

TIN HỌC CƠ SỞ

# ***Chương 1:***

## ***Tin học căn bản***



- Bài 1: Tổng quan
- Bài 2: Hệ điều hành Windows
- Bài 3: Mạng máy tính
- Bài 4: Một số chương trình ứng dụng

# TIN HỌC CƠ SỞ

## ***Bài 1: Tổng quan***

### Tin học căn bản

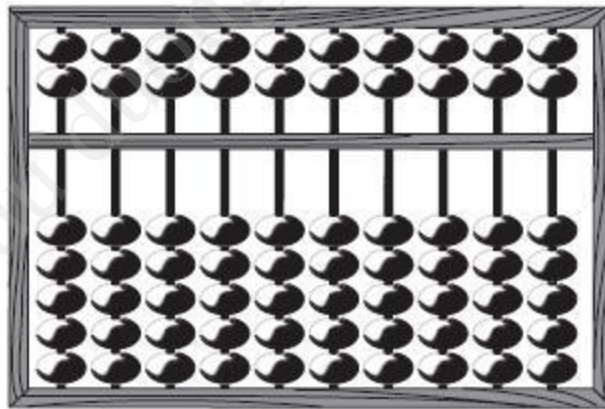


- Giới thiệu về máy tính
- Các thành phần của máy tính
- Hệ thống đếm
- Đơn vị lưu trữ thông tin
- Biểu diễn thông tin

- Giới thiệu về máy tính
  - Sự ra đời
  - Lịch sử phát triển
  - Đặc điểm và chức năng
  - Phân loại máy tính
- Các thành phần của máy tính
- Hệ thống đếm
- Đơn vị lưu trữ thông tin
- Biểu diễn thông tin

# Sự ra đời của máy tính

- Thiết bị tính toán cổ xưa nhất là bàn tính, có thể bắt nguồn từ Babylon vào khoảng 2400 năm trước công nguyên.
- Một phiên bản quen thuộc nhất hiện nay là bàn tính của người Trung Quốc.



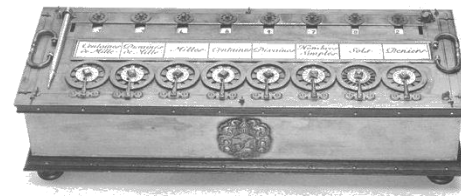
Bàn tính của người Trung Quốc

# Sự ra đời của máy tính

- 1641, Blaise Pascal (1623 – 1662) chế tạo máy cộng cơ học đầu tiên.
- 1671, Gottfried Leibnitz (1646 – 1716) cải tiến máy cộng cơ học của Pascal để thực hiện các phép tính cộng, trừ, nhân, chia đơn giản.



Blaise Pascal



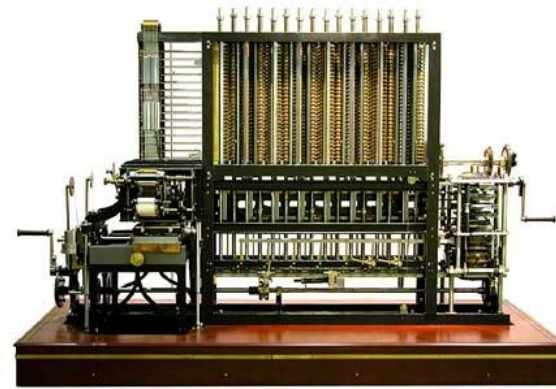
Máy cộng cơ học của Pascal

# Sự ra đời của máy tính

- 1833, Charles Babbage (1792 - 1871) cho rằng không nên phát triển máy cơ học và đề xuất máy tính với chương trình bên ngoài (thẻ đục lỗ).



Charles Babbage



Máy tính của Charles Babbage

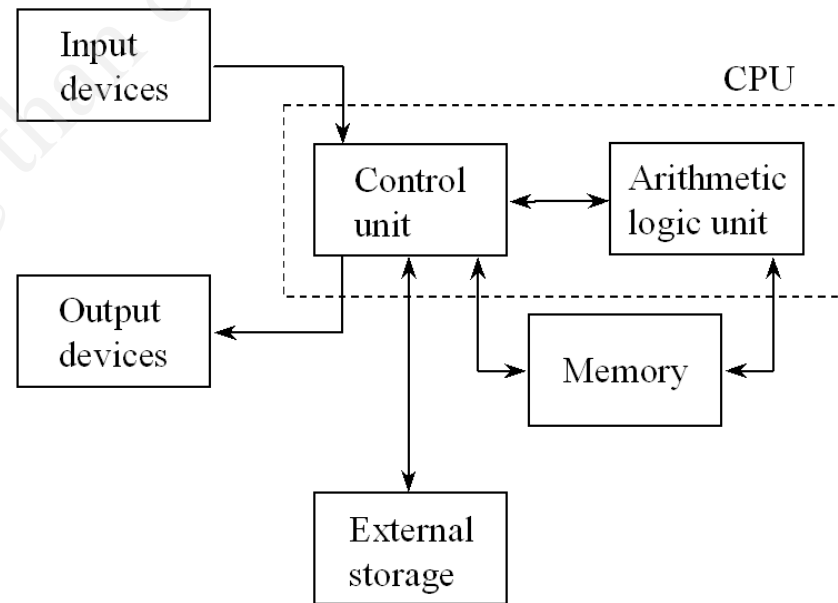


# Sự ra đời của máy tính

- 1945, John Von Neumann đưa ra nguyên lý có tính chất quyết định, đó là chương trình được lưu trữ trong máy và sự gián đoạn quá trình tuần tự.



John Von Neumann



Kiến trúc của J.V. Neumann

- Thế hệ thứ nhất (1945 – 1959)
  - Sử dụng đèn chân không (vacuum tube).
  - Tốc độ thấp:  $10^3$  phép tính/giây.
  - Chương trình viết bằng ngôn ngữ máy.
  - Máy ENIAC (Hoa Kỳ) dài 30.5m, nặng 30 tấn, 18000 bóng chân không, sử dụng thẻ đục lỗ, thực hiện 1900 phép cộng/giây, phục vụ cho mục đích quốc phòng (tính đạn đạo, chế tạo bom nguyên tử,...)
  - Máy UNIVAC nhanh hơn máy ENIAC 10 lần, sử dụng hơn 5000 bóng chân không.

- Thế hệ thứ hai (1960 – 1964)
  - Sử dụng đèn bán dẫn (nhỏ và rẻ hơn, tiêu thụ ít điện năng và tỏa nhiệt ít hơn bóng chân không).
  - Tốc độ nhanh:  $10^6$  phép tính/giây.
  - IBM 7090 đạt 2 triệu phép tính/giây, tham gia vào dự án Mercury (Hoa Kỳ) (đưa con người lên quỹ đạo trái đất), tìm ra số nguyên tố lớn nhất tại thời điểm đó (1961) với 1332 chữ số\*.
  - Máy M-3, Minsk -1, Minsk -2 (Liên Xô)
  - Chương trình viết bằng NNLT cấp cao COBOL, ALGOL.

*\* Đến tháng 10/2009, số nguyên tố tìm được có 12.978.189 chữ số*

- Thế hệ thứ ba (1964 – 1970)
  - Sử dụng mạch tích hợp IC (máy tính nhỏ hơn, tốc độ thực thi nhanh hơn, nhiệt lượng tỏa ra giảm, giá thành rẻ hơn...).
  - Tốc độ cao:  $10^9$  phép tính/giây.
  - Ngôn ngữ lập trình cấp cao & các phần mềm ứng dụng.
  - IBM360 (Mỹ) thực hiện 500.000 phép cộng/giây (gấp 250 lần máy ENIAC).

- Thế hệ thứ tư (1970 – Nay)
  - Mạch tích hợp quy mô lớn (LSI) và mạch tích hợp quy mô rất lớn (VLSI).
    - Intel 4004 năm 1971 (bộ vi xử lý 4 bit)
    - Intel 8008 năm 1972 (bộ vi xử lý 8 bit)
    - Intel 8086 năm 1978 (bộ vi xử lý 16 bit)
    - Intel Core i7 (1.170.000.000 bóng bán dẫn, 6 nhân, xử lý cùng lúc 12 luồng công việc)
  - Cơ chế xử lý song song.

- Thể hệ thứ năm (*tương lai gần?*)
  - Hoạt động trên trí thông minh nhân tạo.
  - Giao tiếp trực tiếp với con người bằng ngôn ngữ tự nhiên, có thể tự học các tri thức của thế giới xung quanh, có thể biểu đạt cảm xúc...



## ■ Đặc điểm

- Là công cụ dùng để lưu trữ và xử lý dữ liệu một cách tự động theo một chương trình được xác định trước mà không cần có sự can thiệp của con người.

## ■ Chức năng

- Lưu trữ mọi loại dữ liệu trong thế giới thực.
- Xử lý tính toán một cách nhanh chóng và chính xác.

- Siêu máy tính (Supercomputer)
  - Nhiều bộ vi xử lý ghép song song, tốc độ cực lớn.
  - Được thiết kế để xử lý các ứng dụng thời gian thực như dự báo thời tiết, mô phỏng vụ nổ hạt nhân...





- Máy tính cái (Mainframe)
  - Được thiết kế để xử lý đa nhiệm.
  - Kích thước vật lý lớn, thực hiện hàng tỉ phép tính/giây.
  - Hệ thống nhập xuất mạnh, tập trung vào các bài toán có lượng dữ liệu vô cùng lớn, ví dụ như số liệu giao dịch tài chính, kinh doanh bảo hiểm...



- Máy tính cỡ trung (Minicomputer)
  - Dòng máy tính nằm giữa dòng máy tính cỡ lớn và máy vi tính.
  - Hiệu suất xử lý cũng như qui mô các ứng dụng cũng nằm giữa hai dòng này.



## ■ Máy vi tính (Microcomputer)

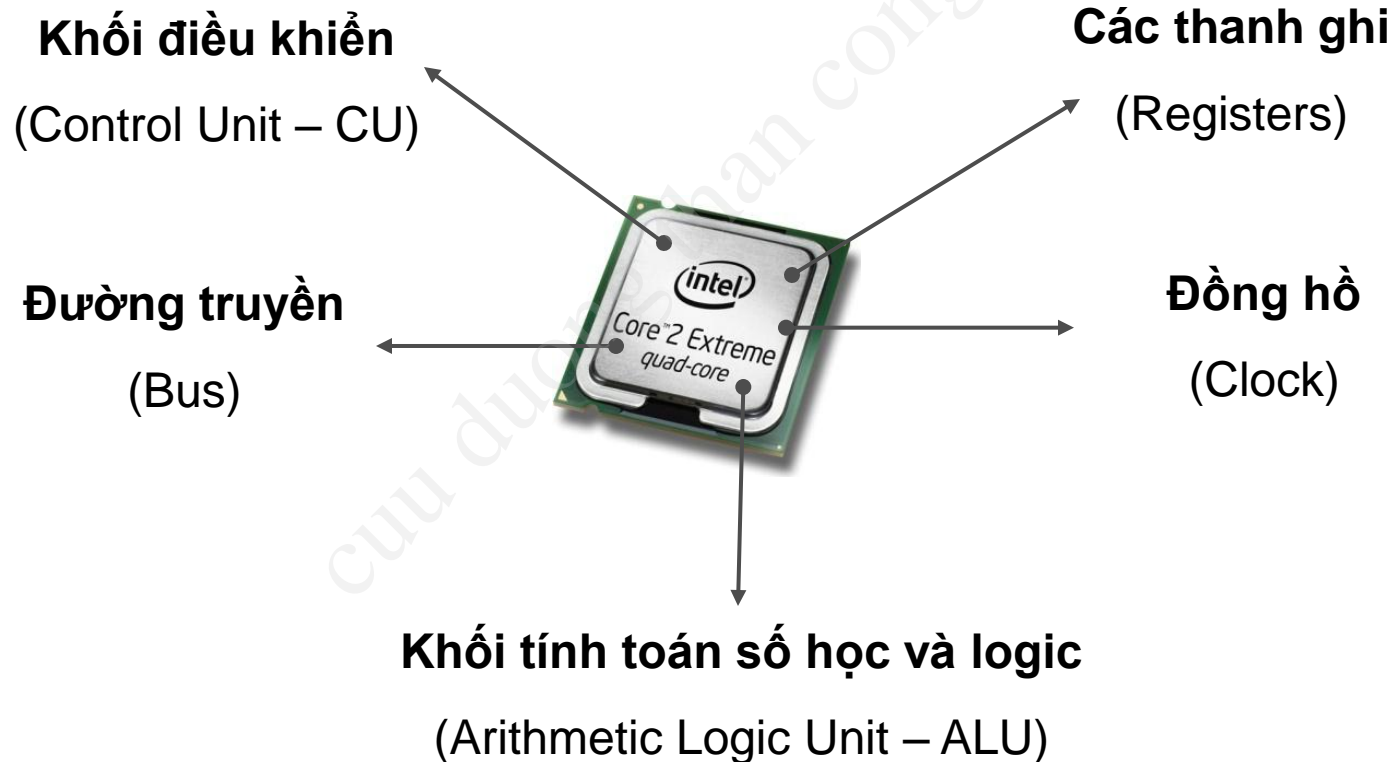
- Máy tính dùng cho cá nhân, gồm ba loại chính:
  - Máy tính để bàn (Desktop)
  - Máy tính xách tay (Laptop)
  - Máy tính cầm tay (Handheld)



- Giới thiệu về máy tính
- Các thành phần của máy tính
  - Phần cứng
  - Phần mềm
- Hệ thống đếm
- Đơn vị lưu trữ thông tin
- Biểu diễn thông tin

- Là các linh kiện / thiết bị vật lý cụ thể của máy tính, như màn hình, bàn phím, bộ nguồn, bộ vi xử lý, bảng mạch chủ...
- Các bộ phận chính:
  - Bộ vi xử lý (**C**entral **P**rocessing **U**nit - CPU)
  - Bộ nhớ trong (Memory)
  - Bộ nhớ ngoài (External Storage)
  - Bảng mạch chủ (Mainboard)
  - Các thiết bị nhập xuất (Input/Output Devices)

- Chỉ huy các hoạt động của máy tính.
- Gồm 3 khối chính:



- Thiết bị lưu trữ thông tin trong quá trình máy tính hoạt động.
  - ROM (Read Only Memory): Bộ nhớ chỉ đọc. ROM lưu chương trình hệ thống và dữ liệu này vẫn tồn tại khi nguồn điện cung cấp bị ngắt.
  - RAM (Random Access Memory): Bộ nhớ truy cập ngẫu nhiên. Dữ liệu lưu trên RAM chỉ là tạm thời, chúng sẽ mất đi khi nguồn điện cung cấp bị ngắt.



ROM



RAM

- Ưu điểm và khuyết điểm so với bộ nhớ trong:
  - Ưu điểm: khả năng lưu trữ lớn hơn rất nhiều, độ tin cậy cao và giá thành thấp.
  - Khuyết điểm: tốc độ truy xuất chậm hơn đáng kể nên chủ yếu dùng để chứa dữ liệu.
- Phân loại dựa trên đặc tính kỹ thuật:
  - Hệ thống từ tính
  - Hệ thống quang học
  - Bộ nhớ flash
  - Đĩa cứng thể rắn



## ■ Hệ thống từ tính:

- Băng từ (Tape): Phương thức lưu trữ ra đời đầu tiên, có tốc độ chậm, thường dùng để sao lưu dữ liệu.
- Đĩa mềm (Floppy Disk): Đường kính 5.25” (1.2MB) hoặc 3.5” (1.4MB), tốc độ chậm, tuổi thọ không cao.
- Đĩa cứng (Hard Disk): Nhiều lớp đĩa đồng tâm, dung lượng lên đến 3TB, tốc độ nhanh, tuổi thọ cao.



## ■ Hệ thống quang học

- Đĩa CD (Compact Disk): Kích thước 12 cm và 8 cm (loại nhỏ), dung lượng khoảng 700M.
- Đĩa DVD (Digital Video/Versatile Disk): Kích thước tương tự CD, dung lượng lên đến 17GB (2 mặt, 2 lớp).
- Một số cải tiến từ DVD:
  - HD DVD/Blu-ray (30/50GB)
  - HVD (500GB lên đến 3,9TB)
  - 5D DVD (10TB)



www.shutterstock.com · 1636782

- Hệ thống flash - Ổ USB flash (USB Flash Drive)
  - Kỹ thuật này được phát triển trong khoảng 10 năm gần đây, loại bỏ tính cơ học của đĩa từ và đĩa quang.
  - Kích thước nhỏ, giao tiếp thuận tiện thông qua cổng USB (Universal Serial Bus) nên sự xuất hiện của nó đã khiến cho đĩa mềm không còn lý do tồn tại.
  - Dung lượng thông dụng hiện nay trong khoảng 1 GB đến 16 GB.



- **Đĩa cứng thể rắn - SSD (Solid State Drive)**
  - Sử dụng bộ nhớ dạng rắn để lưu trữ dữ liệu.
  - Tốc độ đọc nhanh gấp 3 lần, tốc độ ghi nhanh gấp 1.5 lần ổ cứng bình thường.
  - Tiêu thụ ít điện năng, phù hợp với các thiết bị di động.
  - Giá thành cao hơn 10 lần so với ổ cứng thông thường.
  - Dung lượng lớn nhất năm 2010 là 1 TB có giá khoảng 2200\$.



- Bàn phím (keyboard): Thiết bị nhập chuẩn
  - Nhập dữ liệu.
  - Loại phổ biến có 104 phím, gồm 4 nhóm phím chính:
    - Nhóm phím đánh máy: phím chữ, phím số và phím các ký tự đặc biệt (~, !, @, #, \$, %, ^, &, ?, ...).
    - Nhóm phím chức năng: phím F1 đến F12, phím mũi tên, phím PageUp, PageDown, Insert, Delete, Home, End.
    - Nhóm phím số: NumLock, CapsLock, ScrollLock.
    - Nhóm phím điều khiển: Shift, Ctrl, Alt





- Chuột (Mouse)

Dùng để di chuyển con trỏ chuột trong môi trường đồ họa.



- Máy quét hình (Scanner)

Nhập văn bản hay hình vẽ, hình chụp vào máy tính.



- Camera & Webcam

Quay hình ảnh bên ngoài đưa vào máy tính.



- Bàn vẽ

Thiết bị gồm bảng điện tử và bút cảm ứng.

- Màn hình (Monitor): thiết bị xuất chuẩn
  - Gồm hai loại thông dụng là CRT, LCD.
  - Độ phân giải 800x600, 1024x768,...
  - Kích thước màn hình phổ biến hiện nay là 15", 17"...





- Máy chiếu (Projector)  
Kết xuất thông tin ra màn chiếu.



- Máy in (Printer)



- Loa (Speaker)



- Máy vẽ (Plotter)  
In đồ thị, đồ họa vec-tơ.

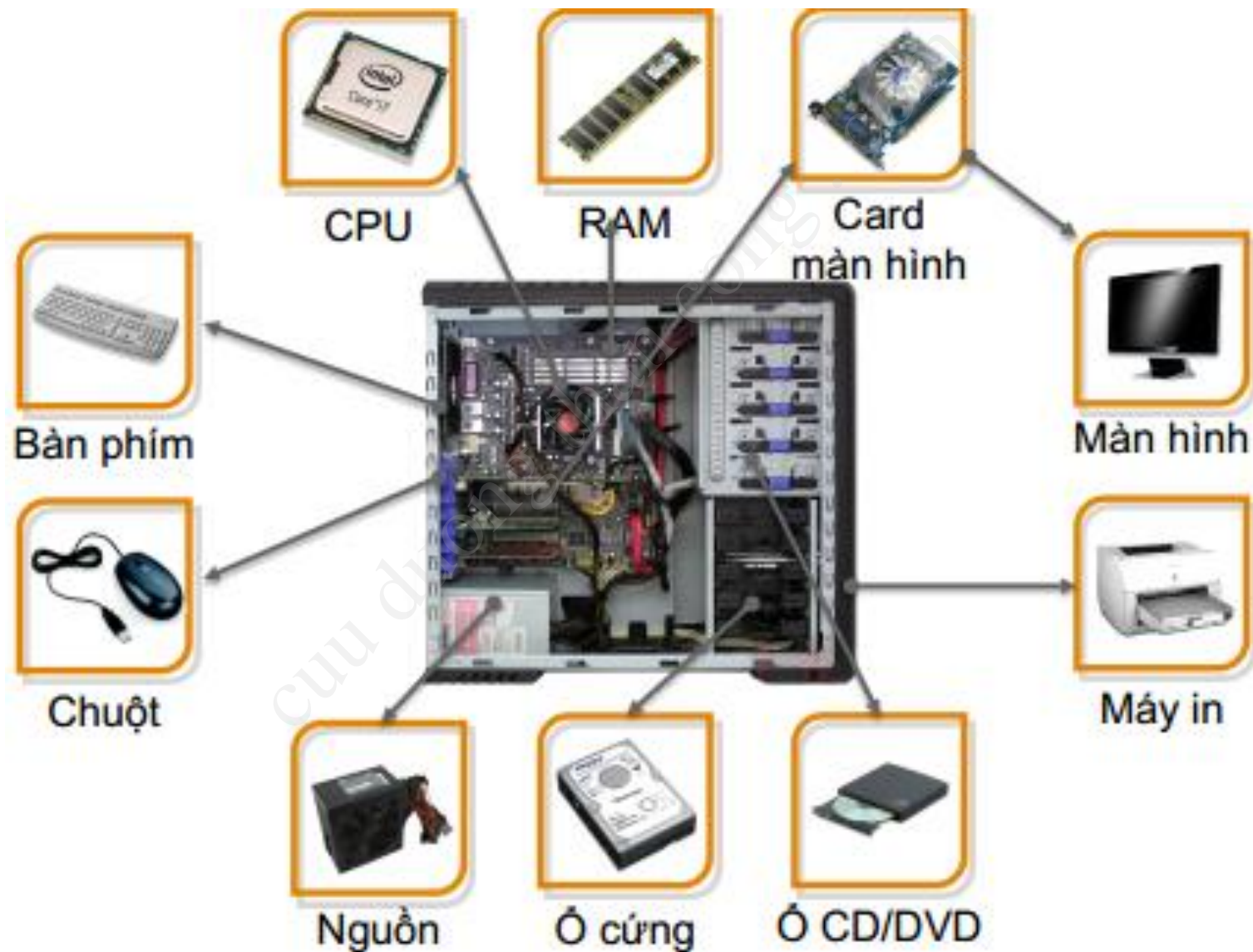


# Bảng mạch chủ

- Bảng mạch chủ (Mainboard) đóng vai trò quan trọng, là cầu nối cho các thành phần khác.
- Tất cả các thiết bị được gắn trên bảng mạch chủ như: nguồn máy tính, CPU, RAM, bảng mạch điều khiển (đồ họa, âm thanh, mạng), ổ đĩa cứng, đầu đọc đĩa (CD, đĩa mềm), cổng **kết nối màn hình**, bàn phím, chuột,...



# Thùng máy vi tính (Case)



- Phần mềm là một tập hợp những *câu lệnh* được viết bằng một hoặc nhiều *ngôn ngữ lập trình* theo một trật tự xác định nhằm tự động thực hiện một số nhiệm vụ hoặc chức năng, hoặc giải quyết một bài toán nào đó.
- Một số khái niệm liên quan
  - Phần mềm mã nguồn mở (open-source software)
  - Phần mềm miễn phí (freeware)
  - Phần mềm chia sẻ (shareware)
  - Bản quyền (copyright, ký hiệu © hay (C))

- Phần mềm hệ thống: Vận hành máy tính và các phần cứng máy tính.
  - Hệ điều hành (OS): MS-DOS, Linux, Windows...
  - Phần mềm quản trị cơ sở dữ liệu.
  - Phần mềm điều khiển thiết bị ngoại vi (driver).



HĐH Microsoft Windows



HĐH Fedora



HĐH MacOS

- Phần mềm hỗ trợ phát triển phần mềm
  - Trình biên dịch và thông dịch (Compiler, Interpreter).
  - Phần mềm gỡ rối (Debugger).
  - Phần mềm kết nối (Linkers, Loader).



Microsoft Visual Studio



Eclipse



Borland C++ Builder 6

## ■ Phần mềm ứng dụng

- Phần mềm hỗ trợ công việc: các ứng dụng văn phòng, thiết kế đồ họa...
- Giải trí: trò chơi, nghe nhạc, xem phim...
- Phần mềm tiện ích: diệt virus, nén dữ liệu...



Microsoft Office



World of Warcraft



Norton Antivirus



- Giới thiệu về máy tính
- Các thành phần cơ bản của máy tính
- **Hệ thống đếm**
  - Khái niệm
  - Một số hệ đếm
  - Chuyển đổi giữa các hệ đếm
- Đơn vị lưu trữ thông tin
- Biểu diễn thông tin

- Hệ đếm là hệ thống các quy tắc giúp xác định và thể hiện độ lớn của một giá trị.
- Ví dụ:
  - Hệ 10 (Decimal – DEC)
  - Hệ 2 (Binary – BIN)
  - Hệ 16 (Hexadecimal – HEX)



- Hệ đếm quen thuộc của con người.
- Sử dụng 10 ký hiệu từ 0 đến 9.
- Ví dụ:
  - $1208_{10} = 1 \times 10^3 + 2 \times 10^2 + 0 \times 10^1 + 8 \times 10^0$   
 $= 1 \times 1000 + 2 \times 100 + 0 \times 10 + 8 \times 1$   
 $= 1000 + 200 + 0 + 8 = 1208_{10}$
  - $1904_{10} = 1 \times 10^3 + 9 \times 10^2 + 0 \times 10^1 + 4 \times 10^0$   
 $= 1 \times 1000 + 9 \times 100 + 0 \times 10 + 4 \times 1$   
 $= 1000 + 900 + 0 + 4 = 1904_{10}$

- Hệ đếm sử dụng trong máy tính.
- Sử dụng hai ký hiệu là 0 và 1.
- Ví dụ:
  - $10001_2 = 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$   
 $= 1 \times 16 + 0 \times 8 + 0 \times 4 + 0 \times 2 + 1 \times 1$   
 $= 16 + 0 + 0 + 0 + 1 = 17_{10}$
  - $11101_2 = 1 \times 2^4 + 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$   
 $= 1 \times 16 + 1 \times 8 + 1 \times 4 + 0 \times 2 + 1 \times 1$   
 $= 16 + 8 + 4 + 0 + 1 = 29_{10}$
- Các phép toán: cộng, trừ, nhân, chia...

## ■ Phép cộng:

- Cộng có nhớ các cặp số cùng vị trí từ phải sang trái.
- Bảng cộng:

+	0	1
0	0	1
1	1	10

## ■ Ví dụ:

$$\begin{array}{rcccc} & 1 & 1 & 1 & 0 \\ + & 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 1 & 1 & 0 \end{array}$$

## ■ Phép trừ:

- Số bù 1: đảo tất cả các bit của 1 số nhị phân ta được số bù 1 của nó.
- Số bù 2: lấy số bù 1 cộng 1 ta được số bù 2 của số nhị phân ban đầu.
- Ví dụ:  $x = 1010$ 
  - Số bù 1 của  $x$ : 0101
  - Số bù 2 của  $x$ : 0110

## ■ Phép trừ:

➤ Cho 2 số nhị phân  $x$  và  $y$ , phép trừ:

$$x - y = x + \text{số bù 2 của } y$$

➤ Ví dụ:  $x=1010$ ,  $y=0101$

- Số bù 1 của  $y$ :  $1010$
- Số bù 2 của  $y$ :  $1011(y_2)$
- $x - y = x + y_2 = 1010 + 1011 = 0101$

## ■ Phép nhân:

➤ Nhân từ phải qua trái theo cách thông thường.

➤ Bảng nhân:

x	0	1
0	0	0
1	0	1

➤ Ví dụ:

$$\begin{array}{r}
 \begin{array}{cccc}
 & 1 & 0 & 1 & 1 \\
 x & & & 1 & 0 \\
 \hline
 & 0 & 0 & 0 & 0 \\
 1 & 0 & 1 & 1 & \\
 \hline
 1 & 0 & 1 & 1 & 0
 \end{array}
 \end{array}$$

## ■ Phép chia:

➤ Thực hiện tương tự như phép chia trong hệ cơ số 10.

➤ Ví dụ:

$$\begin{array}{r}
 \begin{array}{r}
 \text{—} \quad \mathbf{1 \ 1 \ 1 \ 0 \ 1} \mid \mathbf{1 \ 0 \ 1} \\
 \mathbf{1 \ 0 \ 1} \\
 \hline
 \mathbf{0 \ 1 \ 0 \ 0} \\
 \text{—} \quad \mathbf{0 \ 0 \ 0} \\
 \hline
 \mathbf{1 \ 0 \ 0 \ 1} \\
 \text{—} \quad \mathbf{1 \ 0 \ 1} \\
 \hline
 \mathbf{1 \ 0 \ 0} \leftarrow \text{Số dư}
 \end{array}
 \end{array}$$

- Sử dụng 16 ký hiệu từ 0=0000 đến 9 và A, B, C, D, E, F=1111.
- Ví dụ:
  - $4B8_{16} = 4 \times 16^2 + B \times 16^1 + 8 \times 16^0$   
 $= 4 \times 256 + 11 \times 16 + 8 \times 1$   
 $= 1024 + 176 + 8 = 1208_{10}$
  - $770_{16} = 7 \times 16^2 + 7 \times 16^1 + 0 \times 16^0$   
 $= 7 \times 256 + 7 \times 16 + 0 \times 1$   
 $= 1792 + 112 + 0 = 1904_{10}$



- Giá trị  $a \in \mathbb{N}^*$  có biểu diễn duy nhất dưới dạng:
  - $a = a_n b^n + a_{n-1} b^{n-1} + \dots + a_1 b^1 + a_0 b^0$   
hay thường được viết:  $a = (a_n a_{n-1} \dots a_1 a_0)_b$
  - Trong đó:
    - $b$  là cơ sở của biểu diễn,  $b \in \mathbb{N}$ ,  $b \geq 2$ .
    - $a_i$  là các ký số và  $a_i \in \mathbb{N}$ ,  $0 \leq i \leq n$ ,  $0 \leq a_i < b$ .
    - Cách viết trên được gọi là biểu diễn cơ sở  $b$  của  $a$ .
    - Chiều dài của biểu diễn bằng  $n + 1$ .

# Chuyển đổi giữa các hệ đếm

- Từ hệ cơ số 2 và 16 sang hệ cơ số 10

- Ví dụ 1: Đổi  $10001_2$  sang hệ cơ số 10

$$10001_2 = 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 0 \times 2^1 + 1 \times 2^0$$

$$10001_2 = 1 \times 16 + 0 \times 8 + 0 \times 4 + 0 \times 2 + 1 \times 1$$

$$10001_2 = 16 + 0 + 0 + 0 + 1 = 17_{10}$$

- Ví dụ 2: Đổi  $3203_{16}$  sang hệ cơ số 10

$$3203_{16} = 3 \times 16^3 + 2 \times 16^2 + 0 \times 16^1 + 3 \times 16^0$$

$$3203_{16} = 3 \times 4096 + 2 \times 256 + 0 \times 16 + 3 \times 1$$

$$3203_{16} = 12288 + 512 + 0 + 3 = 12803_{10}$$

# Chuyển đổi giữa các hệ đếm

## ■ Từ hệ cơ số 10 sang hệ cơ số 2 và 16

➤ Ví dụ 1: Đổi  $11_{10}$  sang hệ cơ số 2

$$11 : 2 = 5 \text{ dư } 1, \text{ vậy } a_0 = 1$$

$$05 : 2 = 2 \text{ dư } 1, \text{ vậy } a_1 = 1$$

$$02 : 2 = 1 \text{ dư } 0, \text{ vậy } a_2 = 0$$

$$01 : 2 = 0 \text{ dư } 1, \text{ vậy } a_3 = 1$$

$$\Rightarrow 11_{10} = 1011_2$$

➤ Ví dụ 2: Đổi  $1208_{10}$  sang hệ cơ số 16

$$1208 : 16 = 75 \text{ dư } 08, \text{ vậy } a_0 = 8$$

$$0075 : 16 = 04 \text{ dư } 11, \text{ vậy } a_1 = B$$

$$0004 : 16 = 00 \text{ dư } 04, \text{ vậy } a_2 = 4$$

$$\Rightarrow 1208_{10} = 4B8_{16}$$

- Từ hệ cơ số 2 sang hệ cơ số 16
  - Nhóm từng bộ 4 bit trong biểu diễn nhị phân rồi chuyển sang ký số tương ứng trong hệ thập lục phân (0000 → 0, ..., 1111 → F)
  - Ví dụ:  
 $1001011_2 = 0100\ 1011 = 4B_{16}$

HEX	BIN	HEX	BIN	HEX	BIN	HEX	BIN
0	0000	4	0100	8	1000	C	1100
1	0001	5	0101	9	1001	D	1101
2	0010	6	0110	A	1010	E	1110
3	0011	7	0111	B	1011	F	1111

# Chuyển đổi giữa các hệ đếm

- Từ hệ cơ số 16 sang hệ cơ số 2
  - Một ký số ở hệ cơ số 16 tương ứng với 4 ký số (bit) hệ cơ số 2.
  - Ví dụ:

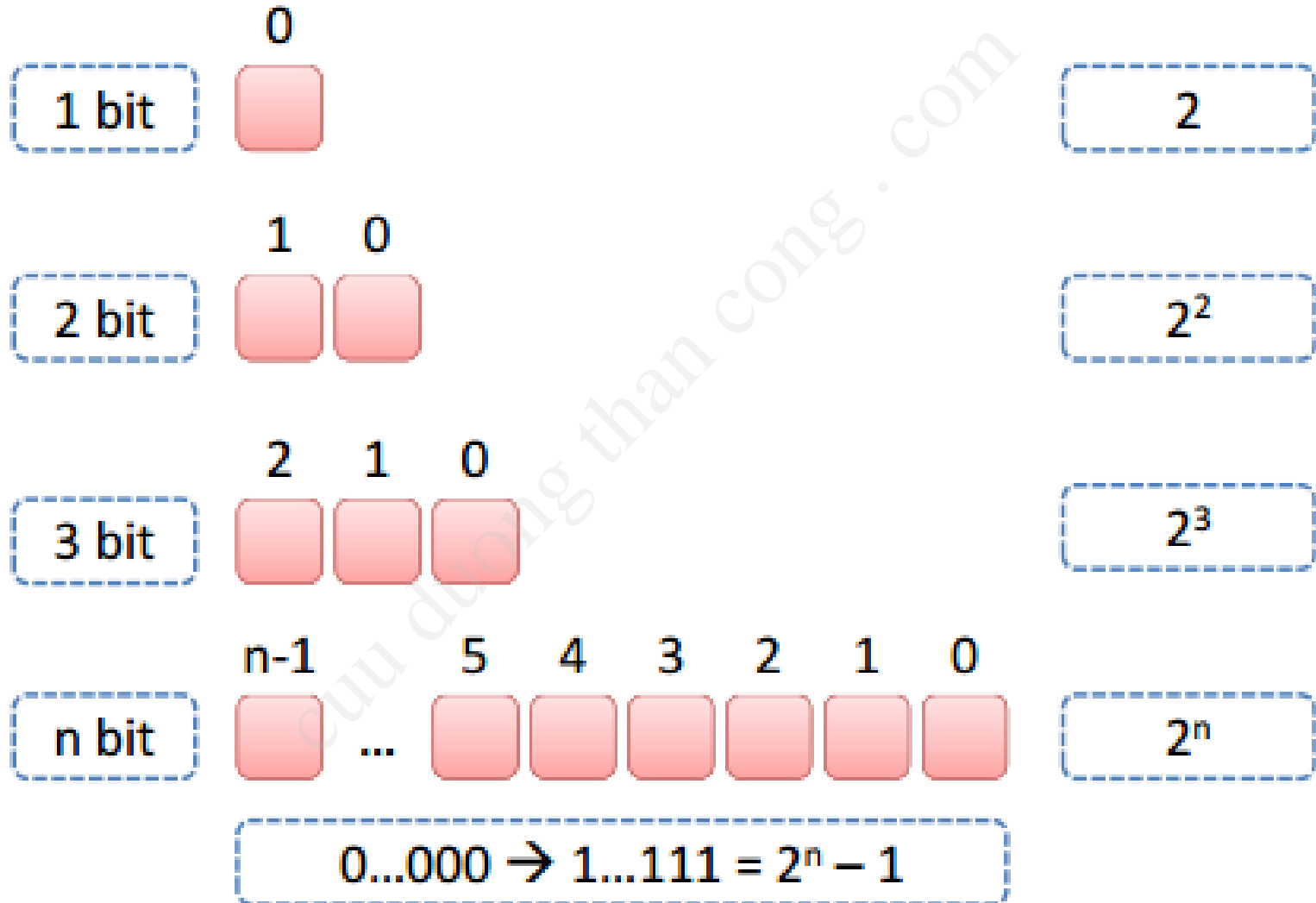
$$4B_{16} = 0100\ 1011 = 1001011_2$$

HEX	BIN	HEX	BIN	HEX	BIN	HEX	BIN
0	0000	4	0100	8	1000	C	1100
1	0001	5	0101	9	1001	D	1101
2	0010	6	0110	A	1010	E	1110
3	0011	7	0111	B	1011	F	1111

- Giới thiệu về máy tính
- Các thành phần của máy tính
- Hệ thống đếm
- Đơn vị lưu trữ thông tin
- Biểu diễn thông tin

- Máy tính chỉ “hiểu” một trong hai trạng thái: bật/tắt hay đúng/sai, được trừu tượng hóa bởi hai ký hiệu 1 và 0 (phù hợp với hệ đếm cơ số 2).
- Các ký hiệu 0 và 1 được gọi là các bit (binary digit), là đơn vị chứa thông tin nhỏ nhất.
- Byte là đơn vị đo cơ bản trong lĩnh vực máy tính, một byte gồm 8 bit.
- Các Bit được lưu trữ trong thanh ghi hoặc ô nhớ.
- Thanh ghi hoặc ô nhớ có kích thước 1 byte hoặc 1 word (16 bit).

# Đơn vị lưu trữ thông tin



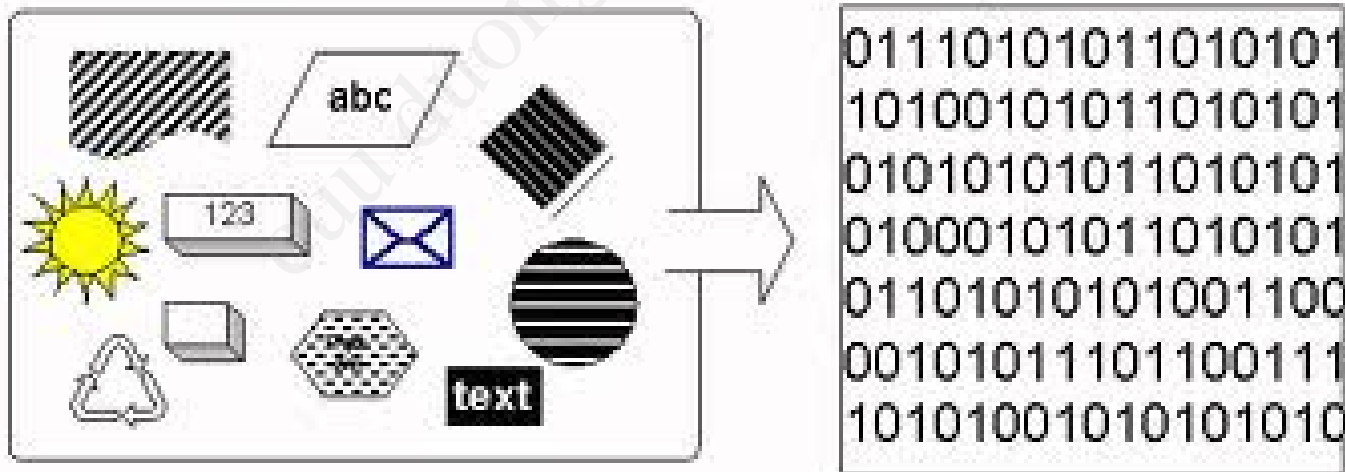


- Một số đơn vị đo:

Tên gọi	Ký hiệu	Biểu diễn
KiloByte	KB	$2^{10} \text{ B} = 1024 \text{ Byte}$
MegaByte	MB	$2^{10} \text{ KB} = 2^{20} \text{ Byte}$
GigaByte	GB	$2^{10} \text{ MB} = 2^{30} \text{ Byte}$
TeraByte	TB	$2^{10} \text{ GB} = 2^{40} \text{ Byte}$
Penta	PB	$2^{10} \text{ TB} = 2^{50} \text{ Byte}$

- Giới thiệu về máy tính
- Các thành phần của máy tính
- Hệ thống đếm
- Đơn vị lưu trữ thông tin
- **Biểu diễn thông tin**
  - Biểu diễn số nguyên
  - Biểu diễn số thực
  - Bảng mã
  - Tập tin

- Mọi dữ liệu lưu trữ trên máy tính đều là dữ liệu số - gồm các *bit*.
- Các dạng dữ liệu truyền thống như chữ viết, hình ảnh, âm thanh, video... phải được số hóa sang chuẩn dữ liệu mà máy tính nhận biết được.



- Biểu diễn số nguyên không dấu
  - Biểu diễn các đại lượng luôn dương.  
Ví dụ: chiều cao, cân nặng, số nguyên tố...
  - Số nguyên không dấu 1 byte lớn nhất là 1  
 $111\ 1111_2 = 2^8 - 1 = 255_{10}$ .
  - Số nguyên không dấu 1 word lớn nhất là  
 $1111\ 1111\ 1111\ 1111_2 = 2_{16} - 1 = 65535_{10}$
  - Tùy nhu cầu có thể sử dụng số 2, 3... word.

- Biểu diễn số nguyên có dấu

- Lưu các số dương hoặc âm.
- Bit cao nhất dùng để biểu diễn dấu.

Ví dụ: đối với số nguyên sử dụng 8 bit lưu trữ (gọi tắt là số nguyên 8 bits)

0 biểu diễn số dương. VD: 0101 0011

1 biểu diễn số âm. VD: 1101 0011

- Số âm được biểu diễn ở dạng số bù 2.

- Biểu diễn số nguyên có dấu (tt)

Số 5 (1 byte)

0	0	0	0	0	1	0	1
---	---	---	---	---	---	---	---

Số bù 1 của 5

1	1	1	1	1	0	1	0
---	---	---	---	---	---	---	---

+

1
---

Số bù 2 của 5

1	1	1	1	1	0	1	1
---	---	---	---	---	---	---	---

+ Số 5

0	0	0	0	0	1	0	1
---	---	---	---	---	---	---	---

Kết quả

1	0	0	0	0	0	0	0
---	---	---	---	---	---	---	---

- Biểu diễn số nguyên có dấu (tt)
  - Nhận xét
  - Số bù 2 của x cộng với x là một dãy toàn bit 0 (*không tính bit 1 cao nhất do vượt quá phạm vi lưu trữ*). Do đó số bù 2 của x chính là giá trị âm của x hay -x.
  - Đổi số thập phân -5 sang nhị phân?
    - => Đổi 5 sang nhị phân rồi lấy số bù 2 của nó.
  - Thực hiện phép toán  $a - b$ ?
    - $a - b = a + (-b) \Rightarrow$  Cộng với số bù 2 của b.

- Tính giá trị có dấu và không dấu của một số?

- Ví dụ: Số word (16 bit): 1100 1100 1111 0000

- Số nguyên không dấu?

Tất cả 16 bit lưu giá trị

=> giá trị là 52464.

- Số nguyên có dấu?

Bit cao nhất bằng 1 do đó số này là số âm

=> độ lớn là giá trị của số bù 2.

Số bù 2 = 0011 0011 0001 0000 = 13072.

=> giá trị là -13072.



- Sử dụng dấu chấm động (floating-point).

- Ví dụ số thực hệ 10:

$$-123.4d = -12.34 \times 10^1 = -1.234 \times 10^2 = -0.1234 \times 10^3$$

- Biểu diễn khoa học:  $-1.234 \times 10^2$

Chia làm 3 phần: *Một bit biểu diễn dấu; Một số bit biểu diễn số mũ; Một số bit biểu diễn phần định trị.*

- Ví dụ số:  $-1.234 \times 10^2$

- Bit biểu diễn dấu có giá trị 1 (giá trị âm).
- Các bit biểu diễn số mũ có giá trị 2.
- Các bit biểu diễn phần định trị có giá trị 1234.

*Quy ước bên trái dấu chấm là 1 ký số khác không.*

- Có nhiều chuẩn biểu diễn số dấu chấm động, chuẩn thông dụng nhất hiện nay là IEEE754.
  - IEEE754 đưa ra nhiều biểu diễn khác nhau, trong đó có 5 dạng cơ bản:
    - Số dấu chấm động độ chính xác đơn 32 bit.
    - Số dấu chấm động độ chính xác kép 64 bit.
    - Số dấu chấm động độ chính xác bậc bốn 128 bit.
    - Số dấu chấm động thập phân 64 bit.
    - Số dấu chấm động thập phân kép 128 bit.
  - Ví dụ: Số độ chính xác đơn gồm 1 bit dấu, 23 bit phần định trị, và 8 bit phần số mũ.

- Bảng mã là công cụ mã hoá các ký hiệu thế giới thực thành dữ liệu số, để lưu trữ và trao đổi thông tin trong lĩnh vực máy tính.
- Bảng mã thông dụng: ASCII và Unicode.
- Ví dụ trong bảng mã ASCII:

Ký tự A có mã 65

Ký tự B có mã 66

Ký tự a có mã 97

Dec	Hex	Char
64	40	@
65	41	A
66	42	B
67	43	C
68	44	D
69	45	E

Dec	Hex	Char
96	60	'
97	61	a
98	62	b
99	63	c
100	64	d
101	65	e

- **American Standard Code for Information Interchange.**
- ASCII được công bố làm tiêu chuẩn lần đầu vào năm 1963.
- Là bộ mã ký tự dựa trên bảng chữ cái tiếng Latinh (a – z, A – Z).
- Ban đầu, bảng mã ASCII chứa 128 mô tả cặp ký tự và số.

## ■ Các ký tự thông dụng

### ➤ Ký tự in được:

‘ ’ (khoảng trắng): 32 (0x20)

‘0’ -> ‘9’: 48 (0x30) -> 57 (0x39)

‘A’ -> ‘Z’: 65 (0x41) -> 90 (0x5A)

‘a’ -> ‘z’: 97 (0x61) -> 122 (0x7A)

### ➤ Ký tự điều khiển:

ký tự rỗng: 0

‘ ’ (tab): 9

ký tự xuống dòng: 10

ký tự về đầu dòng: 13

# **Bảng mã ASCII mở rộng**

- Bảng mã ASCII mở rộng chứa 256 mô tả cặp ký tự - số.
  - 128 ký tự đầu giống ASCII ban đầu.
  - 128 ký tự sau bao gồm một số ký hiệu tiếng Hy Lạp ('α', 'β', 'π'...), các biểu diễn tiền tệ ('£', '¥'...), các ký hiệu hình vẽ ('┐', '└', '┘', '┙'...).
  - Bảng mã ASCII không thể biểu diễn các ký tự ngôn ngữ khác như tiếng Việt (có dấu), Nga, Nhật...

- Unicode là bộ mã chuẩn quốc tế được thiết kế để dùng làm bộ mã duy nhất cho tất cả các ngôn ngữ khác nhau trên thế giới.
- Hiện tại, bảng mã unicode có 1.114.112 mã (code points), được chia thành 17 miền, mỗi miền có 65535 ( $2^{16}$ ) mã.

# **Bảng mã Unicode**

- Trong bảng mã unicode, có phần mã dành riêng tương ứng với những ký tự tiếng Việt.
- Bảng mã unicode các ký tự tiếng Việt:  
<http://vietunicode.sourceforge.net/charset/v3.htm>
- Mỗi bộ mã có nhiều cách hiển thị.
- Font thực hiện chức năng ánh xạ giá trị mã unicode với hình [ký tự] hiển thị tương ứng.
- Các font hỗ trợ unicode (có tiếng Việt) phổ biến gồm: Times New Roman, Arial, Tahoma...