

## Bài 2: Thông Tin Và Dữ Liệu

### A. Lý thuyết

#### 1. Khái niệm thông tin và dữ liệu

- Thông tin là những hiểu biết có thể có được về 1 thực thể nào đó
- Dữ liệu là thông tin đưa vào máy tính để xử lý

#### 2. Đơn vị đo lường thông tin

- Đơn vị cơ bản đo thông tin là bit (Binary Digital)
- Bit là đơn vị nhỏ nhất được lưu trữ trong máy tính để biểu diễn hai trạng thái 0 và 1 (0: không có điện; 1: có điện) ta còn thường gọi là mã nhị phân



Hình 1. Biểu diễn thông tin bằng dãy tám bit

- Ngoài đơn vị bit, ta cũng thường dùng đơn vị đo lường thông tin là Byte (đọc là bai)

- 1 byte = 8 bit

Một số đơn vị bội của Byte

Kí hiệu	Đọc	Độ lớn
Byte	Bai	8 bit
KB	Ki-lô-bai	1024 byte
MB	Mê-ga-bai	1024 KB
GB	Gi-ga-bai	1024 MB
TB	Tê-ra-bai	1024 GB
PB	Pê-ta-bai	1024 TB

Bảng 1. Một số đơn vị bội của Byte

### 3. Các dạng thông tin

Thông tin có 2 loại: số và phi số

- Số: Số nguyên, số thực,...

- Phi số: Văn bản, hình ảnh, âm thanh,...

+ Dạng văn bản: Tờ báo, cuốn sách, tấm bia,...

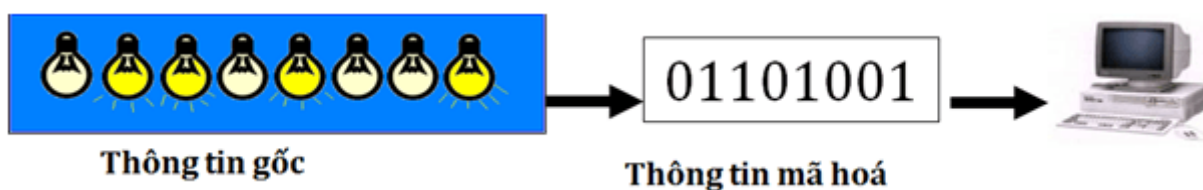
+ Dạng hình ảnh: Bức tranh vẽ, ảnh chụp, bản đồ, biển báo,...

+ Dạng âm thanh: Tiếng nói con người, tiếng sóng biển, tiếng đàn, tiếng chim hót,...

### 4. Mã hóa thông tin trong máy tính

- Để máy tính xử lý được, thông tin cần phải được biến đổi thành dãy bit (biểu diễn bằng các số 0, 1). Cách biến đổi như thế được gọi là mã hoá thông tin

Ví dụ:



Hình 2. Mã hóa thông tin trong máy tính

- Để mã hoá thông tin dạng văn bản ta dùng bộ mã ASCII để mã hoá các ký tự.

Mã ASCII các ký tự đánh số từ: 0 đến 255

- Bộ mã Unicode: có thể mã hóa  $65536 = 2^{16}$  ký tự, có thể mã hóa tất cả các bảng chữ cái trên thế giới

### 5. Biểu diễn thông tin trong máy tính

#### a. Thông tin loại số

Hệ đếm:

+ Hệ thập phân: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

+ Hệ nhị phân: 0, 1

+ Hệ cơ số mười sáu (hexa): 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F

Biểu diễn số trong các hệ đếm:

- Hệ thập phân: Mọi số N có thể biểu diễn dưới dạng

$$N = a_n 10^n + a_{n-1} 10^{n-1} + \dots + a_1 10^1 + a_0 10^0 + a_{-1} 10^{-1} + \dots + a_{-m} 10^{-m}$$

$$0 \leq a_i \leq 9$$

- Hệ nhị phân: Tương tự như trong hệ thập phân, mọi số N cũng có biểu diễn dạng

$$N = a_n 2^n + a_{n-1} 2^{n-1} + \dots + a_1 2^1 + a_0 2^0 + a_{-1} 2^{-1} + \dots + a_{-m} 2^{-m}$$

$$a_i = 0, 1$$

- Hệ hexa: Biểu diễn số trong hệ hexa cũng tương tự

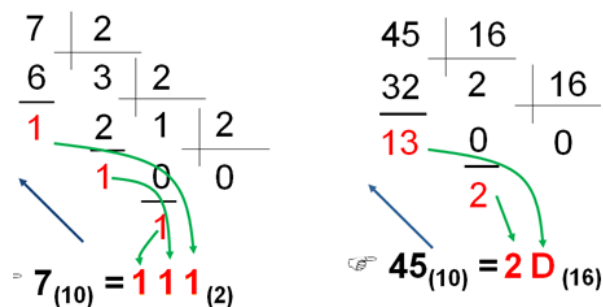
$$N = a_n 16^n + a_{n-1} 16^{n-1} + \dots + a_1 16^1 + a_0 16^0 + a_{-1} 16^{-1} + \dots + a_{-m} 16^{-m}$$

$$0 \leq a_i \leq 15$$

Với quy ước: A = 10, B = 11, C = 12, D = 13, E = 14, F = 15

Chuyển đổi giữa các hệ đếm:

Đổi số trong hệ cơ số 10 sang hệ cơ số 2 và hệ cơ số 16

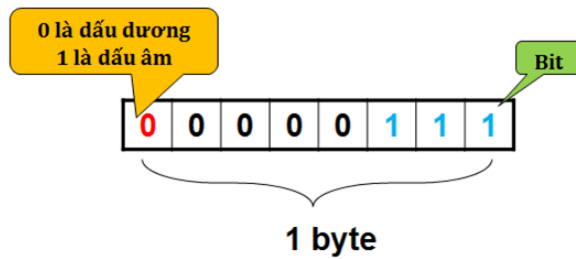


Hình 3. Ví dụ minh họa đổi số trong hệ cơ số 10 sang hệ cơ số 2 và hệ cơ số 16

Biểu diễn số trong máy tính:

-Biểu diễn số nguyên:

$$7_{(10)} = 111_{(2)}$$



Hình 4. Ví dụ minh họa biểu diễn số nguyên

Trong đó:

- + Phần nhỏ nhất của bộ nhớ lưu trữ số 0 hoặc 1: 1 bit
- + Một byte có 8 bit, bit cao nhất thể hiện dấu (bit dấu)
- + Có thể dùng 1 byte, 2 byte, 4 byte... để biểu diễn số nguyên

- Biểu diễn số thực:

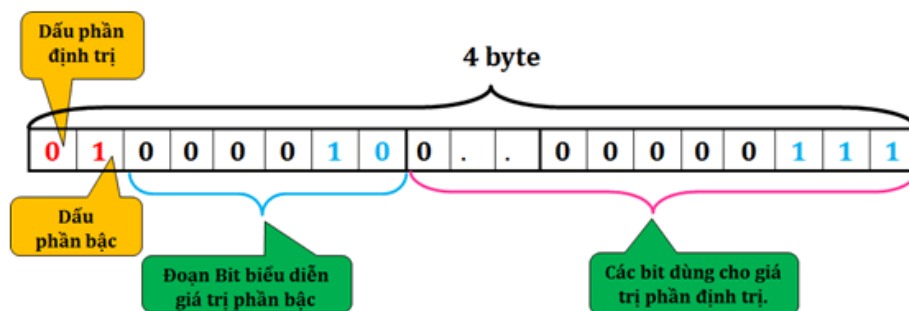
- + Biểu diễn số thực dưới dạng dấu phẩy động:
- + M: Là phần định trị ( $0,1 \leq M < 1$ )
- + K: Là phần bậc ( $K \leq 0$ )

Ví dụ:  $13456,25 = 0.1345625 \times 10^5$

Dạng tổng quát:  $\pm M \times 10^{\pm K}$

Trong đó:

- + Biểu diễn số thực trong một số máy tính:
- + Ví dụ:  $0,007 = 0.7 \times 10^{-2}$



Hình 5. Ví dụ minh họa biểu diễn số thực

## b. Thông tin loại phi số

Biểu diễn văn bản:

- Mã hoá thông tin dạng văn bản thông qua việc mã hóa từng kí tự và thường sử dụng:

- + Bộ mã ASCII: Dùng 8 bit để mã hoá kí tự, mã hoá được  $256 = 2^8$  kí tự
- + Bộ mã Unicode: Dùng 16 bit để mã hóa kí tự, mã hoá được  $65536 = 2^{16}$  kí tự

- Trong bảng mã ASCII mỗi kí tự được biểu diễn bằng 1 byte

Các dạng khác: Hình ảnh, âm thanh cũng phải mã hoá thành các dãy bit

\*Nguyên lí mã hóa nhị phân:

*Thông tin có nhiều dạng khác nhau như số, văn bản, hình ảnh, âm thanh, ... Khi đưa vào máy tính, chúng đều biến đổi thành dạng chung - dãy bit. Dãy bit đó là mã nhị phân của thông tin mà nó biểu diễn.*

## **B. Trắc nghiệm**

**Câu 1:** : Một quyển sách A gồm 200 trang nếu lưu trữ trên đĩa chiếm khoảng 5MB. Hỏi 1 đĩa cứng 40GB thì có thể chứa khoảng bao nhiêu cuốn sách có dung lượng thông tin xấp xỉ cuốn sách A?

- A. 8000
- B. 8129
- C. 8291

**D. 8192**

(**HD** : Đĩa cứng 40GB có thể lưu trữ số cuốn sách là:  $(40 \times 1024) : 5 = 8192$  (cuốn) )

**Câu 2:** Chọn câu đúng tron các câu sau:

- A. 1MB = 1024KB**
- B. 1B = 1024 Bit
- C. 1KB = 1024MB
- D. 1Bit = 1024B

**Câu 3:** Thông tin khi đưa vào máy tính, chúng đều được biến đổi thành dạng chung đó là:

- A. Hình ảnh
- B. Văn bản
- C. Dãy bit**
- D. Âm thanh

**Câu 4:** Phát biểu nào sau đây là phù hợp nhất về khái niệm bit?

A. Đơn vị đo khối lượng kiến thức

B. Chính chữ số 1

**C. Đơn vị đo lượng thông tin**

D. Một số có 1 chữ số

**Câu 5:** Tại sao phải mã hoá thông tin?

A. Để thay đổi lượng thông tin

B. Làm cho thông tin phù hợp với dữ liệu trong máy

C. Để chuyển thông tin về dạng câu lệnh của ngôn ngữ máy

**D. Tất cả đều đúng**

**Câu 6:** Đơn vị đo lượng thông tin cơ sở là:

A. Byte

**B. Bit**

C. GB

D. GHz

**Câu 7:** Mã hoá thông tin là quá trình:

**A. Đưa thông tin vào máy tính**

B. Chuyển thông tin về bit nhị phân

C. Nhận dạng thông tin

D. Chuyển dãy hệ nhị phân về hệ đếm khác

**Câu 8:** Chọn phát biểu đúng trong các câu sau:

**A. Một byte có 8 bits**

B. RAM là bộ nhớ ngoài

C. Dữ liệu là thông tin

D. Đĩa mềm là bộ nhớ trong

**(HD :**

- Bộ nhớ trong gồm RAM và ROM → loại B

- Bộ nhớ ngoài gồm: đĩa cứng, đĩa mềm, đĩa CD, thiết bị nhớ flash... → loại

D

- Dữ liệu là thông tin được mã hóa trong máy tính → loại C

)

**Câu 9:** 1 byte có thể biểu diễn ở bao nhiêu trạng thái khác nhau:

A. 8

B. 255

**C. 256**

D. 65536

(**HD** : 1 byte= 8 bit. Vậy 1 byte có thể biểu diễn các trạng thái khác nhau là  $2^8=256$  trạng thái)

**Câu 10:** Thông tin là gì?

A. Các văn bản và số liệu

**B. Hiểu biết của con người về một thực thể, sự vật, khái niệm, hiện tượng nào đó**

C. Văn bản, Hình ảnh, âm thanh

D. Hình ảnh, âm thanh