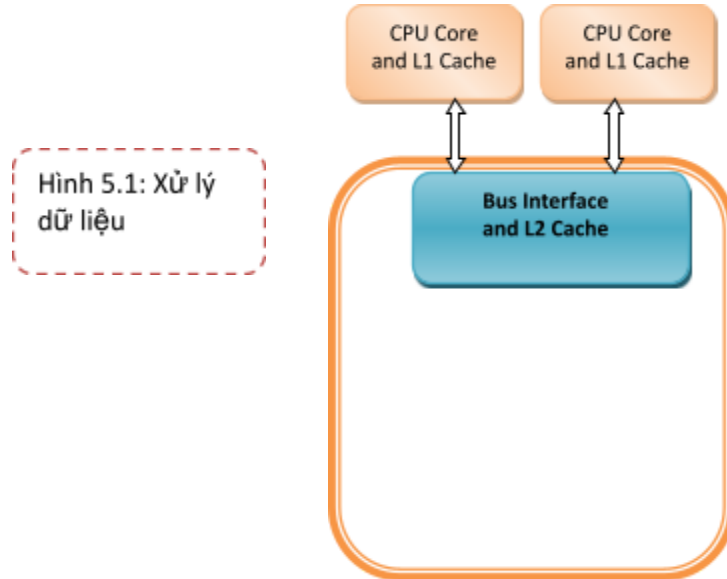
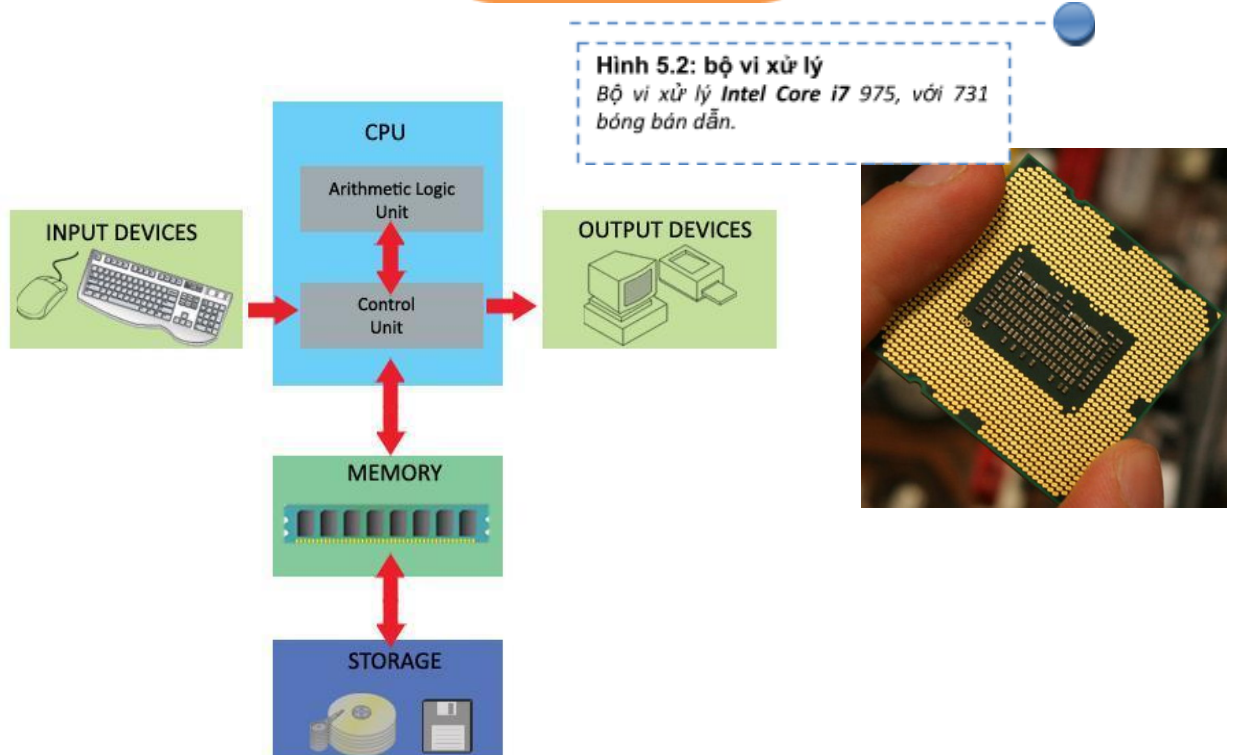


Trong các phần trước, bạn có thể làm việc và chơi với máy tính như một người dùng đầu cuối. Tuy nhiên, để trở thành một lập trình viên, bạn cần phải hiểu thêm một vài kỹ thuật xử lý ở mức thấp hơn bên trong máy tính. Trong phần này sẽ giải thích cách biểu diễn và xử lý dữ liệu bên trong bộ vi xử lý (CPU) và bộ nhớ. Kiến thức này sẽ giúp bạn hiểu “bộ não máy tính” hoạt động như thế nào. Thêm vào đó, trong phần này cũng mô tả về các thiết bị và các kỹ thuật lưu trữ dữ liệu.



Hình 5.1: Xử lý dữ liệu



Hình 5.2: bộ vi xử lý
Bộ vi xử lý Intel Core i7 975, với 731 bóng bán dẫn.

5.1 Dữ liệu và thông tin

Dường như các máy tính hiểu được con người. Tuy nhiên, máy tính thuần túy chỉ là cỗ máy điện tử. Chúng nhận diện hai trạng thái vật lý riêng biệt được tạo ra bởi điện, từ tính hoặc ánh sáng phản xạ. Cơ bản, bộ xử lý trung tâm (CPU), não bộ của máy tính, bao gồm hàng trăm triệu mạch vi điện tử bật/tắt gọi là các transistor. Máy tính hiểu thông tin chỉ bởi vì nó làm việc ở tốc độ rất cao, nhóm các trạng thái bật/tắt riêng biệt thành các mẫu có thể hiểu được với con người. Trong điện toán, dữ liệu là khái niệm được dùng để mô tả thông tin được biểu diễn bởi các nhóm thiết bị chuyển mạch bật/tắt. Mặc dù hai từ dữ liệu và thông tin thường được sử dụng thay thế cho nhau nhưng có một sự khác biệt quan trọng giữa hai từ này. **Dữ liệu** bao gồm các số liệu thô mà máy tính tổ chức để tạo ra thông tin có nghĩa (và hữu ích) cho con người. Do sự khác biệt này, một trong những công việc quan trọng nhất của máy tính đó là chuyển đổi dữ liệu thành thông tin.

5.2 Biểu diễn dữ liệu

Trong máy tính, dữ liệu được biểu diễn dưới dạng con số. Điểm CF bạn đạt được sau khóa học CF, các bức thư, các dấu chấm câu, âm thanh, video, v.v. đều là các con số. Thậm chí các mã lệnh của máy tính cũng là các con số. Khi bạn đọc dòng “*Computing Fundamentals*” trên màn hình, đó là bạn chỉ thấy một cách khác mà máy tính sử dụng để biểu diễn các con số.

Dữ liệu máy tính khá xa lạ bởi con người thường sử dụng cơ số 10 để biểu diễn các con số. Hệ thập phân được gọi là cơ số 10 vì các con số được biểu diễn bằng 10 ký hiệu : 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. Khi bạn cần biểu diễn một số lớn hơn 9, bạn sử dụng hai ký hiệu cùng nhau, chẳng hạn $9+1=10$. Mỗi ký hiệu trong một số được gọi là chữ số, do đó 10 là một số có 2 chữ số.

Tuy nhiên, trong máy tính, dữ liệu được biểu diễn bởi trạng thái của các thiết bị chuyển mạch điện tử trong máy tính. Thiết bị chuyển mạch chỉ có hai trạng thái – bật (on) và tắt (off). Do vậy nó chỉ có thể biểu diễn hai giá trị số. Đối với máy tính, khi thiết bị chuyển mạch ở trạng thái tắt, nó thể hiện giá trị 0; khi bật nó biểu diễn giá trị 1. Bởi vì chỉ có hai giá trị nên máy tính cũng được gọi là thực hiện tính toán trên cơ số 2, cũng còn gọi là **hệ số nhị phân**.

Khi máy tính cần biểu diễn giá trị lớn hơn 1, nó thực hiện giống như bạn biểu diễn giá trị lớn hơn 9 trong cơ số 10 : nó sử dụng hai hay nhiều hơn chữ số nhị phân.

Đối với những bạn yêu thích toán học, các bạn có thể xem chi tiết hơn về cách biểu diễn số trong bài viết **Các hệ thống số** nằm trong cuối chương này.

5.3 Bit (phát âm là “bít”) và Byte (phát âm là “bai”)

Khi nói tới dữ liệu được tin học hóa, giá trị được biểu diễn bởi trạng thái của từng thiết bị chuyển mạch gọi là **bit** (tổ hợp của các chữ số nhị phân). Bit là đơn vị nhỏ nhất của dữ liệu mà một máy tính có thể biểu diễn và sử dụng. Để biểu diễn bất cứ điều gì có ý nghĩa, máy tính sử dụng các nhóm bit.

Một nhóm 8 bit gọi là một byte. Byte là đơn vị cực kỳ quan trọng, vì các tổ hợp đầy đủ 8 bit khác nhau biểu diễn tất cả các ký tự trên bàn phím, bao gồm các từ, các số, các dấu câu và các ký hiệu khác. Bit và byte và đơn vị chẳng hạn kilobits, kilobytes, megabytes, gigabytes, v.v. là các đơn vị thông tin được dùng trong máy tính và truyền thông dữ liệu.