

HỌC VIỆN NÔNG NGHIỆP VIỆT NAM
KHOA CÔNG NGHỆ THÔNG TIN

CHỦ BIÊN: PHẠM QUANG DŨNG

GIÁO TRÌNH
TIN HỌC ĐẠI CƯƠNG

2015

TaiLieu.vn

LỜI NÓI ĐẦU

Cuốn giáo trình Tin học đại cương này được viết bởi các giảng viên Khoa Công nghệ thông tin, Học viện Nông nghiệp Việt Nam. Giáo trình bao gồm những kiến thức cơ bản và những kiến thức cập nhật của một số khía cạnh chủ yếu trong lĩnh vực Khoa học máy tính và Công nghệ thông tin như phần cứng, phần mềm, hệ điều hành, mạng máy tính và Internet, cơ sở dữ liệu, thuật toán, ngôn ngữ lập trình, an toàn thông tin. Mỗi chương được viết bởi giảng viên có kinh nghiệm giảng dạy nhiều năm các học phần chuyên về nội dung tương ứng.

Cuốn giáo trình này được dùng để làm tài liệu giảng dạy và học tập chính cho học phần Tin học đại cương gồm 2 tín chỉ trong chương trình đào tạo của các ngành không chuyên Tin học thuộc Học viện Nông nghiệp Việt Nam. Tất nhiên giáo trình cũng có thể được dùng làm tài liệu tham khảo cho nhiều đối tượng bạn đọc khác muốn mở mang kiến thức.

Giáo trình gồm 7 chương, mỗi chương là một mảng kiến thức tương đối độc lập với các chương khác. Bạn đọc có thể lựa chọn đọc những phần phù hợp với nhu cầu mà hầu như không gặp trở ngại về sự đòi hỏi kiến thức từ các phần trước.

Chương 1: Giới thiệu chung, do TS. Phạm Quang Dũng và ThS. Trần Thị Thu Huyền viết. Chương này giới thiệu một số kiến thức nền tảng chung gồm: (1) Các khái niệm về dữ liệu, thông tin, tin học, công nghệ thông tin; (2) Các hệ thống số dùng trong máy tính; (3) Mã hóa và biểu diễn thông tin trong máy tính; và (4) Các ứng dụng của công nghệ thông tin.

Chương 2: Cấu trúc máy tính, do TS. Phạm Quang Dũng biên soạn. Chương 2 đề cập đến: (1) Chức năng, sơ đồ tổ chức, nguyên lý hoạt động của máy tính; (2) Cấu trúc và chức năng của các thành phần cơ bản của máy tính, thuộc 4 khối chức năng gồm: bộ xử lý trung tâm, bộ nhớ, thiết bị ngoại vi và liên kết hệ thống.

Chương 3: Phần mềm máy tính và hệ điều hành, được viết bởi ThS. Phạm Thủy Vân. Qua chương này bạn đọc có thể hiểu được các vấn đề về: (1) Khái niệm và phân loại phần mềm máy tính, quy trình phát triển phần mềm; (2) Khái niệm, lịch sử phát triển hệ điều hành, một số hệ điều hành thông dụng cho máy tính và các thiết bị di động.

Chương 4: Mạng máy tính và Internet, do ThS. Phan Thị Thu Hồng đảm nhiệm. Chương này giới thiệu đến bạn đọc các kiến thức cơ bản về mạng máy tính và Internet bao gồm: (1) Khái niệm, các thành phần cơ bản, các mô hình kết nối và giao thức mạng, phân loại mạng máy tính; (2) Các khái niệm, các dịch vụ phổ biến như world wide web, tìm kiếm, thư điện tử, lưu trữ đám mây.

Chương 5: Cơ sở dữ liệu, do ThS. Hoàng Thị Hà biên soạn. Trong chương này bạn đọc có thể nắm được những kiến thức về: (1) Cơ sở dữ liệu: khái niệm, lợi ích các mức thể hiện của cơ sở dữ liệu, hệ cơ sở dữ liệu; (2) Khái niệm và chức năng của hệ quản trị cơ sở dữ liệu; và (3) Các câu lệnh của ngôn ngữ truy vấn có cấu trúc (SQL).

Chương 6: Thuật toán và ngôn ngữ lập trình, được viết bởi ThS. Lê Thị Nhung. Chương này giới thiệu với bạn đọc: (1) Khái niệm, các tính chất, các cách diễn đạt thuật toán, thiết kế thuật toán và đánh giá độ phức tạp; (2) Khái niệm về ngôn ngữ lập trình, trình biên dịch và trình thông dịch.

Chương 7: Các vấn đề xã hội của công nghệ thông tin; do ThS. Nguyễn Văn Hoàng đảm nhiệm. Qua chương này bạn đọc có thể thu nhận được các kiến thức về: (1) An toàn thông tin như các tài nguyên có thể bị xâm phạm, các hình thức tấn công để lấy cắp hay phá hoại thông tin; (2) Một số điều trong Bộ luật hình sự về tội phạm trong lĩnh vực tin học; và (3) Vấn đề sở hữu trí tuệ nói chung và sở hữu trí tuệ trong công nghệ thông tin nói riêng.

Do đây là lần đầu tiên viết giáo trình Tin học đại cương theo hướng cải tiến nên chắc chắn chúng tôi không thể tránh khỏi những thiếu sót. Nhóm tác giả mong nhận được những ý kiến góp ý từ các bạn đọc để lần tái bản sau giáo trình sẽ tốt hơn. Mọi ý kiến đóng góp xin gửi về địa chỉ pqdung@vnua.edu.vn.

Chúng tôi xin chân thành cảm ơn!

Hà Nội, tháng 1 năm 2015

Nhóm tác giả

MỤC LỤC

LỜI NÓI ĐẦU.....	III
MỤC LỤC	V
DANH MỤC CHỮ VIẾT TẮT.....	VIII
CHƯƠNG 1. GIỚI THIỆU CHUNG.....	1
1.1. MỘT SỐ KHÁI NIỆM CƠ BẢN	1
1.1.1. Dữ liệu	1
1.1.2. Thông tin	1
1.1.3. Tin học.....	2
1.1.4. Công nghệ thông tin	3
1.2. CÁC HỆ THỐNG SỐ VÀ CÁC PHÉP TOÁN DÙNG TRONG MÁY TÍNH.....	3
1.2.1. Các hệ thống số	3
1.2.2. Chuyển đổi giữa các hệ cơ số	4
1.2.3. Các phép toán số học trên hệ 2	6
1.3. BIỂU DIỄN VÀ MÃ HÓA THÔNG TIN	8
1.3.1. Biểu diễn thông tin trong máy tính và các đơn vị thông tin	8
1.3.2. Khái niệm về mã hóa.....	9
1.3.3. Mã hóa tập ký tự.....	10
1.3.4. Mã hóa số nguyên và số thực	11
1.3.5. Mã hóa dữ liệu logic.....	13
1.3.6. Mã hóa hình ảnh tĩnh.....	13
1.3.7. Mã hóa âm thanh và phim ảnh	15
1.4. ỨNG DỤNG CỦA CÔNG NGHỆ THÔNG TIN.....	15
1.4.1. Các bài toán khoa học kỹ thuật.....	15
1.4.2. Các bài toán quản lý	16
1.4.3. Tự động hóa.....	17
1.4.4. Công tác văn phòng	17
1.4.5. Giáo dục.....	17
1.4.6. Thương mại điện tử	18
CHƯƠNG 2. CẤU TRÚC MÁY TÍNH	19
2.1. GIỚI THIỆU	19
2.2. CHỨC NĂNG VÀ SƠ ĐỒ CẤU TRÚC CỦA MÁY TÍNH	20
2.2.1. Chức năng của máy tính:	20
2.2.2. Sơ đồ cấu trúc chung của máy tính.....	20
2.2.3. Nguyên lý hoạt động của máy tính.....	21
2.3. CÁC THÀNH PHẦN CƠ BẢN CỦA MÁY TÍNH	22
2.3.1. Bộ xử lý trung tâm.....	23
2.3.2. Bộ nhớ	26

2.3.3. Thiết bị vào/ra.....	35
2.3.4. Liên kết hệ thống	40
CHƯƠNG 3. PHẦN MỀM MÁY TÍNH VÀ HỆ ĐIỀU HÀNH	42
3.1. PHẦN MỀM MÁY TÍNH.....	42
3.1.1. Khái niệm về phần mềm.....	42
3.1.2. Phân loại phần mềm	43
3.1.3. Quy trình phát triển phần mềm.....	45
3.1.4. Phần mềm mã nguồn đóng và mã nguồn mở.....	46
3.2. HỆ ĐIỀU HÀNH.....	47
3.2.1. Khái niệm hệ điều hành	47
3.2.2. Lịch sử phát triển và phân loại hệ điều hành	48
3.2.3. Một số hệ điều hành điển hình.....	51
3.2.4. Quản lý dữ liệu trên bộ nhớ ngoài	57
CHƯƠNG 4. MẠNG MÁY TÍNH VÀ INTERNET	59
4.1. MẠNG MÁY TÍNH.....	59
4.1.1. Các thành phần cơ bản của mạng máy tính	59
4.1.2. Mô hình kết nối và giao thức mạng	60
4.1.3. Phân loại mạng máy tính	62
4.2. INTERNET	63
4.2.1. Một số khái niệm	63
4.2.2. Kết nối Internet	67
4.3. MỘT SỐ DỊCH VỤ CƠ BẢN CỦA INTERNET	67
4.3.1. WWW (World Wide Web).....	67
4.3.2. Tìm kiếm.....	68
4.3.3. Thư điện tử.....	72
4.3.4. Lưu trữ dữ liệu đám mây	76
CHƯƠNG 5. CƠ SỞ DỮ LIỆU	80
5.1. CƠ SỞ DỮ LIỆU	80
5.1.1. Khái niệm cơ sở dữ liệu.....	80
5.1.2. Các mức thể hiện của cơ sở dữ liệu	81
5.1.3. Mô hình dữ liệu quan hệ.....	83
5.1.4. Hệ cơ sở dữ liệu.....	85
5.1.5. Lợi ích của hệ cơ sở dữ liệu.....	85
5.2. HỆ QUẢN TRỊ CƠ SỞ DỮ LIỆU.....	86
5.2.1. Khái niệm.....	86
5.2.2. Phân loại hệ quản trị cơ sở dữ liệu	87
5.2.3. Chức năng cơ bản của hệ quản trị cơ sở dữ liệu.....	88
5.3. NGÔN NGỮ TRUY VẤN SQL.....	88
5.3.1. Câu lệnh truy vấn dữ liệu.....	89
5.3.2. Câu lệnh cập nhật dữ liệu	99

5.3.3. Thêm dữ liệu.....	99
5.3.4. Xóa dữ liệu	100
5.3.5. Các hàm của SQL	100
CHƯƠNG 6. THUẬT TOÁN VÀ NGÔN NGỮ LẬP TRÌNH	103
6.1. PHƯƠNG PHÁP GIẢI QUYẾT VẤN ĐỀ BẰNG MÁY TÍNH.....	103
6.2. THUẬT TOÁN	103
6.2.1. Khái niệm thuật toán	103
6.2.2. Các tính chất của thuật toán.....	105
6.2.3. Cách diễn đạt thuật toán	106
6.2.4. Thiết kế thuật toán	108
6.2.5. Độ phức tạp của thuật toán và vấn đề đánh giá thuật toán	110
6.3. NGÔN NGỮ LẬP TRÌNH.....	113
6.3.1. Khái niệm về ngôn ngữ lập trình	113
6.3.2. Lịch sử phát triển của ngôn ngữ lập trình.....	113
6.3.3. Trình biên dịch và trình thông dịch	117
6.3.4. Các công việc của người lập trình	117
CHƯƠNG 7. CÁC VẤN ĐỀ XÃ HỘI CỦA CÔNG NGHỆ THÔNG TIN.....	122
7.1. CÁC TÀI NGUYÊN CÓ THỂ BỊ XÂM PHẠM.....	122
7.1.1. Nội dung thông tin.....	122
7.1.2. Tài nguyên hạ tầng công nghệ thông tin	122
7.1.3. Định danh người dùng	122
7.2. CÁC HÌNH THỨC TẤN CÔNG	123
7.2.1. Tận dụng các lỗ hổng phần mềm.....	123
7.2.2. Sử dụng các phần mềm độc hại	123
7.2.3. Tấn công từ chối dịch vụ	127
7.2.4. Lừa đảo	129
7.3. SỞ HỮU TRÍ TUỆ.....	130
7.3.1. Tài sản trí tuệ	130
7.3.2. Quyền sở hữu trí tuệ	130
7.3.3. Luật sở hữu trí tuệ.....	130
7.4. CÁC QUY ĐỊNH, ĐIỀU LUẬT VỀ AN TOÀN THÔNG TIN VÀ SỞ HỮU TRÍ TUỆ	132
7.4.1. Các điều trong Bộ luật hình sự	132
7.4.2. Điều trong Nghị định Chính phủ	135
7.4.3. Các điều trong Luật Công nghệ thông tin	137
TÀI LIỆU THAM KHẢO.....	140

DANH MỤC CHỮ VIẾT TẮT

Chữ viết tắt	Từ tiếng Anh	Nghĩa tiếng Việt
AAC	Advanced Audio Coding	Mã hóa âm thanh tiên tiến
ADSL	Asymmetric Digital Subscriber Line	Đường dây thuê bao số bất đối xứng
ALU	Arithmetic and Logic Unit	Đơn vị toán học và logic, nằm trong CPU
AMD	Advanced Micro Devices	Các vi thiết bị tiên tiến, cũng là tên của một tập đoàn phát triển các thiết bị loại này như bộ vi xử lý máy tính và các công nghệ liên quan
ARM	Advanced RISC Machine	Một loại cấu trúc vi xử lý 32-bit kiểu RISC
ASCII	American Standard Code for Information Interchange	Bảng mã chuẩn của Mỹ dùng để trao đổi thông tin
ATA	Advanced Technology Attachment	Cách gọi ngắn gọn của Paralell ATA, là một chuẩn giao tiếp kết nối giữa máy tính và các ổ đĩa cứng, ổ đĩa quang trong máy tính
ATI	Array Technology Inc.	Tên tập đoàn được thành lập năm 1985, chuyên sản xuất các đồ họa cho máy tính cá nhân. Năm 2006, ATI được tập đoàn AMD mua lại.
BD	Bluray Disk	Một dạng đĩa quang cho phép ghi/phát lại hình ảnh/âm thanh với chất lượng cao
BIOS	Basic Input/Output System	Hệ thống vào/ra cơ bản
CCFL	Cold-Cathode Fluorescent Lamp	Đèn huỳnh quang catốt lạnh
CD	Compact Disc	Đĩa CD (gọn nhẹ), một dạng đĩa quang
CIDR	Classless Inter-Domain Routing	Lược đồ địa chỉ mới của Internet
CISC	Complex Instruction Set Computer [Architecture]	Kiến trúc tập lệnh phức tạp
CMOS	Complementary Metal–Oxide–Semiconductor,	Công nghệ bán dẫn kim loại bù, dùng để chế tạo các vi mạch tích hợp
CMYK		Hệ màu gồm 4 màu cơ sở: Cyan, Magenta, Yellow, Black
CNTT		Công nghệ thông tin
CPU	Central Processing Unit	Bộ xử lý trung tâm
CSDL		Cơ sở dữ liệu
CU	Control Unit	Đơn vị điều khiển, nằm trong CPU
DBMS	Database Management System	Hệ quản trị cơ sở dữ liệu

Chữ viết tắt	Từ tiếng Anh	Nghĩa tiếng Việt
DDR-SDRAM	Double Data Rate SDRAM,	SDRAM có tốc độ truyền dữ liệu gấp đôi
DIB	Device-Independent Bitmap	Dạng biểu diễn ảnh Bitmap mà không phụ thuộc vào thiết bị lưu trữ
DNS	Domain Name System,	Hệ thống tên miền Internet, cho phép thiết lập tương ứng giữa địa chỉ IP và tên miền
DRAM	Dynamic RAM	RAM động, được dùng để chế tạo bộ nhớ chính
DVD	Digital Video Disc, hoặc Digital Versatile Disc	Đĩa video số, hoặc đĩa đa năng số, một dạng đĩa quang
EEPROM	Electrically Erasable Programmable ROM	Bộ nhớ chỉ đọc có thể lập trình và xóa bằng điện
ENIAC	Electronic Numerical Interpolator and Computer	Máy tính và bộ tích hợp số điện tử, tên của chiếc máy tính điện tử đầu tiên trên thế giới (năm 1946)
EOF	End Of File	Tên của ký tự đặc biệt đánh dấu kết thúc tệp tin
FSB	Front Side Bus	Tên bus bên trong CPU
GPL	General Public License	Một điều kiện áp dụng cho việc sử dụng phần mềm nguồn mở
HD (1)	High Definition	Độ rõ nét cao
HD (2)	Hard Disk	Đĩa cứng, ngầm hiểu là dùng công nghệ từ tính. Một dạng ổ cứng khác dùng công nghệ flash là SSD – ổ cứng thể rắn.
HQTCSDL		Hệ quản trị cơ sở dữ liệu
HTML	Hypertext Markup Language	Ngôn ngữ đánh dấu siêu văn bản
IDE (1)	Integrated Drive Electronics	Một chuẩn kết nối khác giữa máy tính với các loại ổ đĩa nêu trên
IDE (2)	Integrated Development Environment	Môi trường phát triển tích hợp
IP	Internet Protocol	Giao thức mạng Internet
ISP	Internet Service Provider	Nhà cung cấp dịch vụ Internet
KHKT		Khoa học kỹ thuật
LAN/ MAN/ WAN/ VAN	Local/ Metropolitan/ Wide/ Vast Area Network	Mạng cục bộ / đô thị / diện rộng / toàn cầu
LCD	Liquid Crystal Display	Màn hình tinh thể lỏng
LED	Light-Emitting Diode	Điốt phát sáng
LPT	Line Printer Terminal	Loại cổng song song để kết nối với máy in theo dòng

Chữ viết tắt	Từ tiếng Anh	Nghĩa tiếng Việt
MIPS	Million Instructions per Second	Số triệu lệnh trên một giây
MPEG	Moving Picture Experts Group	Một định dạng tệp video
MS-DOS	Microsoft Disk Operating System	Hệ điều hành hướng đĩa của hãng Microsoft
NAT	Network Address Translation	Biên dịch địa chỉ mạng
NCP	Network Control Protocol	Giao thức điều khiển mạng
NIC	Network Interface Card	Cạc giao diện mạng
ODBMS	Object BDMS	Hệ quản trị cơ sở dữ liệu hướng đối tượng
OS	Operating System	Hệ điều hành
PC	Personal Computer	Máy tính cá nhân
RAM	Random Access Memory	Bộ nhớ truy nhập ngẫu nhiên
RDBMS	Relational BDMS	Hệ quản trị cơ sở dữ liệu quan hệ
RGB		Hệ màu gồm 3 màu cơ sở: Red, Green, Blue
RISC	Reduced Instruction Set Computer [Architecture],	Kiến trúc tập lệnh tập lệnh rút gọn
ROM	Read Only Memory	Bộ nhớ chỉ cho phép đọc
SATA	Serial ATA	Một chuẩn gắn kết nối tiếp, để nối ổ đĩa cứng hoặc ổ đĩa quang với bo mạch chủ
SDRAM	Synchronous DRAM	RAM đồng làm việc được đồng bộ bởi xung đồng hồ
SQL	Structured Query Language	Ngôn ngữ truy vấn có cấu trúc
SRAM	Static RAM	RAM tĩnh, được dùng để chế tạo bộ nhớ cache
SSD	Solid State Drive	Ổ cứng thể rắn, dùng công nghệ flash
TCP/IP	Transmission Control Protocol/ Internet Protocol	Giao thức điều khiển truyền dữ liệu/giao thức Internet
USB	Universal Serial Bus	Bus nối tiếp đa năng, là một chuẩn kết nối các thiết bị ngoại vi với máy tính
VGA	Video Graphics Array	Tên cạc đồ họa video hoặc cổng kết nối máy tính với màn hình hoặc máy chiếu (projector)
VXL		Vi xử lý
XML	eXtensible Markup Language,	Ngôn ngữ đánh dấu có thể mở rộng, có khả năng mô tả nhiều loại dữ liệu khác nhau, với mục đích chính là đơn giản hóa việc chia sẻ dữ liệu giữa các hệ thống khác nhau, đặc biệt là các hệ thống được kết nối với Internet

Chương 1

GIỚI THIỆU CHUNG

Chương 1 giới thiệu những kiến thức cơ bản và nền tảng nhất của Tin học. Mục 1.1 nêu những khái niệm về dữ liệu, thông tin, tin học và công nghệ thông tin. Mục 1.2 trình bày về biểu diễn dữ liệu trong máy tính, các hệ thống số và chuyển đổi giữa các hệ cơ số. Các mã hóa một số dạng dữ liệu thông dụng sẽ được trình bày trong mục 1.3. Cuối cùng, mục 1.4 sẽ giới thiệu những ứng dụng của công nghệ thông tin trong các lĩnh vực đời sống.

1.1. MỘT SỐ KHÁI NIỆM CƠ BẢN

Đối với chúng ta, quá trình hình thành trí tuệ bắt đầu từ việc thu nhận và xử lý dữ liệu rồi rạc để có thông tin, rồi kiểm nghiệm thông tin để có thể vận dụng vào mục đích cụ thể nào đó ta gọi là tri thức. Trí tuệ là khả năng sử dụng tri thức một cách khôn ngoan nhằm đạt được mục đích. Một ví dụ cụ thể về quá trình trên như sau:

- Dữ liệu: mưa, nắng, râm, cao, thấp, vừa, bay, chuồn chuồn, trời.
- Thông tin: khi thấy chuồn chuồn bay thấp thường thấy một lúc sau trời mưa, bay cao vừa phải thì trời râm mát, còn khi bay cao thì trời nắng.
- Tri thức: chuồn chuồn bay thấp trời mưa, bay cao trời nắng, bay vừa trời râm.
- Trí tuệ: khi thấy chuồn chuồn bay thấp thì ta cất quần áo đang phơi. Ví dụ đỉnh cao lợi dụng thời tiết thời Tam Quốc như Chu Du lợi dụng gió đánh hỏa công trận Xích Bích, Gia Cát Lượng lợi dụng mưa tuyết phá trận xe thiết xa của rợ Khương.

Trong lĩnh vực công nghệ thông tin, máy tính trợ giúp con người chủ yếu ở khâu đầu tiên, từ dữ liệu đến thông tin. Phần này sẽ giới thiệu tới bạn đọc các khái niệm về dữ liệu, thông tin, tin học và công nghệ thông tin.

1.1.1. Dữ liệu

Dữ liệu (Data) là những con số hoặc dữ kiện thuần túy, rời rạc do quan sát hoặc đo đếm được, không có ngữ cảnh hay diễn giải. Dữ liệu sau khi được tổ chức lại và xử lý sẽ cho ra thông tin.

Ví dụ: Với một quyển sách thì chữ, hình ảnh là dữ liệu còn nội dung của quyển sách là thông tin. Để biết được nội dung thì phải đọc sách. Việc đọc sách chính là xử lý dữ liệu.

Trong thực tế dữ liệu có thể là:

- Văn bản: Sách, báo, truyện, công văn...
- Các loại số liệu: Số liệu thống kê về nhân sự, thời tiết, kho tàng...
- Âm thanh, hình ảnh: Tiếng nói, âm nhạc, phim ảnh, tranh vẽ...

1.1.2. Thông tin

Thông tin (Information) là một khái niệm trừu tượng được thể hiện qua các thông báo, các biểu hiện..., đem lại một nhận thức chủ quan cho một đối tượng nhận tin. Thông tin là dữ liệu đã được xử lý xong, mang ý nghĩa rõ ràng.

Tương tự như dữ liệu, thông tin có thể tồn tại dưới nhiều hình thức khác nhau như âm thanh, hình ảnh, ký tự..., có thể được nén, giải nén, mã hóa, giải mã và được truyền tải qua các môi trường vật lý khác nhau như ánh sáng, sóng âm, sóng điện từ.

Ví dụ: Khi chúng ta nói chuyện trực tiếp với nhau, thông tin được thể hiện dưới dạng âm thanh và được truyền tải qua môi trường sóng âm. Còn khi chúng ta trao đổi với nhau qua điện thoại, thông tin được biểu diễn dưới dạng âm thanh nhưng được truyền tải qua môi trường sóng điện từ.

Tuy nhiên, giữa dữ liệu và thông tin không phải lúc nào cũng đồng nhất với nhau.

Ví dụ: Một số kí hiệu trong hệ đếm La Mã mang ý nghĩa thông tin là số nhưng trong hệ thống chữ La-tinh lại mang ý nghĩa là chữ cái.

Cùng một dữ liệu nhưng tùy thuộc vào đối tượng tiếp nhận dữ liệu khác nhau lại có thể cho ra khối lượng và chất lượng thông tin khác nhau.

Ví dụ: Cùng một tài liệu có sinh viên hiểu được 100% nhưng cũng có sinh viên chỉ hiểu được 50%, có sinh viên có thể phát triển liên hệ được với các vấn đề khác nhưng có sinh viên lại không...

Hình thức vật lý của thông tin được gọi là tín hiệu. Giữa thông tin và tín hiệu không phải lúc nào cũng đồng nhất với nhau. Cùng một thông tin có thể được biểu diễn bởi nhiều tín hiệu khác nhau và ngược lại cùng một tín hiệu có thể biểu diễn nhiều dạng thông tin khác nhau.

Ví dụ: Cùng một thông tin về chỉ dẫn giao thông nhưng có thể được biểu diễn dưới các tín hiệu khác nhau như là đèn tín hiệu giao thông hay là chỉ dẫn của cảnh sát giao thông... Hay cùng tín hiệu là gạt đầu trong từng trường hợp khác nhau lại biểu diễn thông tin khác nhau tùy thuộc vào câu hỏi...

1.1.3. Tin học

Thông tin nằm trong dữ liệu, xử lý thông tin bao gồm nhiều quá trình xử lý dữ liệu để rút ra thông tin hữu ích phục vụ con người. Khi xã hội càng phát triển thì khối lượng thông tin, dữ liệu ngày càng nhiều và con người không thể xử lý thông tin một cách thủ công được mà cần tới sự hỗ trợ của máy móc để xử lý thông tin một cách tự động. Trước yêu cầu đó của con người, một ngành khoa học mới đã ra đời, đó là Tin học.

Tin học (Informatics) là một ngành khoa học chuyên nghiên cứu các phương pháp, công nghệ và các kỹ thuật xử lý thông tin một cách tự động.

Hay nói một cách khác: Tin học là một ngành khoa học chuyên nghiên cứu về khả năng lưu trữ, truyền tải và xử lý thông tin. Điều này đã được thể hiện rõ qua quá trình hình thành và phát triển của ngành tin học. Trước đây, những thiết bị lưu trữ chưa đa dạng có dung lượng nhỏ với dung lượng tính theo Megabyte, Gigabyte được thay thế dần bởi các thiết bị lưu trữ đa dạng, gọn nhẹ hơn nhưng có dung lượng lưu trữ lớn hơn rất nhiều, tính theo Terabyte; Đường truyền thông tin có tốc độ thấp dần được thay thế bởi đường truyền tốc độ cao, không dây; Những bộ vi xử lý có tốc độ thấp dần được thay thế bởi những bộ vi xử lý tốc độ cao, bộ đa xử lý...

Sản phẩm mà tin học phát minh ra để giúp con người xử lý thông tin tự động là máy vi tính hay máy tính (computer).

Từ "tin học" đã được dịch từ từ *informatique* trong tiếng Pháp. Từ *informatics* trong tiếng Anh cũng bắt nguồn từ từ tiếng Pháp này, nhưng theo thời gian *informatics* đã mang nghĩa khác dần với nghĩa ban đầu. Ngày nay, thuật ngữ tiếng Anh tương đương với *informatique* là *computer science*, nghĩa là "khoa học máy tính". Thuật ngữ này được sử dụng rất rộng rãi trên thế giới và đa số các trường đại học ở nước ngoài sử dụng cụm từ *Computer Science* để đặt tên cho khoa chuyên môn.

Khoa học máy tính là ngành nghiên cứu các cơ sở lý thuyết về thông tin và tính toán cùng sự thực hiện và ứng dụng của chúng trong các hệ thống máy tính. Khoa học máy tính gồm nhiều ngành hẹp; một số ngành tập trung vào các ứng dụng thực tiễn cụ thể chẳng hạn như đồ họa máy tính, trong khi một số ngành khác lại tập trung nghiên cứu đến tính chất cơ bản của các bài toán tính toán như lý thuyết độ phức tạp tính toán. Ngoài ra còn có những ngành khác nghiên cứu các vấn đề trong việc thực thi các phương pháp tính toán. Ví dụ, ngành lý thuyết ngôn ngữ lập trình nghiên cứu những phương thức mô tả cách tính toán khác nhau, trong khi ngành lập trình nghiên cứu cách sử dụng các ngôn ngữ lập trình và các hệ thống phức tạp và ngành tương tác người-máy tập trung vào những thách thức trong việc làm cho máy tính và công việc tính toán hữu ích và dễ sử dụng đối với mọi người dùng. (trích từ wikipedia)

1.1.4. Công nghệ thông tin

Thuật ngữ *Công nghệ thông tin (Information Technology)* mang ý nghĩa về khía cạnh kỹ thuật, công nghệ hơn là khía cạnh khoa học. Ở Việt Nam, thuật ngữ này được sử dụng rộng rãi, dễ gây nhầm lẫn là “Công nghệ thông tin” mang nghĩa rộng hơn “Khoa học máy tính”.

Luật Công nghệ thông tin do Quốc hội nước Cộng hòa XHCN Việt Nam ban hành ngày 29/6/2006 có đưa ra khái niệm: “*Công nghệ thông tin là tập hợp các phương pháp khoa học, công nghệ và công cụ kỹ thuật hiện đại để sản xuất, truyền đưa, thu thập, xử lý, lưu trữ và trao đổi thông tin số*”.

Ở đây, thông tin số là thông tin được tạo lập bằng phương pháp dùng tín hiệu số. Trước đây, khi muốn gửi thư người ta thường dùng phương pháp chuyển thư tay nhưng giờ đây đã được thay thế bằng thư điện tử. Với khối lượng lớn công văn giấy tờ trong các cơ quan, xí nghiệp, trường học... việc lưu trữ số trở nên đơn giản và gọn nhẹ hơn.

1.2. CÁC HỆ THỐNG SỐ VÀ CÁC PHÉP TOÁN DÙNG TRONG MÁY TÍNH

Trong máy tính có sử dụng 3 hệ thống số là hệ cơ số 10 (gọi tắt là *hệ 10*), hệ cơ số 2 và hệ cơ số 16. Trong đó, hệ 2 là hệ cốt lõi được bộ vi xử lý sử dụng để tính toán, xử lý (ta sẽ xét kỹ hơn ở mục 1.3). Trong mục này sẽ giới thiệu các nội dung về 3 hệ thống số và chuyển đổi giữa chúng, các phép toán số học trên hệ 2 và các phép toán logic.

1.2.1. Các hệ thống số

a. Hệ cơ số 10 (Hệ thập phân - Decimal Numeral System)

Hệ 10 là hệ đếm được sử dụng để đếm và tính toán trong đời sống hàng ngày. Hệ 10 sử dụng 10 ký hiệu số 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 để biểu diễn các số. Các chương trình máy tính thường cho phép người dùng nhập vào các số hệ 10, xuất kết quả ở hệ 10, nhưng quá trình tính toán trung gian ở bên trong là các số hệ 2.

Khi làm việc với nhiều hệ thống số khác nhau, để phân biệt một số viết trong hệ cơ số này với một số viết trong hệ cơ số khác người ta thường viết kèm theo chỉ số có giá trị bằng cơ số của hệ đếm.

Ví dụ: 2092_{10} ; $789_{12_{10}}$; $12A_{16}$; 10110_2

Một số hệ 10 có thể biểu diễn ở dạng khai triển theo cơ số 10. Ví dụ:

$$8623,56_{10} = 8 \times 10^3 + 6 \times 10^2 + 2 \times 10^1 + 3 \times 10^0 + 5 \times 10^{-1} + 6 \times 10^{-2}$$

Trong đó:

- 8, 6, 2, 3, 5, 6 là các chữ số thành phần của số 8623,56

- 10 là cơ số của hệ
- Số mũ tương ứng với vị trí của chữ số thành phần: bằng 0 với chữ số phần nguyên nhỏ nhất, tăng dần về phía trái (1, 2, 3...), giảm dần về phía phải (-1, -2...).

Từ hệ cơ số 10 ta tổng quát hóa cho hệ cơ số a ($a \geq 2$). Số hệ a có các chữ số là b_i , $N_a = b_n b_{n-1} \dots b_1 b_0, b_{-1} b_{-2} \dots b_{-m}$ có thể biểu diễn theo cơ số a như sau:

$$N_a = b_n \times a^n + b_{n-1} \times a^{n-1} + \dots + b_1 \times a^1 + b_0 \times a^0 + b_{-1} \times a^{-1} + b_{-2} \times a^{-2} + \dots + b_{-m} \times a^{-m} \quad (1.1)$$

Giá trị của tổng ở vế phải trong công thức (1.1) được gọi là giá trị của số N_a . Công thức (1.1) sẽ được sử dụng để chuyển đổi số hệ a sang hệ 10.

b. Hệ cơ số 2 (Hệ nhị phân - Binary Numeral System)

- Hệ 2 hay hệ nhị phân chỉ sử dụng 2 ký hiệu số là 0 và 1 để biểu diễn các số. Đây là hệ cơ số cơ sở của máy tính. Máy tính chỉ lưu trữ và xử lý các dữ liệu ở dạng số nhị phân.

- Có thể biểu diễn một số trong hệ 2 ra thành tổng các hệ số nhân theo quy tắc nêu ở phần trên:

$$\text{Ví dụ: } 10011_2 = 1 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 19_{10}$$

c. Hệ cơ số 16 (Hexadecimal Numeral System)

- Hệ 16 sử dụng 16 ký hiệu để biểu diễn các số: 10 ký hiệu số từ 0, 1..., 9 để biểu diễn các giá trị từ 0 đến 9 và 6 ký hiệu chữ A, B, C, D, E, F để biểu diễn các giá trị từ 10 đến 15.

- Hệ 16 được dùng để đánh địa chỉ các ô nhớ, địa chỉ vật lý của các máy tính trong mạng (địa chỉ MAC), địa chỉ của các cổng vào-ra trong máy tính. Các địa chỉ này hiển thị cho người dùng ở dạng số hệ 16 mà không phải hệ 2 vì lý do làm cho địa chỉ ngắn gọn hơn, dễ nhớ, dễ sử dụng hơn.

- Có thể biểu diễn một số trong hệ 16 ra thành tổng các hệ số nhân với lũy thừa của cơ số.

$$\text{Ví dụ: } 12A_{16} = 1 \times 16^2 + 2 \times 16^1 + A \times 16^0 = 298_{10}$$

1.2.2. Chuyển đổi giữa các hệ cơ số

a. Chuyển từ hệ a sang hệ 10

Quy tắc: Muốn chuyển một số hệ a (2 hoặc 16) sang hệ 10 ta đem triển khai số trong hệ a ra thành tổng các hệ số nhân với lũy thừa của cơ số, khi đó ta sẽ được biểu thức trong hệ 10. Tính giá trị của biểu thức đó ta sẽ được số tương ứng trong hệ 10.

Ví dụ:

$$110101_2 = 1 \times 2^5 + 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = 32 + 16 + 4 + 1 = 53_{10}$$

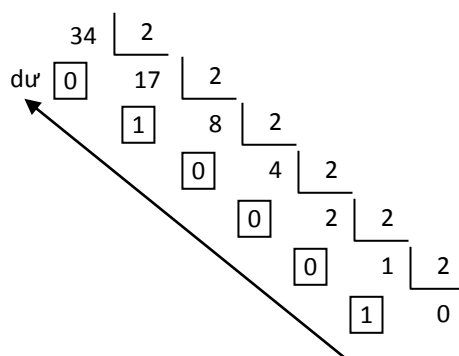
$$10F_{16} = 1 \times 16^2 + 0 \times 16^1 + F \times 16^0 = 256 + 15 \times 16^0 = 256 + 15 = 271_{10}$$

b. Chuyển từ hệ 10 sang hệ a

Ta chỉ xét trường hợp chuyển số nguyên hệ 10 sang hệ a .

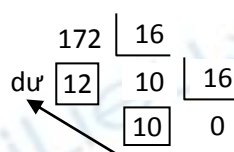
* Quy tắc: Đem số hệ 10 chia nguyên liên tiếp cho cơ số a cho tới khi thương bằng 0 thì dừng lại, với mỗi phép chia ta nhận được một số dư. Lấy các số dư của phép chia theo thứ tự ngược lại ta được số trong hệ a (số dư của phép chia cuối cùng là chữ số có trọng số lớn nhất, chữ số đầu tiên bên trái).

Ví dụ 1: Đổi số 34_{10} sang hệ 2, ta thực hiện các phép chia như sau:



Kết quả là $34_{10} = 100010_2$

Ví dụ 2: Đổi số 172_{10} sang hệ 16



Kết quả là $172_{10} = AC_{16}$ (10 ứng với A, 12 ứng với C)

c. Chuyển từ hệ 2 sang hệ 16

Bảng 1.1. Mười sáu số đầu tiên trong 3 hệ cơ số

Hệ 10	Hệ 2	Hệ 16
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F

Nhận xét: Từ bảng trên ta thấy khi hệ 16 dùng đến chữ số lớn nhất thì hệ 2 phải dùng đến 4 chữ số. Vì $16 = 2^4$, mỗi chữ số hệ 16 sẽ tương đương với 4 chữ số hệ 2.

Quy tắc: Để chuyển một số từ hệ 2 sang hệ 16 ta nhóm thành các nhóm 4 chữ số hệ 2 từ phải qua trái, sau đó chuyển từng nhóm 4 chữ số hệ 2 thành các chữ số hệ 16.

Ví dụ: $110\ 1100\ 1011_2 = ?_{16}$

$0110\ 1100\ 1011 \Rightarrow 6CB_{16}$

d. Chuyển từ hệ 16 sang hệ 2

Quy tắc: chuyển từng chữ số hệ 16 thành 4 chữ số hệ 2.

Ví dụ: $9C0A_{16} = ?_2 = 1001\ 1100\ 0000\ 1010_2$

1.2.3. Các phép toán số học trên hệ 2

a. Phép cộng

Bảng cộng hai bit:

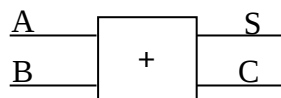
A	B	Tổng (Sum, S)	Số nhớ (Carry, C)
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

Cách thức thực hiện phép cộng hai số hệ 2 thủ công: Thực hiện cộng như trong hệ 10, cộng từng cột bit từ phải qua trái, có nhớ sang cột bit cao hơn.

Ví dụ:

$$\begin{array}{r} \text{C} \quad \quad 1\ 1 \\ \text{A} \quad 0\ 1\ 0\ 1\ 1 \\ \text{B} \quad 1\ 0\ 0\ 1\ 1 \\ \hline \text{A+B} \quad 1\ 1\ 1\ 1\ 0 \end{array}$$

Trong máy tính, phép cộng hai bit được thực hiện bằng mạch cộng như sau:



b. Phép trừ

Trong kỹ thuật máy tính, để tận dụng các mạch cộng đã có sẵn người ta thực hiện phép trừ thông qua phép cộng và phép lấy số đối: cộng số bị trừ với số đối của số trừ.

$$A - B = A + (-B)$$

Vấn đề đặt ra là phải có cách biểu diễn số âm trong hệ 2 trong máy tính để các phép tính toán vẫn cho kết quả đúng. Cách biểu diễn số nguyên và số thực (dương, âm) sẽ được trình bày ở mục 1.3.

c. Phép nhân và phép chia

Trong máy tính, phép nhân và phép chia được thực hiện qua phép cộng, phép trừ và phép dịch bit. Ở mức đại cương, chúng ta không xét sâu hơn hai phép toán này ở đây.

1.2.4. Các phép toán logic

Ngoài việc hỗ trợ các phép toán số học, máy tính cũng phải hỗ trợ các phép toán logic như NOT, AND, OR vì sự cần thiết của chúng. Ta hay phải làm việc với các phép toán logic trong bảng tính Excel hay trong khi lập trình. Ví dụ với ngôn ngữ lập trình Pascal, để kiểm tra một biến x có nằm trong khoảng $[7,8)$ hay không, nếu đúng thì viết dòng “Điểm B”:

```
If (x>=7) and (x<8) then  
    Writeln("Diem B");
```

Giá trị logic biểu diễn một trong hai trạng thái đối lập là đúng/sai, có/không. Trong đại số logic giá trị đúng được gọi là TRUE, giá trị sai được gọi là FALSE. Các phép toán logic tác động trên các giá trị logic TRUE, FALSE gồm có:

a. Phép toán NOT (phủ định hay đảo)

Bảng chân lý:

X	NOT X
FALSE	TRUE
TRUE	FALSE

b. Phép toán AND (và)

Bảng chân lý:

X	Y	X AND Y
FALSE	FALSE	FALSE
FALSE	TRUE	FALSE
TRUE	FALSE	FALSE
TRUE	TRUE	TRUE

Nhận xét: Phép toán AND chỉ cho kết quả “đúng” khi cả hai toán hạng đều “đúng”.

c. Toán tử OR (hoặc)

Bảng chân lý:

X	Y	X OR Y
FALSE	FALSE	FALSE
FALSE	TRUE	TRUE
TRUE	FALSE	TRUE
TRUE	TRUE	TRUE

Nhận xét: Phép toán OR chỉ cho kết quả “sai” khi cả hai toán hạng đều “sai”.

d. Toán tử XOR (eXclusive OR, hoặc loại trừ)

Bảng chân lý:

X	Y	X XOR Y
FALSE	FALSE	FALSE
FALSE	TRUE	TRUE
TRUE	FALSE	TRUE
TRUE	TRUE	FALSE

Nhận xét: Phép toán XOR cho kết quả “đúng” khi hai toán hạng khác nhau, cho kết quả “sai” khi hai toán hạng giống nhau.

e. Biểu thức logic và thứ tự ưu tiên các phép toán

- Biểu thức logic là sự kết hợp các giá trị logic bằng các phép toán logic để tạo ra một giá trị logic mới. Mỗi biểu thức logic có kết quả là một giá trị hoặc đúng (TRUE) hoặc sai (FALSE).

- Nếu trong biểu thức logic có chứa nhiều phép toán logic thì các phép toán logic được thực hiện theo thứ tự ưu tiên sau: NOT $\rightarrow$ AND $\rightarrow$ OR, XOR (OR và XOR cùng mức ưu tiên). Các phép toán cùng mức ưu tiên được thực hiện từ trái qua phải.

Ví dụ:

(Các SV có hộ khẩu Hà Nội) AND NOT (Các SV dân tộc Kinh)

Trong ví dụ này cần ưu tiên thực hiện phép NOT trước rồi mới thực hiện phép AND.

1.3. BIỂU DIỄN VÀ MÃ HÓA THÔNG TIN

1.3.1. Biểu diễn thông tin trong máy tính và các đơn vị thông tin

a. Biểu diễn thông tin trong máy tính

Như đã đề cập ở mục 1.1, dữ liệu thực có thể tồn tại ở nhiều dạng khác nhau: dạng số (số nguyên, số thực) hoặc phi số (chữ viết, âm thanh, hình ảnh). Tuy nhiên, máy tính chỉ xử lý được dữ liệu ở dạng số nhị phân (hệ cơ số 2). Do vậy, dữ liệu thực bất kể ở dạng nào muốn đưa vào máy tính để lưu trữ, xử lý, hay truyền tải cần phải được mã hóa (số hóa thành số nhị phân). Sau khi xử lý, dữ liệu kết quả sẽ được khôi phục lại ở dạng dễ hiểu, dễ cảm nhận đối với con người.

Sở dĩ trong máy tính chỉ dùng được số nhị phân để biểu diễn thông tin là vì các linh kiện và vật liệu điện tử dùng để chế tạo bộ nhớ trong của máy tính (flip-flop, tụ điện) chỉ có hai trạng thái đối lập (ví dụ như *có điện* hoặc *không có điện*), tương ứng được biểu diễn là 1 và 0.

b. Các đơn vị thông tin

Trong khoa học máy tính, mỗi chữ số nhị phân được gọi là một bit (viết tắt của từ tiếng Anh là BInary digiT). Mỗi ô nhớ trong máy tính lưu trữ được một bit. Bản thân mỗi ô nhớ đó còn được gọi là một bit. Các bit được đánh số thứ tự bắt đầu từ 0.

Một nhóm 8 bit bắt đầu từ bit thứ 8i ($i \geq 0$, nguyên) được gọi là một byte. Các byte được đánh địa chỉ bắt đầu từ 0. Byte 0 gồm các bit từ 0 đến 7, byte 1 gồm các bit từ 8 đến 15... Một đơn vị nữa cũng hay được sử dụng là word (từ nhớ). Một word gồm 2, 4, hay 8 byte tùy thuộc vào bộ vi xử lý (CPU) cụ thể có thể xử lý mỗi lần được bao nhiêu byte.

Các đơn vị bội của byte hay được sử dụng và cách quy đổi như sau:

1 Kilobyte (1 KB) = 2^{10} byte = 1024 byte

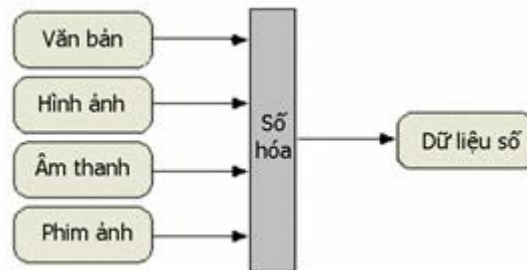
1 Megabyte (1 MB) = 2^{10} KB = 2^{20} byte = 1.048.576 byte

1 Gigabyte (1 GB) = 2^{10} MB = 2^{20} KB = 2^{30} byte

1 Terabyte (1 TB) = 2^{10} GB = 2^{20} MB = 2^{30} KB = 2^{40} byte

1.3.2. Khái niệm về mã hóa

Mã hóa thông tin trong máy tính thực chất là số hóa dữ liệu thành các chuỗi số nhị phân (hình 1.1) theo những quy ước chung để các máy tính có thể lưu trữ, xử lý và trao đổi thông tin với nhau.



Hình 1.1. Sơ đồ số hóa dữ liệu

Trong máy tính người ta dùng các số nhị phân có độ dài (số bit) cố định để biểu diễn thông tin. Các số nhị phân này được gọi là từ mã. Với độ dài từ mã là n , ta có thể biểu diễn được 2^n thông tin khác nhau.

Ví dụ:

- Nếu dùng 1 byte (8 bit) để biểu diễn các số nguyên không dấu thì ta có thể biểu diễn được $2^8 = 256$ số có giá trị từ 0 đến 255 như sau:

Từ mã	Số nguyên
0000 0000	0
0000 0001	1
0000 0010	2
...	
1111 1111	255

- Nếu dùng 1 byte để biểu diễn các ký tự (chữ cái, chữ số thập phân, các dấu chấm câu, các ký hiệu phép toán...) thì có thể biểu diễn được $2^8 = 256$ ký tự khác nhau.

Cách mã hóa các loại dữ liệu được tuân theo những chuẩn chung để các máy tính có thể “hiểu” được nhau khi trao đổi, xử lý thông tin.

- Các ký tự: mã hóa theo bảng mã ASCII hoặc Unicode
- Các số nguyên: mã hóa theo một số chuẩn quy ước
- Các số thực: mã hóa theo số dấu phẩy động
- Dữ liệu ảnh, âm thanh, phim: mã hóa rời rạc thành các ma trận số thực biểu diễn cường độ sáng, tần số âm.

Câu hỏi đặt ra là làm sao máy tính phân biệt được chuỗi số nhị phân nào ứng với dữ liệu dạng số, chuỗi nào ứng với dạng ký tự... Ví dụ với một chuỗi 8 bit 0100 0001, khi nào thì máy tính cần hiểu đó là biểu diễn của số 65_{10} , khi nào cần hiểu là biểu diễn của ký tự ‘A’? Để thực hiện được điều này, bằng cách nào đó các chương trình máy tính hoặc người sử dụng phải khai

báo kiểu và cấu trúc dữ liệu của các thành phần trong chương trình để hệ điều hành ghi nhớ vào các vùng nhớ thích hợp có địa chỉ và kích thước xác định. Ví dụ, với file ảnh thì các thông tin đó được chương trình tạo ảnh số lưu ở đầu file. Với các ngôn ngữ lập trình, người lập trình sẽ khai báo các hằng, biến qua các câu lệnh. Ta lấy ví dụ trong ngôn ngữ Pascal như sau:

```
Var   ch: char;           {biến ch có kiểu ký tự}
      st: string;         {biến st có kiểu chuỗi ký tự}
      i: byte;            {biến i có kiểu số nguyên không dấu 8 bit}
      j: shortint;        {biến j có kiểu số nguyên có dấu 8 bit}
      k1: word;           {biến k1 có kiểu số nguyên không dấu 16 bit}
      k2: integer;        {biến k2 có kiểu số nguyên có dấu 16 bit}
      m: longint;         {biến m có kiểu số nguyên có dấu 32 bit}
      r: real;            {biến r có kiểu số thực 6 byte}
      t: double;          {biến t có kiểu số thực 8 byte}
```

Các mục con tiếp theo sẽ trình bày việc mã hóa các loại dữ liệu thông dụng gồm: ký tự, số nguyên, số thực, ảnh tĩnh, âm thanh và phim ảnh.

1.3.3. Mã hóa tập ký tự

Về nguyên tắc, mỗi quốc gia đều có thể tự thiết kế một bảng mã riêng để biểu diễn các ký tự của nước mình. Nếu làm như vậy thì các máy tính và thậm chí các bộ phận của cùng một máy tính sẽ không hiểu nhau khi kết nối với nhau. Bởi vậy, các nước phải quy định dùng chung một bảng mã để biểu diễn ký tự, bảng mã này được gọi là bảng mã chuẩn. Trong thực tế có nhiều bảng mã chuẩn nhưng được sử dụng phổ biến nhất trên máy tính hiện nay là bảng mã ASCII (America Standard Code for Information Interchange) và bảng mã Unicode.

Bảng mã ASCII có 256 từ mã 8 bit, biểu diễn 256 ký tự khác nhau. Bảng mã ASCII được chia thành phần tiêu chuẩn (gồm các từ mã có giá trị trong hệ 10 từ 0 đến 127) và phần mở rộng (có mã từ 128 đến 255). Các máy tính trên thế giới có bảng mã ASCII tiêu chuẩn giống nhau, phần mở rộng có thể khác nhau vì được dùng để biểu diễn các ký tự của riêng từng nước.

Bảng 1.2. Bảng mã ASCII tiêu chuẩn

		Các bit cao							
		000	001	010	011	100	101	110	111
Các bit thấp	0000	NUL	DLE	SP	0	@	P	\	p
	0001	SOH	DC1	!	1	A	Q	a	q
	0010	STX	DC2	"	2	B	R	b	r
	0011	ETX	DC3	#	3	C	S	c	s
	0100	EDT	DC4	\$	4	D	T	d	t
	0101	ENQ	NAK	%	5	E	U	e	u
	0110	ACK	SYN	&	6	F	V	f	v
	0111	BEL	ETB	,	7	G	W	g	w
	1000	BS	CAN	(	8	H	X	h	x
	1001	HT	EM	)	9	I	Y	i	y
	1010	LF	SUB	*	:	J	Z	j	z
	1011	YT	ESC	+	;	K	[	k	{
	1100	FF	FS	.	<	L	C	l	ù
	1101	CR	GS	-	=	M	]	m	}
	1110	SO	RS	.	>	N	↑	n	≈
	1111	SI	US	/	?	O	<--	o	DEL
		code ASCII							

Ví dụ: ký tự 'A' được mã hóa thành 0100 0001 (= 65₁₀). Khi ta ấn Shift+A trên bàn phím, một xung điện truyền đến bộ xử lý máy tính có dạng tương ứng là: