

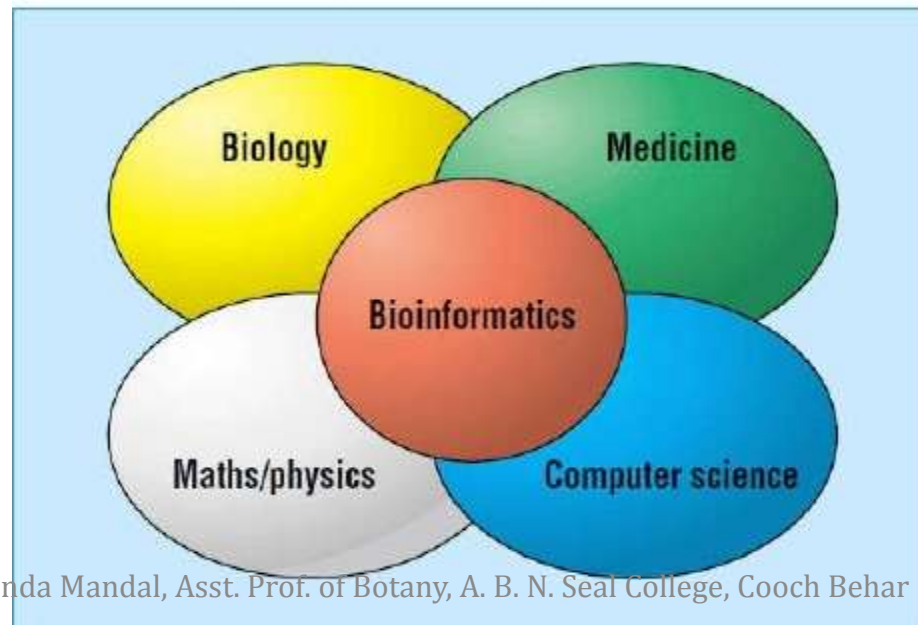
Bioinformatics

Dr. Aninda Mandal
Assistant Professor of Botany
A. B. N. Seal College
Cooch Behar 736101, WB

1. Introduction to the concept of hardware and software.
2. Introduction to Windows, UNIX and Linux; Introduction to Perl and Python.

Definition

- The marriage between computer science and molecular biology.
- Bioinformatics is an interdisciplinary field that combines biology, computer science, information engineering, mathematics and statistics to analyze and interpret biological data.
- The bioinformatics has been considered **“the mathematical, statistical and computing methods that aim to solve biological problems using DNA and amino acid sequences and related information.”**

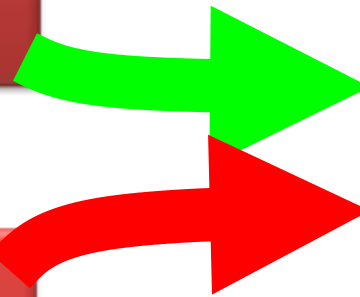


Biologists

collect molecular data:
DNA & Protein sequences,
gene expression, etc.

Computer scientists

(+Mathematicians,
Statisticians, etc.)
Develop tools, softwares,
algorithms
to store and analyze the
data.



Bioinformaticians

Study biological
questions by analyzing
molecular data

The Human Genome Project

- The Human Genome Project (HGP) was an international research effort to determine the sequence of the human genome and identify the genes that it contains. The Project was coordinated by the **National Institutes of Health and the U.S. Department of Energy**. Additional contributors included universities across the United States and international partners in the United Kingdom, France, Germany, Japan, and China. The Human Genome Project formally began in 1990 and was completed in 2003, 2 years ahead of its original schedule.
- The work of the Human Genome Project has allowed researchers to begin to understand the blueprint for building a person. As researchers learn more about the functions of genes and proteins, this knowledge will have a major impact in the fields of medicine, biotechnology, and the life sciences.

It is **important** because it uses information from DNA to develop new ways to treat, cure, or even prevent the thousands of diseases that afflict humankind.

Bioinformatics is playing an increasingly important role in the Human Genome Project, which aims to determine the complete DNA sequence of the human genome and identify all of the genes it contains.



- **Hardware**
- Hardware refers to the physical elements of a computer. This is also sometime called the machinery or the equipment of the computer. Examples of hardware in a computer are the **keyboard**, the **monitor**, the **mouse** and the **central processing unit**. However, most of a computer's hardware cannot be seen; in other words, it is not an external element of the computer, but rather an internal one, surrounded by the computer's casing (tower). A computer's hardware is comprised of many different parts, but perhaps the most important of these is the **motherboard**. The motherboard is made up of even more parts that power and control the computer.
- In contrast to software, *hardware is a physical entity*. Hardware and software are interconnected, without software, the hardware of a computer would have no function. However, without the creation of hardware to perform tasks directed by software via the central processing unit, software would be useless.
- **Hardware** is limited to specifically designed tasks that are, taken independently, very simple. **Software** implements *algorithms* (problem solutions) that allow the computer to complete much more complex tasks.

- **Software**
- Software, commonly known as programs or apps, consists of all the instructions that tell the hardware how to perform a task. These instructions come from a software developer in the form that will be accepted by the *platform* (operating system + CPU) that they are based on. For example, a program that is designed for the Windows operating system will only work for that specific operating system. Compatibility of software will vary as the design of the software and the operating system differ. Software that is designed for Windows XP may experience a compatibility issue when running under Windows 2000.
- Software is capable of performing many tasks, as opposed to hardware which can only perform mechanical tasks that they are designed for. Software provides the means for accomplishing many different tasks with the same basic hardware. Practical computer systems divide software systems into two major classes:
- **System software:** Helps run the computer hardware and computer system itself. System software includes operating systems, device drivers, diagnostic tools and more. System software is almost always pre-installed on your computer.
- **Application software:** Allows users to accomplish one or more tasks. It includes word processing, web browsing and almost any other task for which you might install software. (Some application software is pre-installed on most computer systems.)
- Software is generally created (written) in a high-level programming language, one that is (more or less) readable by people. These high-level instructions are converted into "machine language" instructions, represented in binary code, before the hardware can "run the code". When you install software, it is generally already in this machine language, binary, form.

Windows

- **MS Windows** is a **multitasking operating** system which uses a **Graphical User Interface (GUI)** to link the user to the computer.
- A **GUI** is an interface that helps users to interact with the computer by use of windows, icons, and menus. The interaction between the user and the computer heavily relies on the use of a mouse and keyboard which make it possible to enter data into the computer and also manipulate it the way we want.
- The **oldest of all Microsoft's operating systems is MS-DOS** (Microsoft Disk Operating System). MS-DOS is a text-based operating system. Users have to type commands rather than use the more friendly GUI's available today. Despite its very basic appearance, **MS-DOS is a very powerful operating system.** There are many advanced applications and games available for MS-DOS. A version of MS-DOS underpins Windows. Many advanced administration tasks in Windows can only be performed using MS-DOS.

Windows versions through the years

- **1985: Windows 1.0:** Windows 1.0 was essentially a program that ran on top of DOS.
- **1987: Windows 2.0 and 2.11:** Windows 2.0 was faster, more stable and had more GUI features. The system introduced the control panel and ran the first versions of Excel and Word.
- **1990: Windows 3.0:** Microsoft optimized the Windows 3.0 operating system, which still ran on top of DOS, for the 386 processor for a more responsive system. Windows 3.0 supported 16 colors and included the casual games familiar to most Windows users: Solitaire, Minesweeper and Hearts.
- Microsoft offered **Windows 3.1** as a paid sub-release in **1993**. Windows 3.1 features included support for TrueType fonts and peer-to-peer networking.
- **1993: Windows NT:** Windows NT's (New Technology) release marked the completion of a side project to build a new, advanced OS. NT was 32-bit and had a hardware abstraction layer. DOS was available through the command prompt, but it did not run the Windows OS. Microsoft designed NT as a workstation OS for businesses rather than home users. The system introduced the Start button.
- **1995: Windows 95:** Windows 95 introduced the Windows operating system to a wider audience with a marketing campaign that featured The Rolling Stones song "Start Me Up" to celebrate the Start button's arrival to the masses. Windows 95 facilitated hardware installation with its Plug and Play features.

- **1998: Windows 98:** Microsoft improved speed and Plug and Play hardware support in Windows 98. The company also debuted USB (Universal Serial Bus) support and the Quick Launch bar in this release.
- **2000: Windows ME:** Windows ME (Millennium Edition) was the last use of the Windows 95 codebase. Its most notable new feature was System Restore. Microsoft released the professional desktop **OS Windows 2000** the same year. Microsoft based this OS on the more stable Windows NT code.
- **2001: Windows XP:** Microsoft delivered Windows XP (eXPerience) as the first NT-based system with a version aimed squarely at the home user. XP offered the first Windows support for 64-bit computing, but it was not very well supported, lacking drivers and applications to run.
- **2006: Windows Vista:** Vista had interesting visual effects but the OS was slow to start and run.
- **2009: Windows 7:** Windows 7 picked up Vista's visual capabilities but featured more stability.
- **2012: Windows 8:** Microsoft released Windows 8 with a number of enhancements and debuted its tile-based Metro user interface. Windows 8 took better advantage of multicore processing, solid-state drives (SSD), touchscreens and other alternate input methods.
- **2015: Windows 10:** Microsoft announced Windows 10 in September 2014, skipping Windows 9. Version 10 includes the Start menu, which was absent from Windows 8. A responsive design feature called Continuum adapts the interface depending on whether the user works with a touchscreen or a keyboard and mouse for input.

UNIX

- UNIX is an operating system which was first developed in the 1960s, and has been under constant development ever since. The UNIX operating system is a set of programs that act as a link between the computer and the user.
- Types: There are many different versions of UNIX, although they share common similarities. The most popular varieties of UNIX are Sun Solaris, GNU/Linux, and MacOS X.
- The UNIX operating system is made up of three parts; the kernel, the shell and the programs.

Kernel – The kernel is the heart of the operating system. It interacts with the hardware and most of the tasks like memory management, task scheduling and file management.

Shell – The shell is the utility that processes your requests. When you type in a command at your terminal, the shell interprets the command and calls the program that you want. The shell uses standard syntax for all commands. **C Shell, Bourne Shell and Korn Shell** are the most famous shells which are available with most of the Unix variants.

What is Linux Operating system?

- Linux is an open source operating system based on UNIX, created in 1991. It is software which sits underneath of all other software on a computer. Users can modify the existing code and create distributions from it as it is an open source operating system. Linux operating system also comes with a graphical user interface (GUI) with some necessary software's which are used on a daily basis. Linux is mostly used as a server – as most of the web pages over the internet are generated from Linux servers and also used in desktop computers, mobile devices, gaming consoles, digital storing devices, eBook readers, cameras, video recorders have Linux running.

Key Differences Between Linux vs Windows

- Linux is open source operating system whereas Windows OS is commercial.
- Linux has access to source code and alters the code as per user need whereas Windows does not have access to source code.
- Linux will run faster than windows latest editions even with a modern desktop environment and features of the operating system whereas windows are slow on older hardware.
- Linux distributions don't collect user data whereas Windows collect all the user details which lead to privacy concern.
- Linux is more reliable then windows as in Linux we can kill application if they hung through *x kill* command whereas, in windows, we need to try multiple times to kill it.
- Linux supports a wide variety of free software's than windows but windows have a large collection of video game software.
- In Linux software cost is almost free as all programs, utilities, complex applications such as open office are free but windows also have many free programs and utilities but most of the programs are commercial.
- Linux is highly secure because it's easy to identify bugs and fix whereas Windows has a large user base and becomes a target for developers of viruses and malware.
- Linux is used by corporate organizations as servers and operating system for security purpose at *Google, Facebook, twitter* etc. whereas windows are mostly used by gamers and business users.
- Linux and windows have same priority over hardware and driver support in the present situation.

Comparison of Linux vs Windows

Basis for comparison	Linux	Windows
Access	In Linux user has access to the source code of kernel and alter the code according to his need. It has its own advantages like bugs in OS will fix at a rapid pace and disadvantages like developers may take advantage of any weakness in OS if they found.	In windows every user won't have access to the source code, only members of the selected group will have access to it.
Flavors or Variety	Linux has various distributions which are highly customizable based on user needs.	Windows has very few customization options available.
Licensing	In Linux with GPL- Licensed operating system, users are free to modify the software, can re-use in any number of systems and even they can sell the modified version.	In windows, with Microsoft license, users won't have access to source code (can't modify the software) and based on a number of licenses – we can install only on those number of computers.
Command line	In Linux, command line is a very useful tool for administration and daily tasks but for end users, it doesn't make much difference.	In windows, we have command line but can't use as Linux command line. We need to go running and enter cmd then command line will open.
Run level	Linux has inbuilt ability to stop at different run levels with this we can work using a command line and GUI if anyone has an issue.	In windows, if we encounter any problem in order to fix it, we need to reboot at run level 3 as an administrator/ root to find and fix the problem.
Usability	Linux is complicated to install but has the ability to complete complex tasks easier.	Windows gives user's a simple system to operate but it will take a longer time to install.
Support	Linux has support via a huge community of user forums/websites and online search.	Windows has support which is easily accessible, online forums/ websites and it has paid support also.
Updates	In Linux, users have full control of updates, we can install whenever we needed and it will take less time without any reboot.	In windows, updates will come at inconvenient times such as you are giving a print to the printer but suddenly update pop up will come which makes users frustrate and took more time to install.
Security	Linux is more secure than windows where hackers or developers of viruses will find difficult to break through Linux.	Windows is the major target for developers of viruses and malware and it is most vulnerable without anti-virus software.

Perl

- Perl is a general-purpose **programming language** originally developed for text manipulation and now used for a wide range of tasks including system administration, web development, network programming, etc.
- Though Perl is not officially an acronym, there are *various backronyms in use, including "Practical Extraction and Reporting Language"*. Perl was originally developed by **Larry Wall in 1987** as a general-purpose Unix scripting language to make report processing easier. Since then, it has undergone many changes and revisions. **Perl 6**, which began as a redesign of **Perl 5** in 2000, eventually evolved into a separate language. I development, and more.

Python

- **Python** is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991.
- Rossum was fan of a comedy series from late seventies. The name "Python" was adopted from the same series "Monty Python's Flying Circus".
- Sites like *Mozilla*, *Reddit*, *Instagram* and *PBS (Public Broadcasting Site)* are written in Python.
- Python 1.0 (first standard release): January 1994
- Python 3.5 (Last updated version): September 13, 2015

Both the languages have following features in common:

- Both were developed to target UNIX machine.
- Both are interpreted not compiled.
- Both are OO (Object oriented) but Python is ahead of Perl when we talk about OOPs.
- Both are open source in nature and you can modify it according to your need.
- Both have vast community support over the internet.

Key Differences Between Perl vs Python

- Perl, we can say “Practical Extraction and Report Language” but it is not official on the other hand Python is named after famous artist of his time “Monty Python”.
- Perl was invented by Larry Wall in 1987 while Python by Guido van Rossum in 1989.
- Perl is considered for text processing which helps in reporting while Python is now a general purpose language and currently being used in many fields.
- Python code is easy to use for beginners due to its intuitive design.
- Python is unlike other programming languages in that it uses indentation to identify blocks of code.
- Perl is a much older language with a wider range of modules available.

Thank You