first digit is for special permissions like setuid, setgid, and sticky bit.

 $^{^{\}it a}$ You can create a file normally, or using the ${\bf mktemp}$ command.

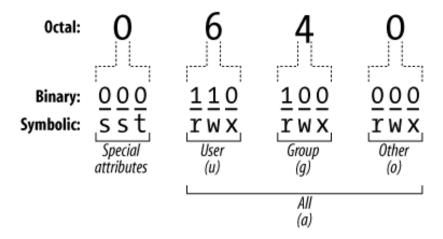


Figure 1.9: Octal Permissions

Table 1.8: Octal Representation of Permissions

Octal	Read	Write	Execute	Symbolic Representation	Description
0	0	0	0	_	No permissions
1	0	0	1	-x	Execute only
2	0	1	0	-W-	Write only
3	0	1	1	-wx	Write and execute
4	1	0	0	r–	Read only
5	1	0	1	r-x	Read and execute
6	1	1	0	rw-	Read and write
7	1	1	1	rwx	Read, write, and execute

^{1 \$} chmod 754 myscript.sh # this sets the permissions of myscript.sh to rwxr-xr--

However, if you want to add or remove a permission without changing the other permissions, the symbolic format is more useful.

```
s chmod u+x myscript.sh # this adds execute permission to the owner of myscript.sh
```

Question 1.4.1 How to list the permissions of a file?

Answer 1.4.1 ls -l

The permissions are the first 10 characters of the output. stat -c %A filename will list only the permissions of a file. There are other format specifiers of stat to show different statistics which can be found in man stat.

Question 1.4.2 How to change permissions of a file? Let's say we want to change file1's permissions to rwxr-xr- What is the octal form of that?

^{2 \\$./}myscript.sh

³ Hello World!

^{2 \$./}myscript.sh

³ Hello World!

Answer 1.4.2 chmod u=rwx, g=rx, o=r file1 will change the permissions of file1

The octal form of rwxr-xr- is 754.

So we can also use chmod 754 file1

Providing the octal is same as using = to set the permissions.

We can also use + to add permissions and - to remove permissions.

1.5 Types of Files

We had briefly seen that the output of ls -l shows the type of the file as the first character of the permissions.

There are 7 types of files in a linux file system as shown in Table Table 1.9.

1.5.1 Regular Files

Regular files are the most common type of file. Almost all files are regular files. Scripts and executable binaries are also regular files. All the configuration files of the system are regular files as well. The regular files are actually the only files that contain data and are stored on the disk.

1.5.2 Directories

Directories are files that contain a list of other files. Directories do not contain data, they contain references to other files. Usually the size of a directory is equal to the block size of the filesystem. Directories have some special permissions that are different from regular files as discussed in Section 1.4.

1.5.3 Symbolic Links

Symbolic links are files that point to other files. They only consume the space of the path they are pointing to. Symlinks ⁴² are useful to create shortcuts to files or directories. They are dependent on the original file and will stop working if the original file is deleted or moved. They are discussed in detail in Section 1.6.

1.5.4 Character Devices

Character devices are files that represent devices that are accessed as a stream of bytes. For example the keyboard, mouse, webcams, and most USB devices are character devices. These are not real files stored on the disk, but are files that represent devices. They can interacted with like a file using the read and write system calls to interact with the hardware directly. These files are made available by the kernel and are stored in the /dev directory. Any read/write operation on a character device is monitored by the kernel and the data is sent to the device.

```
$ cd /dev/input
2
  $ ls -l
3
  total 0
4
  drwxr-xr-x 2 root root
                             220 May 22 13:49 by-id
                             420 May 22 13:49 by-path
  drwxr-xr-x 2 root root
5
  crw-rw---- 1 root input 13, 64 May 22 13:49 event0
7
  crw-rw---- 1 root input 13, 65 May 22 13:49 event1
8 crw-rw---- 1 root input 13, 74 May 22 13:49 event10
9 crw-rw---- 1 root input 13, 75 May 22 13:49 event11
10 crw-rw---- 1 root input 13, 76 May 22 13:49 event12
11 crw-rw---- 1 root input 13, 77 May 22 13:49 event13
```

Table 1.9: Types of Files

Туре	Symbol
Regular Files	-
Directories	d
Symbolic Links	1
Character Devices	c
Block Devices	b
Named Pipes	p
Sockets	S

42: Symlinks is short for symbolic links.

There are another type of links called hard links. However, hard links are not files, they are pointers to the same inode. They do not consume extra space, and are not dependent on the original file. Hard links do not have a separate type, they are just regular files.

```
12 crw-rw---- 1 root input 13, 78 May 22 13:49 event14
13 crw-rw---- 1 root input 13, 79 May 22 13:49 event15
14 crw-rw---- 1 root input 13, 80 May 22 13:49 event16
15 crw-rw---- 1 root input 13, 81 May 22 13:49 event17
16 crw-rw---- 1 root input 13, 82 May 22 13:49 event18
17 crw-rw---- 1 root input 13, 83 May 22 13:49 event19
18 crw-rw---- 1 root input 13, 66 May 22 13:49 event2
19 crw-rw---- 1 root input 13, 84 May 22 13:49 event20
20 crw-rw---- 1 root input 13, 67 May 22 13:49 event3
21 crw-rw---- 1 root input 13, 68 May 22 13:49 event4
  crw-rw---- 1 root input 13, 69 May 22 13:49 event5
22
23 crw-rw---- 1 root input 13, 70 May 22 13:49 event6
24 crw-rw---- 1 root input 13, 71 May 22 13:49 event7
25 crw-rw---- 1 root input 13, 72 May 22 13:49 event8
26 crw-rw---- 1 root input 13, 73 May 22 13:49 event9
27 crw-rw---- 1 root input 13, 63 May 22 13:49 mice
28 crw-rw---- 1 root input 13, 32 May 22 13:49 mouse0
29 crw-rw---- 1 root input 13, 33 May 22 13:49 mouse1
```

Here the event and mouse files are character devices that represent input devices like the keyboard and mouse. Note the **c** in the permissions, which indicates that these are character devices.

1.5.5 Block Devices

Block devices are files that represent devices that are accessed as a block of data. For example hard drives, SSDs, and USB drives are block devices. These files also do not store actual data on the disk, but represent devices. Any block file can be mounted as a filesystem. We can interact with block devices using the read and write system calls to interact with the hardware directly. For example, the /dev/sda file represents the first hard drive in the system.

This makes it easy to write an image to a disk directly using the dd command. 43

The following example shows how we can use the dd command to write an image 44 to a USB drive. It is this easy to create a bootable USB drive for linux.

```
$ dd if=~/Downloads/archlinux.iso of=/dev/sdb bs=4M status=
progress
```

Here if is the input file, of is the output file, bs is the block size, and status is to show the progress of the operation.

Warning 1.5.1 Be very careful when using the dd command. Make sure you are writing to the correct disk. Writing to the wrong disk can cause data loss.

Example of how /sys works:

Try running the following code snippet in a terminal if you have a caps lock key on your keyboard and are running linux directly on your bare-metal. ⁴⁵

43: The dd command is a powerful tool that can be used to copy and convert files. It is acronym of *data duplicator*. However, it is also known as the *disk destroyer* command, as it can be used to overwrite the entire disk if you are not careful with which disk you are writing the image to.

44: ISO file

45: This will only work on your own system, not on the system commands VM, since you do not have the priviledge to modify the files there. Make sure you have the ability to run commands as root and are able to use sudo. It is also unlikely to work on a virtual machine. It will also not work on linux systems older than 2.6.

```
s cd /sys/class/leds
s echo 1 | sudo tee *capslock/brightness
```

If you are running a linux system directly on your hardware, you will see the caps lock key light up. Most modern keyboards will quickly turn off the light again as the capslock is technically not turned on, only the led was turned on manually by you.

1.5.6 Named Pipes

Named pipes ⁴⁶ are files that are used for inter-process communication. They do not store the data that you write to them, but instead pass the data to another process. A process can only write data to a named pipe if another process is reading from the named pipe. ⁴⁷

```
1  $ mkfifo pipe1
2  $ ls -l pipe1
3  prw-r--r-- -1 sayan sayan 0 May 22 18:22 pipe1
```

Here the p in the permissions indicates that this is a named pipe. If you now try to write to the named pipe, the command will hang until another process reads from the named pipe. Try the following in two different terminals:

Terminal 1:

```
1 | $ echo "hello" > pipe1
```

Terminal 2:

```
1 | $ cat pipe1
```

You will notice that whichever command you run first will hang until the other command is run.

1.5.7 Sockets

Sockets are a special file type, similar to TCP/IP sockets, providing inter-process networking protected by the file system's access control.

This is similar to named pipes, but the difference is that named pipes are meant for IPC between processes in the same machine, whereas sockets can be used for communication across machines.

Try out the following in two different terminals:

Terminal 1:

```
1 | $ nc -lU socket.sock
```

Terminal 2:

```
1 | $ echo "hello" | nc -U socket.sock
```

Notice here, that if you run the command in terminal 2 first, it will error out with the text:

```
1 nc: socket.sock: Connection refused
```

46: Also known as FIFOs

47: and vice versa

48: The nc command is the netcat command. It is a powerful tool for network debugging and exploration. It can be used to create sockets, listen on ports, and send and receive data over the network. This will be discussed in more detail in the networking section and in the Modern Application Development course.

Only if we run them in correct order can you see the message "hello" being printed in terminal 1. ⁴⁸

You can press Ctrl+C to stop the nc command in both terminals.

1.5.8 Types of Regular Files

Regular files can be further classified into different types based on the data they contain. In Linux systems, the type of a file is determined by its MIME type. The extension of a file does not determine its type, the contents of the file do. It is thus common to have files without extensions in Linux systems, as they provide no value.

The file command can be used to determine the type of a file.

identify the type of file are called the magic bytes. More details can be found at: 1 | \$ file /etc/passwd

The bytes at the start of a file used to

https://en.wikipedia.org/wiki/ List_of_file_signatures

```
2 /etc/passwd: ASCII text
3 | $ file /bin/bash
4 /bin/bash: ELF 64-bit LSB pie executable, x86-64, version 1 (SYSV)
      , dynamically linked, interpreter /lib64/ld-linux-x86-64.so
      .2, BuildID[sha1]=165d3a5ffe12a4f1a9b71c84f48d94d5e714d3db,
      for GNU/Linux 4.4.0, stripped
```

Question 1.5.1 What types of files are possible in a linux file system?

Answer 1.5.1 There are 7 types of files in a linux file system:

- ► Regular Files (starts with -)
- ► Directories (starts with d)
- ► Symbolic Links (starts with 1)
- ► Character Devices (starts with c)
- ► Block Devices (starts with b)
- ► Named Pipes (starts with p)
- ► Sockets (starts with s)

Question 1.5.2 How to know what kind of file a file is? Can we determine using its extension? Can we determine using its contents? What does MIME mean? How to get that?

Answer 1.5.2 The file command can be used to determine the type of a file.

The extension of a file does not determine its type.

The contents of a file can be used to determine its type.

MIME stands for Multipurpose Internet Mail Extensions.

It is a standard that indicates the nature and format of a document.

file -i filename will give the MIME type of filename.

1.6 Inodes and Links

Question 1.6.1 How to list the inodes of a file?

Answer 1.6.1 ls -i will list the inodes of a file. The inodes are the first column of the output of ls -i This can be combined with other flags like -l or -a to show more details.

Question 1.6.2 How to create soft link of a file?

Answer 1.6.2 In -s sourcefile targetfile will create a soft link of sourcefile named targetfile. The soft link is a pointer to the original file.

Definition 1.6.1 (Soft Links) Soft Links are special kinds of files that just store the path given to them. Thus the path given while making soft links should either be an absolute path, or relative **from** the location of the soft link **to** the location of the original file. It should not be relative from current working directory.^a

Question 1.6.3 How to create hard link of a file?

Answer 1.6.3 In sourcefile targetfile will create a hard link of sourcefile named targetfile. The hard link is same as the original file. It does not depend on the original file anymore after creation. They are equals, both are hardlinks of each other. There is no parent-child relationship. The other file can be deleted and the original file will still work.

Definition 1.6.2 (Hard Links) Hard Links are just pointers to the same inode. They are the same file. They are not pointers to the path of the file. They are pointers to the file itself. They are not affected by the deletion of the other file. When creating a hard link, you need to provide the path of the original file, and thus it has to be either absolute path, or relative from the current working directory, not relative from the location of the hard link.

Question 1.6.4 How to get the real path of a file?

Assume three files:

- ▶ file1 is a soft link to file2
- ► file2 is a soft link to file3
- ► file3 is a regular file

Real path of all these three should be the same. How to get that?

 $[\]overline{a}$ This is a common mistake.

Answer 1.6.4 realpath filename will give the real path of filename. You can also use readlink -f filename to get the real path.

1.7 Types of Commands

terminal is opened.

Question 1.7.1 How to create aliases? How to make them permanent? How to unset them?

Answer 1.7.1 alias name='command' will create an alias. unalias name will unset the alias. To make them permanent, add the alias to the \sim /.bashrc file. The \sim /.bashrc file is a script that is executed whenever a new

Question 1.7.2 How to run the normal version of a command if it is aliased?

Answer 1.7.2 \command will run the normal version of command if it is aliased.

Question 1.7.3 What is the difference between which, whatis, whereis, locate, and type?

Answer 1.7.3 Each of the commands serve a different purpose:

- ▶ which will show the path of the command that will be executed.
- ▶ whatis will show a short description of the command.
- ▶ whereis will show the location of the command, source files, and man pages.
- ▶ locate is used to find files by name.
- ► type will show how the command will be interpreted by the shell.

Exercise 1.7.1 Find the path of the true command using which. Find a short description of the true command using whatis. Is the executable you found actually the one that is executed when you run true? Check using type true

Definition 1.7.1 (Types of commands) A command can be an alias, a shell built-in, a shell function, keyword, or an executable. The type command will show which type the command is.

- ▶ alias: A command that is an alias to another command defined by the user or the system.
- ▶ builtin: A shell built-in command is a command that is built

- into the shell itself. It is executed internally by the shell.
- ► file: An executable file that is stored in the file system. It has to be stored somewhere in the PATH variable.
- ► **function**: A shell function is a set of commands that are executed when the function is called.
- ▶ **keyword**: A keyword is a reserved word that is part of the shell syntax. It is not a command, but a part of the shell syntax.

Shell Variables 2

2.1 Creating Variables

Definition 2.1.1 (Environment Variables) An environment variable is a variable that is accessbile to all processes running in the environment. It is a key-value pair. It is created using the export command. It can be accessed using the \$ followed by the name of the variable (e.g. \$HOME) or using the printenv command. They are part of the environment in which a process runs. For example, a running process can query the value of the TEMP environment variable to discover a suitable location to store temporary files, or the HOME or USER variable to find the directory structure owned by the user running the process.

Definition 2.1.2 (Shell Variables) A shell variable is a variable that is only accessible to the shell in which it is created. It is a key-value pair. It is created using the = operator. It can be accessed using the \$ followed by the name of the variable (e.g. \$var). They are local to the shell in which they are created.

Question 2.1.1 How to create a shell variable? How to create an environment variable?

Answer 2.1.1 var=value will create a shell variable. export var=value will create an environment variable.

Question 2.1.2 How to set the prompt string? How to set the secondary prompt string?

Answer 2.1.2 PS1="newprompt" will set the prompt string. PS2="newprompt" will set the secondary prompt string. This is useful when writing multi-line commands.

Question 2.1.3 Where are commands searched for? How to add a new path to search for commands? What symbol is used to separate the paths?

Answer 2.1.3 The **PATH** variable stores the paths to search for commands.

export PATH=\$PATH:newpath will add newpath to the **PATH** variable.

2.1 Creating Variables	•	45
2.2 Printing Variables to the		
Terminal		46
2.3 Bash Flags		47
2.4 Signals		47
2.5 Variable Manipulation .		48
2.6 Arrays		50

The colon: is used to separate the paths.

2.2 Printing Variables to the Terminal

Question 2.2.1 List the frequently used shell variables already existing in your shell.

Answer 2.2.1 Frequently used shell variables:

- ▶ USER stores the currently logged in user.
- ▶ **HOME** stores the home directory of the user.
- ▶ PWD stores the current working directory.
- ► **SHELL** stores the path of the shell being used.
- ▶ PATH stores the paths to search for commands.
- ▶ **PS1** stores the prompt string for the shell.
- ▶ PS2 stores the secondary prompt string for the shell
- ► HOSTNAME stores the network name of the system

Question 2.2.2 List the special shell variables useful in scripts.

Answer 2.2.2 Special shell variables useful in scripts:

- ▶ \$0 stores the name of the script.
- ▶ \$1, \$2, \$3, ... store the arguments to the script.
- ▶ \$# stores the number of arguments to the script.
- ▶ \$* stores all the arguments to the script as a single string.
- ▶ \$@ stores all the arguments to the script as array of strings.
- ▶ \$? stores the exit status of the last command.
- ▶ \$\$ stores the process id of the current shell.
- ▶ \$! stores the process id of the last background command.
- ▶ \$- stores the current options set for the shell.
- ▶ **\$IFS** stores the Internal Field Separator.
- ▶ \$LINENO stores the current line number of the script.
- ▶ \$RANDOM stores a random number.
- ▶ \$SECONDS stores the number of seconds the script has been running.

Question 2.2.3 How to print things to the terminal? How to print the value of a variable? How to print escape characters? How to supress the newline?

Answer 2.2.3 echo "Hello World" will print Hello World to the terminal.

echo \$var will print the value of var to the terminal.

echo -e "Hello\nWorld" will print Hello and World in two lines.

echo -n "Hello" will print Hello without a newline.

Question 2.2.4 How to print all the environment variables? Only some variables?

Answer 2.2.4 printenv will print all the environment variables. env will also print all the environment variables. printenv var1 var2 var3 will print only var1, var2, and var3.

Question 2.2.5 How to escape a character in a string? How to escape a variable in a string?

Answer 2.2.5 echo -e "Hello\nWorld" will print Hello and World in two lines. echo "Hello var" will print Hello and the value of var.

2.3 Bash Flags

Question 2.3.1 How to list the bash flags set in the shell? How to set or unset them? What does the default flags mean? (himBHs)

Answer 2.3.1 echo \$- will list the bash flags set in the shell.

set -o flag will set the flag.

set +o flag will unset the flag.

The default flags are:

h: locate and remember (hash) commands as they are looked up.

i: interactive shell

m: monitor the jobs and report changes

B: braceexpand - expand the expression in braces

H: histexpand - expand the history command

s: Read commands from the standard input.

2.4 Signals

Remark 2.4.1 If you press Ctrl+S in the terminal, some terminals might stop responding. You can resume it by pressing Ctrl+Q. This is because ixoff is set. You can unset it using stty -ixon. This is a common problem with some terminals, thus it can placed inside the /.bashrc file.

Other Ctrl+Key combinations:

- ▶ Ctrl+C: Interrupt the current process, this sends the SIGINT signal.
- ► Ctrl+D: End of input, this sends the EOF signal.
- ► Ctrl+L: Clear the terminal screen.
- ► Ctrl+Z: Suspend the current process, this sends the SIGTSTP signal.

- ► Ctrl+R: Search the history of commands using reverse-i-search.
- ► Ctrl+T: Swap the last two characters
- ► Ctrl+U: Cut the line before the cursor
- ► Ctrl+V: Insert the next character literally
- ► Ctrl+W: Cut the word before the cursor
- ► Ctrl+Y: Paste the last cut text

2.5 Variable Manipulation

Question 2.5.1 How to remove a variable that is set? How to check if a variable is present?

Answer 2.5.1 unset var will remove the variable var.

echo \$var will print if the variable var is present.

test -v var will check if the variable var is present.

test -z \$var+x will check if the variable var is present.

Question 2.5.2 How to print a default value if variable is absent? if present? how to set a default value if variable is absent?

Answer 2.5.2 echo \$var:-default will print default if var is absent

echo \$var:+default will print default if var is present. echo \$var:=default will set var to default if var is absent.

Remark 2.5.1 echo \$var:-default will print default if var is absent or empty.

echo \$var-default will print default if var is absent, but not if var is empty.

echo \$var:+default will print default if var is present and not empty.

echo \$var+default will print default if var is present, regardless of whether var is empty or not.

Question 2.5.3 How to print the number of characters in a variable? How about a substring of the variable?

Answer 2.5.3 echo \$#var will print the number of characters in var. echo \$var:0:5 will print the first 5 characters of var.

Remark 2.5.2 The second number in the substring is the length of the substring. Thus, echo \$var:5:3 will print the 3 characters starting

from the 5th character. If the length of substring is more than the length of the variable, it will print till the end. Thus, echo \$var:5:100 will print the characters from the 5th character till the end.

Definition 2.5.1 (Variable Match and Delete) We can match some part of the variable and delete it from left or right:

\${var%pattern} will delete the shortest match of pattern from the end of var.

\${var%pattern} will delete the longest match of pattern from the end of var.

\${var#pattern} will delete the shortest match of pattern from the start of var.

\${var##pattern} will delete the longest match of pattern from the start of var.

Question 2.5.4 If we have a variable with value "abc.def.ghi.xyz". How can we get the following:

- ▶ "abc.def.ghi"
- ▶ "abc"
- ► "def.ghi.xyz"
- ► "xyz"

Answer 2.5.4 The variable is var="abc.def.ghi.xyz".

- ► echo \${var%.*} will print "abc.def.ghi".
- ▶ echo \${var%.*} will print "abc".
- ► echo \${var#*.} will print "def.ghi.xyz".
- ► echo \${var##*.} will print "xy

Question 2.5.5 How to replace some part of a variable by something else? How to do for all occurences?

Answer 2.5.5 echo \${var/pattern/string} will replace the first occurence of pattern with string.

echo ${\rm squares} = {\rm square$

Question 2.5.6 How to make a variable lowercase? How to make it uppercase? How to do for entire string? How to do for first character? How to make a variable Sentence case? (First letter is capital, rest lower)

Answer 2.5.6 The following commands can be used to manipulate the case of a variable:

```
echo ${var,,} # will make the variable lowercase.
echo ${var^^} # will make the variable uppercase.
echo ${var,} # will make the first character lowercase.
```

```
echo ${var^} # will make the first character uppercase.
lower=${var,,} # will store the lowercase variable in lower.
echo ${lower^} # will make the first character uppercase. (
Sentence case)
```

Question 2.5.7 How to make a variable integer only? How to make it lowercase only? Uppercase only? Read only?

```
Answer 2.5.7 declare -i var will make the variable integer only. declare -l var will make the variable lowercase only. declare -u var will make the variable uppercase only. declare -r var will make the variable read only.
```

Question 2.5.8 How to remove those restrictions from a variable?

Answer 2.5.8 declare +i var will remove the integer only restriction.

declare +l var will remove the lowercase only restriction.

declare +u var will remove the uppercase only restriction. declare +r var will not work, as read only cannot be removed.

Question 2.5.9 How to store output of a command in a variable?

Answer 2.5.9 var=\$(command) will store the output of command in var.

2.6 Arrays

Question 2.6.1 How to create an array? How to access an element of an array? How to print all elements of an array?

```
Answer 2.6.1 arr=(1 2 3 4 5) will create an array. echo ${arr[2]} will access the 3rd element of the array. echo ${arr[@]} will print all elements of the array.
```

Question 2.6.2 How to get the number of elements in an array? How to get the indices of the array? How to add an element to the array?

Answer 2.6.2 echo ${\text{gen}}$ will print the number of elements in the array.

```
echo ${!arr[@]} will print the indices of the array. arr+=(6) will add the element 6 to the array.
```

Question 2.6.3 What are associative arrays? How to declare and use them?

Answer 2.6.3 Associative arrays are arrays with key-value pairs. declare -A arr will declare an associative array. arr[key]=value will set the value of key to value. echo \${arr[key]} will print the value of key.

Question 2.6.4 How to store output of a command in an array?

Answer 2.6.4 arr=(\$(command)) will store the output of command in arr.

3.1 What are processes?

3.1 What are processes? 53

Definition 3.1.1 (Process) A process is an instance of a program that is being executed. It contains the program code and its current activity. Depending on the operating system (OS), a process may be made up of multiple threads of execution that execute instructions concurrently. Several processes may be associated with the same program; for example, opening up several instances of the same program often means more than one process is being executed. Each process has its own 'process id' or **PID** to uniquely identify it.

Question 3.1.1 How to get a snapshot of all the processes running? What are the commonly used flags used with it?

Answer 3.1.1 The **ps** command is used to get a snapshot of all the processes running.

- ▶ ps will get a snapshot of all the processes running.
- ▶ ps -e will show all the processes.
- ▶ ps -f will show full format listing.
- ▶ ps -1 will show long format listing.
- ▶ ps -u will show user-oriented format listing.
- ▶ ps -x will show processes without controlling terminals.
- ▶ ps -a will show all processes with a terminal.
- ▶ ps -A will show all processes
- ▶ ps aux is a common command to see all processes
- ▶ ps –forest will show the processes in a tree form



Notation

The next list describes several symbols that will be later used within the body of the document.

- c Speed of light in a vacuum inertial frame
- h Planck constant

Greek Letters with Pronunciations

Character	Name	Character	Name
α	alpha <i>AL-fuh</i>	ν	nu NEW
β	beta BAY-tuh	ξ,Ξ	xi KSIGH
γ, Γ	gamma GAM-muh	O	omicron OM-uh-CRON
δ , Δ	delta DEL-tuh	π , Π	pi <i>PIE</i>
ϵ	epsilon EP-suh-lon	ρ	rho ROW
ζ	zeta ZAY-tuh	σ, Σ	sigma SIG-muh
η	eta AY-tuh	τ	tau TOW (as in cow)
θ, Θ	theta THAY-tuh	v, Υ	upsilon OOP-suh-LON
ι	iota eye-OH-tuh	ϕ , Φ	phi FEE, or FI (as in hi)
к	kappa KAP-uh	Χ	chi KI (as in hi)
λ , Λ	lambda <i>LAM-duh</i>	ψ , Ψ	psi SIGH, or PSIGH
μ	mu MEW	ω, Ω	omega oh-MAY-guh

Capitals shown are the ones that differ from Roman capitals.

Alphabetical Index

preface, v