

## Bài 5.

# DỮ LIỆU TRONG MÁY TÍNH

*Học xong bài này, em sẽ:*

- Biết máy tính dùng dây bit biểu diễn các số trong tính toán.
- Biết được trong máy tính có những loại dữ liệu gì.
- Nêu được các bước cơ bản trong xử lý thông tin của máy tính.
- Biết quy đổi được gần đúng các đơn vị đo lường dữ liệu.

### 1. Biểu diễn số để tính toán trong máy tính



Trong hệ thập phân, cùng là chữ số “1” nhưng giá trị của nó khi ở hàng trăm gấp mười lần giá trị của nó ở hàng chục. Tức là nếu chữ số “1” dịch sang trái một vị trí thì nó biểu diễn giá trị mới gấp mười lần so với khi ở vị trí cũ (khi chưa dịch sang trái một vị trí). Bạn Minh Khuê nhận xét: Quy luật này chỉ đúng với chữ số “1”. Em có đồng ý với bạn Minh Khuê không?”

Kết quả đo đếm sự vật trong thế giới quanh ta là thông tin định lượng. Thông tin định lượng biểu diễn bằng các con số. Con người dùng 10 kí hiệu khác nhau “0”, “1”, “2”, “3”, “4”, “5”, “6”, “7”, “8”, “9” mà ta quen gọi là các “chữ số” để biểu diễn các số. Ta học đếm và tính toán số học theo các quy tắc của hệ thập phân.

Ban đầu con người sáng tạo ra máy tính chính là để tính toán cho nhanh. Máy tính làm việc bằng dòng điện nên dễ thể hiện hai trạng thái dòng điện khác nhau. Do vậy, các nhà tin học đã tìm cách chỉ dùng hai kí hiệu “0” và “1” (trùng ứng với hai trạng thái dòng điện) để biểu diễn số trong máy tính. Cách làm cũng tương tự như trong hệ thập phân, chỉ khác là nếu chữ số “1” dịch sang trái một vị trí thì nó biểu diễn giá trị gấp **hai** lần so với khi ở vị trí cũ, chứ không phải gấp **mười** lần như trong hệ thập phân.

1	8	3
↑	↑	↑
× 100	× 10	× 1

Biểu diễn số 183 trong hệ thập phân

Do quy ước của hệ thập phân, số 183 có giá trị là một trăm cộng tám chục cộng ba đơn vị:

$$183 = 1 \times 100 + 8 \times 10 + 3 \times 1$$

1	1	0
↑	↑	↑
× 4	× 2	× 1

Biểu diễn số 6 chỉ với hai kí hiệu “0” và “1”

Quy ước dịch sang trái một vị trí thì giá trị gấp **hai** lần so với khi ở vị trí cũ:

$$110 \rightarrow 1 \times 4 + 1 \times 2 + 0 \times 1 = 6$$

Số tạo thành từ cách biểu diễn chỉ dùng hai kí hiệu “0” và “1” như vậy được gọi là *Số nhị phân*.

Máy tính dùng dãy bit để biểu diễn các số trong tính toán.

## 2. Dữ liệu và các bước xử lí thông tin trong máy tính

Trong máy tính có dữ liệu văn bản, dữ liệu âm thanh, dữ liệu hình ảnh. Số nhị phân là một loại dữ liệu khác trong máy tính. Số nhị phân biểu diễn thông tin số lượng và được dùng trong các tính toán số học.

Mọi dữ liệu trong máy tính đều là dãy bit (bit kí hiệu là “b”). Với máy tính, thông tin và dữ liệu số là một, đều chỉ là các dãy bit.

Chu trình xử lí thông tin của máy tính bao gồm các bước:

1) Xử lí đầu vào: đầu vào được chuyển thành dữ liệu mà máy tính “hiểu được”, tức là dữ liệu số.

2) Xử lí dữ liệu: các phần mềm ứng dụng xử lí dữ liệu phục vụ mục đích của người dùng máy tính. Vì mọi dữ liệu đều là dãy bit nên mọi thao tác xử lí thông tin trong máy tính đều là thao tác với các bit.

3) Xử lí đầu ra: từ dãy bit xuất ra thông tin dưới dạng con người hiểu được hoặc ghi lưu dữ liệu vào thiết bị lưu trữ hay gửi lên mạng.

## 3. Dung lượng lưu trữ dữ liệu của một số thiết bị thường gặp

Vì bit quá nhỏ nên người ta dùng *byte* (đọc là “bai”) làm đơn vị đo lường dữ liệu, kí hiệu là B.

byte: một dãy 8 bit liên nhau.

Số byte trong một tệp dữ liệu thường khá lớn nên các bội số của byte hay được dùng hơn.

Các bội số của byte dùng để đo lường dữ liệu được tạo ra bằng cách nhân thêm  $2^{10}$  (bằng 1024) lần. Để dễ hình dung có thể xấp xỉ là nhân với 1000 lần. Ta có:

Viết là	Đọc là	Xấp xỉ
KB (Kilobyte)	Ki-lô-bai	Một nghìn byte
MB (Megabyte)	Mê-ga-bai	Một triệu byte
GB (Gigabyte)	Gi-ga-bai	Một tỉ byte
TB (Terabyte)	Tê-ra-bai	Một nghìn tỉ byte

Bảng 1. Các bội số của byte

Thẻ nhớ được dùng phổ biến cho điện thoại thông minh, máy ảnh số. Dung lượng lưu trữ của thẻ nhớ rất đa dạng. Ta thường thấy thẻ nhớ có dung lượng từ 1 GB đến 8 GB, 16 GB, 32 GB, 64 GB, 128 GB.

■ *Dung lượng lưu trữ:* khả năng lưu trữ của các thiết bị nhớ.

USB được dùng rất phổ biến để lưu trữ, trao đổi thông tin nhờ các loại máy tính đều có sẵn cổng cắm. Dung lượng lưu trữ của USB tương đương như thẻ nhớ, hoặc lớn hơn đến 256 GB, 512 GB. Thậm chí còn có các USB có dung lượng đến 1 TB, 2 TB.

Đĩa CD thường có dung lượng 700 MB. Đĩa DVD có thể lưu trữ từ 5 GB đến 17 GB tùy thuộc từng loại.

Điện thoại thông minh có thẻ nhớ trong 16 GB, 32 GB, 64 GB,...

Máy tính có ổ đĩa cứng với dung lượng vài trăm GB đến vài TB.



Trong các câu sau, câu nào đúng, câu nào sai? Giải thích tại sao.

- 1) Một MB xấp xỉ một nghìn byte.
- 2) Một TB xấp xỉ một triệu KB.
- 3) Một GB xấp xỉ một tỉ byte.
- 4) Một KB xấp xỉ một nghìn GB.



USB, thẻ nhớ được dùng phổ biến cho máy tính, điện thoại thông minh, máy ảnh số có nhiều mức dung lượng 8 GB, 16 GB, 32 GB, 64 GB, 128 GB,.... Em nên chọn dung lượng bao nhiêu là thích hợp cho mỗi trường hợp sau?

- 1) Chủ yếu dùng để chứa tài liệu văn bản.
- 2) Chủ yếu dùng để chứa các tệp ảnh du lịch, tham quan.
- 3) Chủ yếu dùng để chứa các tệp bài hát.



**Câu 1.** Số đếm biểu diễn bằng dãy bit 111 có bằng với số 111 ở hệ thập phân không? Vì sao?

**Câu 2.** Có bạn nói: 'Trong máy tính điện tử, các số được biểu diễn như trong hệ thập phân chúng ta quen dùng, vì người ta vẫn nhập các số thập phân vào máy tính để tính toán'. Em có đồng ý với ý kiến đó không? Vì sao?

## TÓM TẮT BÀI HỌC

- ❖ Mọi dữ liệu trong máy tính đều là dãy bit.
- ❖ Xử lý thông tin của máy tính gồm các bước: xử lý đầu vào; xử lý dữ liệu; xử lý đầu ra.
- ❖ Đơn vị đo lường dữ liệu lớn là các bội số của byte: KB, MB, GB, TB.