CSCI 120 Introduction to Computation
History of computing (cont.) (draft)

Saad Mneimneh
Visiting Professor
Hunter College of CUNY

1 The electrical age

The technology in the years 1600 - 1900 was unable to produce the complex gear driven machines of Pascal, Leibniz, and Babbade in an economically feasible manner. But with the advances of electronics in the early 1900s, this barrier was overcome. It was clear at that point in time that what is needed is a programmable machine that uses electrical signals instead of punched cards to create programs. With electronics, however, the only way to simulate holes in punched cards is through the presence or absence of electric current. Therefore, the concept of a switch emerged. A number of switches is used to configure and program the machine. Needless to say, those switches must be operated by the machine itself and not by a human being. At first, the switches were electro-mechanical. For example, an electromechanical machine was completed by George Stibitz at Bell Laboratories in 1940. It was designed to perform calculations of complex numbers, and it used the binary system described by Leibniz. Similarly, Mark I was completed in 1944 at Harvard University by Howard Aiken and a group of IBM engineers. This was a general purpose computer influenced by Babbage's Analytical Engine (and hence used the decimal system). Mark I was followed by Mark II. The following figure shows Mark I.

Figure 1: Mark I

These machines made heavy use of electronically controlled mechanical relays. The relays act as on-off switches to control various operations of the machine. In some sense, they encode the logic or the "smartness" of the computer. For instance, one could say *if A OR B then C*. Here's a schematic for this statement using relays (electric current flows when either switch is closed).

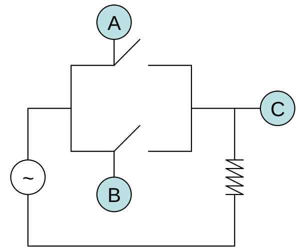

Figure 2: OR logic

A concept of an electromechanical relay is show below:

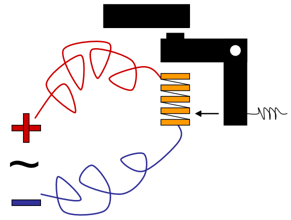

Figure 3: Electromechanical relay

The metal piece is held in place by means of a spring. The electric current creates a magnetic field through the coil that attracts the metal piece forward against the spring tension, causing it to be in contact with the upper piece. When the current ceases to exist, the spring brings back the piece to its original position.

Therefore, by the time these machines were built, they became obsolete, since other researchers were applying the technology of vacuum tubes to replace the mechanical relays and to construct totally electronic computers! So what is a vacuum tube? A vacuum tube is an pure electronic relay as shown below (influenced by the work of Thomas Edison):

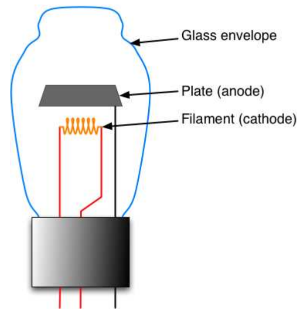

Figure 4: Vacuum tube

Vacuum tubes resemble light bulbs in that they have a filament sealed in a glass envelope which has been evacuated of all air. When hot, the filament releases electrons into the vacuum. This results in a negatively charged cloud of electrons. These electrons will be drawn to a metal ”plate” inside the envelope if the plate (also called the anode) is positively charged relative to the filament (or cathode). The result is a current of electrons flowing from filament to plate. This cannot work in the reverse direction because the plate is not heated and cannot emit electrons. Thus it operates as a switch: If the current flows in one direction, it conducts the current. If the current flows in the other direction, it does not.

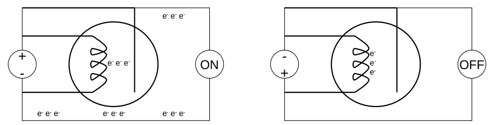

Figure 5: On and Off vacuum tube

The vacuum tubes could burn easily (much like light bulbs) and were unreliable. Although vacuum tubes are still in use in special purpose electrical equipments, they were replaced in the 1940s by *diodes* and *transistors*, which made it possible to construct very small logical circuits like AND, OR, and NOT gates (like the one described above for the OR logic). These are known as the Boolean operations (after the mathematician George Boole) and we will talk about them later. For the time being, let's look at the basic mechanism behind the operation of a diode (since it represents the modern vacuum tube).

A diode is made out of a semiconductor. A semiconductor is a material with an electrical conductivity that is intermediate between that of an insulator (does not conduct) and a conductor: it behaves as an insulator under normal conditions, but can be made to behave as a conductor by applying enough current. Thus it provides the desired behavior of a switch. Let's see how such a semiconductor material is constructed. Semiconductors are produced by a process called *doping*. This is the process of *intentionally introducing impurities* [1] into a pure material in order to change its electrical properties. A commonly used material is Silicon, *Si* (this is not silicone!). The following figure shows the Silicon atom *Si* which has 4 electrons on its outer shell, the Phosphorus atom *P* which has 5 electrons on its outer shell, and the Boron atom *B*, which has 3 electrons on its outer shell.

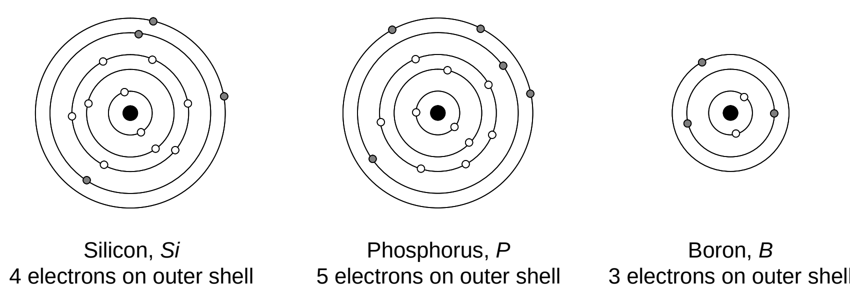

Figure 6: Silicon, Phosphrus, and Boron

[1] Apparently, a process the human race is good at, we keep introducing impurities to the environment around us :).

With only 4 electrons, *Si* is missing 4 electrons on the outer shell to become stable. Therefore, the silicon atoms are usually arranged in a lattice where each atom shares 4 electrons with its 4 neighbors as shown in Figure 7 (a).

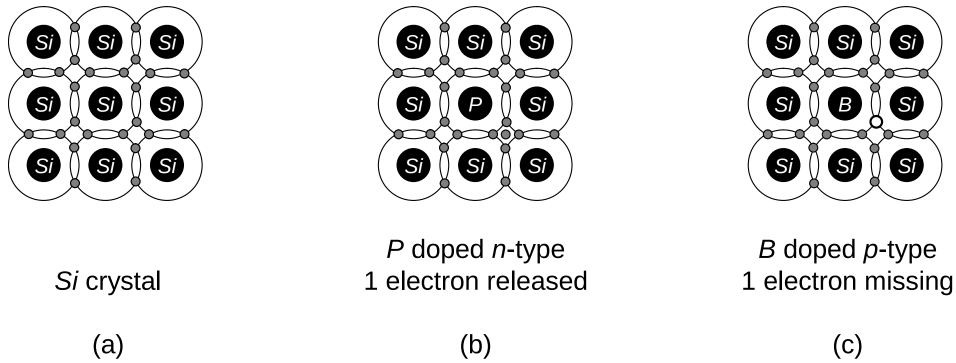

Figure 7: Silicon crystal with and without doping

When doped with phosphorus, the fifth electron on the outer shell of the phosphorus is released (this is an extra electron that is not needed), which becomes unbonded from the individual atoms and hence can move freely, Figure 7 (b). The electron donor is now positively charged. The doped crystal is called *n*-type.

When doped with boron, an electron is missing since boron has only 3 electrons on its outer shell. This creates a *hole* that can accept an electron, an acceptor. The accepter is neutral (neither positive nor negative). The doped crystal is called *p*-type, Figure 7 (c).

A diode, often called a PN junction, consists of putting together an *n*-type and a *p*-type crystals. The free electrons of the *n*-type can now diffuse to the other side, the *p*-type, which contains acceptors. Once an electron meets an acceptor, it sits in the hole, hence making the acceptor negatively charged. This diffusion of electrons continues until there is an equilibrium, where the positively charged donors in the *n*-type attract the electrons and the negatively charged acceptors in the *p*-type repel the electrons. Thus, creating a depletion region where no electron can exist. At this equilibrium point, the PN junction is an insulator, see figure below:

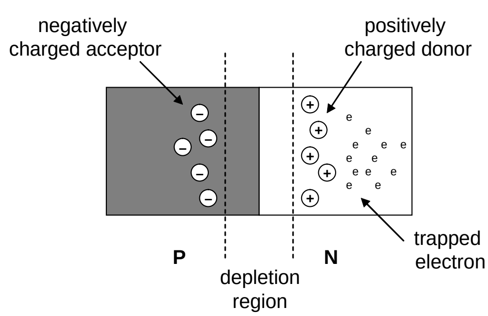

Figure 8: PN junction at equilibrium

By applying a potential in the opposite direction, i.e. positive at P and negative at N, it is possible to overcome this bias, allowing electrons to flow from

N to P. This makes the PN junction a conductor. Once the applied potential ceases to exists, the PN junction goes back to the equilibrium state where it is an insulator again.

Now we come back to those machines that were built using vacuum tubes. At Harvard, Howard Aiken followed Mark II by a pair of vacuum tube machines Mark III and Mark IV. Between 1938 and 1941, John Atanasoff and Clifford Berry of now Iowa State University constructed the Atanasoff-Berry Computer (ABC), an electronic computer using 300 vacuum tubes for solving systems of linear equations. Although it was not programmable, it was the first modern computer in many respects, most importantly the binary representation of numbers. Another machine called the Colossus was built under the direction of Tommy Flowers in England to decode German messages. In 1946, John Mauchy and Presper Eckert built the ENIAC (Electronic Numerical Integrator And Calculator) at the University of Pennsylvania. This is considered the first electronic general purpose computer. ENIAC used 18000 vacuum tubes and was 1000 times faster than its contemporaries. To program the ENIAC, however, meant to rewrite it, and it was done manually.

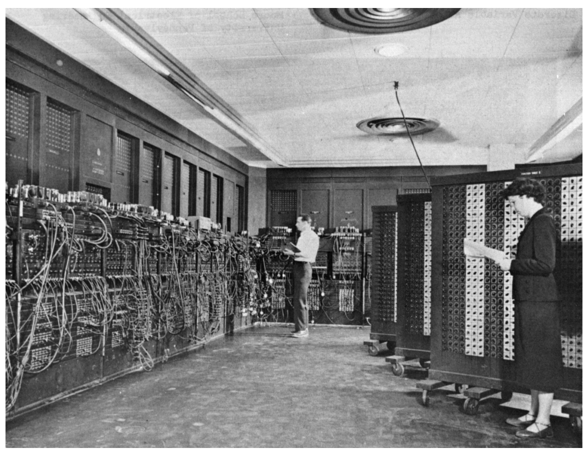

Figure 9: The ENIAC

Shortly after, Jon von Neumann, based on ideas developed by Eckert and Mauchly after recognizing the limitations of ENIAC, wrote a widely-circulated report describing a computer design (the EDVAC design) in which the programs and working data were both stored in a single, unified store. This basic design would serve as the basis for the development of the first really flexible, general purpose digital computers (although too much credit is given to von Neumann). Among the first of these machines were EDVAC (Mauchy and Eckert again) and IAS (Julian Bigelow) at Princeton's Institute of Advanced Studies. IAS was 10 times faster than ENIAC. The resulting paper discussed most of the architectural concepts seen in modern computers today. In the 1940s, first generation machines using vacuum tubes and drum memory (similar in concept to what is known today as hard disk) were produced. In late 1940s, diodes and transistors were invented. Therefore, the 1950s and the 1960s witnessed the second generation of machines with much smaller circuits and electronic memories (drum memory started to become the hard disk, the mass storage device). After that, with companies such as Intel (Ted Hoff), Apple (Steve Wozniak), and Microsoft (Bill Gates), the personal computer was born.

2 The modern computer

The early computers that we've mentioned in the previous sections had:

- either fixed programs (like the Pascaline, or special purpose computers)
- or their programs fed separately, by punched cards for instance

This makes it hard to program the computer to perform various tasks. For instance, it is impossible to use a calculator for text processing. One has to re-write, re-structure, and re-design the whole calculator (physical changes). Moreover, feeding the computer a program manually by means of a special machinery makes the process of programming very hard.

The idea of the stored-program computer (or von Neumann architecture, but we do not want to give too much credit to von Neumann) changed all that. By creating an instruction set (a set of instructions that the computer can perform) and detailing the computation as a series of instructions (the program), the machine becomes much more flexible. Furthermore, by treating those instructions in the same way as data, it is possible to store the program in memory, and hence a stored-program machine can easily change the program, and can do so under the control of another program!

Today most computers follow this architecture, were the program and data are stored in memory. Instructions are fetched from memory for execution by a Control until. The control unit decodes the instruction and determines what needs to be done. If for instance some arithmetic operations are to be performed, the control unit forwards the desired operands to the arithmetic and logic unit ALU. Together, they form the Central Processing Unit, CPU. The control unit contains a number of registers [modern form of Babbage's piles of disks :)]. These registers are used as place holders for the operands and the results of arithmetic or logical operations. Instructions can specify also that operands must be fetched from memory, in which case they are placed into a temporary register. Similarly, the ALU can also write the result to the memory, by copying it form a special register.

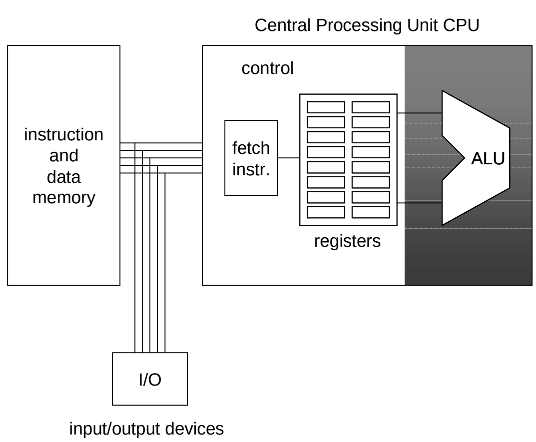

Figure 10: Computer architecture

3 Simplified time line

1200: The Abacus (in china).

1614: John Napier invented the Log line

1623: Wilhelm Schickard invented the first primitive mechanical calculator, first developed for clocks.

1642: Pascal's Pascaline.

1671: Leibniz's machine. Leibniz also described the binary system (base 2).

1801: Joseph-Marie Jacquard used punched cards to program a looming machine.

1820: Charles Xavier Thomas first to mass produced a mechanical calculator with addition, subtraction, multiplication, and division.

1833: Charles Babbage's Difference Engine.

1835: Charles Babbage's Analytical Engine.

1890: Herman Hollerith performed US census using punched cards, then started IBM.

1909: Percy Ludgate, an accountant from Dublin, designed a programmable mechanical computer similar to Babbage's (although unaware of the latter).

1930: The word "computer" was a job title assigned to people who used mechanical calculators to perform mathematical calculations.

1937: Claude Shannon produced his master thesis at MIT that implemented Boolean Algebra using electronic relays and switches "A Symbolic Analysis of Relay and Switching Circuits". This was the birth of practical digital circuit design.

1938: John Atanasoff and Clifford Berry of now Iowa State University built the Atanasoff-Berry Computer (ABC), an electronic computer using 300 vacuum tubes for solving systems of linear equations. Although it was not programmable, it was the first modern computer in many respect e.g. binary representation of numbers.

1940: George Stibitz at Bell Labs completed his Complex Number Calculator, an electro mechanical relay based computer using the binary system.

1941: A team directed by Tommy Flowers in England worked on the Colossus, a computer to decode the German messages.

1944: Howard Aiken from Harvard university and IBM produced Mark I, a general purpose computer also based on electro mechanical relays. However, influenced by the Analytical Engine, it used the decimal system. It was also programmable by punched cards. He followed this by Mark II, Mark III, and mark IV.

1945: John Mauchy and Presper Eckert built the ENIAC (Electronic Numerical Integrator And Calculator) at the University of Pennsylvania.

1946: Jon von Neumann, wrote a widely-circulated report describing a computer design (the EDVAC design) in which the programs and working data were both stored in a single, unified store. This basic design would serve as the basis for the development of the first really flexible, general purpose digital computers.

1940s: First generation machines: Vacuum tubes and Drum memory (similar in concept to what is known today as hard disk).

1950 - 1960: Second generation machines: Invention of diodes and transistors in late 1940s made it possible to construct much smaller circuits and electronic memories.

1960 - present: Third generation and beyond: Intel (Ted Hoff), Apple (Steve Wozniak), Microsoft (Bill Gates), and personal computers.