Basic Computer Structure

A computer is composed of wires and switches. Data is transmitted down wires and each wire can have two different values: on (has current) or off (no current).

So a wire transmits two different values: 1 (on) or 0 (off).

If we put 2 wires together, we get 4 different values: both off (00), both on (11), the first on and the second off (10), the first off and the second on (01).

If we put 3 wires together, we get 8 different values: 000, 001, 010, 011, 100, 101, 110, 111. (Note that 8 = 2^3.)

If we put 8 wires together, we get 2^8 = 256 different values.

Definition: A bit is the smallest unit of information. A bit is two values, usually denoted as 0 and 1.

A byte is 8 bits. A byte has 2^8 = 256 different values.

The main processing in a computer is done in the ALU (Arithmetic-Logic Unit) that is a part of the CPU (Central Processing Unit). The ALU is made up of separate circuits

to perform different tasks like addition, multiplication, etc. The ALU has three sets of input wires and one set of output wires. Two of the sets of input wires are

for the data to the ALU (for example, if you are going to add two numbers, one number arrives in the first set of wires, and the other number on the second set).

The third input set of wires controls which operation the ALU should perform.

x-bit Machines

We sometimes refer to machines as x-bit machines. For example, your computer was probably advertised as a 64-bit machine (newer ones) or a 32-bit machine (older ones).

The number refers to how many wires go into the computer circuits, and that means how much data the computer can process in one step.

For example, in a 32-bit machine's each set of input wires to the ALU will consist of 32 wires and the output will also consist of 32 wires.

32 wires together is 32 bits of information, and that means each input can be one of 2^32 different values.

As a result, there is a limit to how many different values a computer can add (or subtract or multiply, etc) directly on the chip.

If we need to do arithmetic on larger numbers, we need to write programs for it.

With a 64-bit machine, we get a lot more values, but the same idea holds. There are a limited number of values we can manipulate directly on the chip. Anything more requires a program.

Key idea:

All data on a computer is transferred by a group of wires. All data is represented in bits. Therefore all data (whether an integer, a fraction, a music file, etc.) is just a number (and just an integer).

Even the signals that control the ALU are sent on groups of wires, and so those signals are just numbers.

A Quick Introduction to Binary

The rules are exactly the same for binary numbers as for decimal (normal) numbers.

In decimal, we run through the ten unit digits. When we run out of digits, we return to 0 and increase the "ten's" digit.

In binary, we run through the two unit digits. When we run out of digits, we return to 0 and increase the "two's" digit.

Decimal number (base 10) Binary number (base 2)

0 0

1 1

2 10

3 11

4 100

5 101

6 110

7 111

8 1000

9 1001

10 1010

11 1011

Note that the different digits in decimal are the "one's" digit, the "ten's" digit, the "hundred's" digit, and so on.

For binary, the digits are powers of 2 instead of powers of 10, so we have the "one's" digit, the "two's digit", the "four's" digit, and so on.

A Short History of Programming Languages

Remember that a computer is just wires and switches. All data is represented by wires that are ON or OFF. (1 or 0) So, all data is just a number in binary.

Likewise, the commands to control the ALU and CPU circuits are input on wires. Thus, each command the computer understands is also just a number in binary.

These number commands that control the computer are called "machine code". Each operation a machine can do is a number and every computer make will have a different

machine code. It is not easy for a human to write machine code and almost impossible for a human to read it.

To be continued next lecture.