PRINCIPLES AND OPERATION OF ELECTRONIC DIGITAL COMPUTER

By FRANK HIRST, Ph.D., M.Sc.

1. Introduction

A digital computer is a machine which performs arithmetic calculations using discrete digits or numbers arrayed in registers, for example the cash register. Here the smallest discrete entity would be the halfpenny. More generally, digital machines for arithmetic purposes are arranged with registers for performing addition, subtraction, multiplication, division, and so on. The precision of such instruments is defined merely by register capacity, that is the number of digits present in the registers. For example, in a desk calculator, very often the keyboard register has a capacity of 10 decimal digits but the latter could be increased if desired.

In contrast to the digital computer, the analogue computer uses a physical quantity to represent a number. A simple example is the slide rule. With this instrument an analogy is made between the logarithms of numbers and lengths along the rule. Again large analogue machines represent numbers by means of voltage levels. The precision of a recorded number is limited by the accuracy with which the voltage can be measured (of the order of 1 part in 10^4).

2. History of the Electronic Digital Computer

The first calculating machine capable of performing arithmetical operations was constructed by Pascal in 1642. This machine could add and subtract, and was applied successfully to tax computations in France.

Leibnitz in 1694 constructed an extended version of Pascal's machine which had the facilities of multiplication and division. In 1820 Thomas de Colmar constructed his arithmometer but the first really practical machine was produced by Ohdner in 1891.

The first machine to undertake a sequence of operations was the difference engine, designed by Charles Babbage in 1822, which was able to compute a polynomial of the 2nd degree to 8 places of decimals using 2nd order differences. Such a machine requires one more register than the order of the differences. Thus x^2 can be computed by commencing with the registers set in the following manner:

1st register	2nd register	3rd register
1	1	2

The first operation of the sequence would be to add the contents of the 3rd register to the 2nd and the new contents of the 2nd to the first and then repeat the sequence. Under these circumstances, x^2 is generated in the first register as shown in the following table.

1st register	2nd register	3rd register
1	1	2
4	3	2
9	5	2
16	7	2

It is seen that the 2nd register holds the first order differences and the 3rd register the 2nd order differences which remain constant.

In 1823 Babbage proposed a difference engine to work to 20 decimal places and 7th order differences but this was never completed. Scheutz in Sweden constructed a machine working with 4th order differences and 14 places of decimals. A copy of this machine was made for the Registrar-General's department in England for sub-tabulation purposes.

Hollerith in 1889 developed a punch card machine for data sorting in the American census. From this developed the complex calculating and tabulating punch card machinery that we know today.

The large scale digital computer was first conceived by Charles Babbage in 1833. Following the invention of the difference engine, Babbage considered an Analytical Engine which would be capable of performing any computation whatsoever. This was to have been a mechanical machine with (a) storing registers or memory, (b) a mill or arithmetic unit, (c) a device for initiating the operations of the machine, including the transfer of numbers from one part of the machine to another and in their correct sequence (the control unit), (d) a system of punch card input, and (e) printed or punch card output. The Analytical Engine was never completed; however in 1937, 70 years after Babbage's death, Howard Aitken of Harvard University commenced the construction of a mechanical computer with the help of the International Business Machine Company. This machine, called the Automatic Sequence Controlled Calculator Mk. I was completed in 1944.

A Mk. II machine was started in 1945, and in this model operations were performed by means of electromagnetic relays allowing a considerable increase in speed of operation. The storage capacity was 100 numbers each of 10 decimal digits and sign. Addition and subtraction required 200 msec, whilst multiplication required 700 msec. In 1944 the Bell Telephone Laboratories produced a general purpose relay operated computer using 9000 telephone relays. Many other machines of this type have been built, but in general the relay computer has been superseded by the faster, electronic computer.

The first electronic computer was developed to compute ballistic tables during the last war, being sponsored by the U.S. War Department. This machine, a special purpose computer, was named the ENIAC (Electronic Numerical Integrator and Calculator) and was designed by Eckert and Mauchley at the Moore School of Engineering in Philadelphia. On completion it was transferred to the Ballistic Research Laboratory, Aberdeen, Maryland in 1946. With ENIAC addition and subtraction required 200 μsec and multiplication 2.3 msec. There was provision for 10 decimal digits and sign in the storing registers which were constructed from rings of flip-flops. The latter accounted for the large number of valves (viz.18000) used in its construction.

Whilst ENIAC was under construction von Neumann et al. studied the problem of the optimum design of a general purpose electronic computer. They showed that such a machine should have a storage capacity of 1000 words and with this store both operating instructions and numbers could be retained in the same unit. To do this, a command must denote the address of the datum to be operated upon, and also there must be some provision for deciding the next command to be performed. Furthermore, since the command would be in the store, it could be operated on arithmetically thereby allowing great flexibility in programming the machine.

To obtain a storage of 1000 words presented a considerable problem since it was necessary also for the information to be read from and into the memory at electronic speeds.

At the time the only device which seemed capable of such operation was the mercury acoustic delay line. The proposals of von Neumann et al. led to the design of a general purpose computer the EDVAC (the Electronic Discrete Variable Automatic Computer) which was commenced in 1945. Two machines were commenced also in England. The first built at Cambridge (1946-1949), was called the EDSAC (Electronic Delayed Storage Automatic Computer). This machine has a storage capacity of 512 words of 34 binary digits. Addition and subtraction requires 1½ msec and multiplication 4 msec. The second machine, called the ACE (Automatic Computing Engine) was constructed at the National Physical Laboratory, London, being completed in 1950. The storage capacity of this machine was 512 words of 32 binary digits. Addition and subtraction required 32 μsec and multiplication 1 msec.

In 1949 the CSIRAC machine, which will be referred to subsequently, was commenced in the Radiophysics Laboratory of the Commonwealth Scientific and Industrial Research Organisation. It has a register capacity of 20 binary digits and addition and subtraction require 1 msec and multiplication 4 msec. It was transferred to the University of Melbourne in 1955 from Sydney.

Since these early machines there have been many much faster machines built and command performance times now are of the order of microseconds.

In conclusion we can define several rather arbitrary levels of machines:

(1) Single shot - Desk machine operator controlled.
(2) Sequenced - Preset sequence such as division on Marchant.
(3) Flexible Sequence - Plugboard machine preset for problem.
(4) Large scale fully automatic electronic. - Commands subject to arithmetic transformation by other commands during a computation

3. Analogy Between Desk Machine and Electronic Computer

The electronic computer consists essentially of five parts:

(1) Input
(2) Memory
(3) Arithmetic Unit
(4) Control Unit
(5) Output.

By analogy, we can consider that when a computation is carried out on a desk machine five similar facilities are used. With this instrument we input numbers on a keyboard register, we then (say) press the add key and the arithmetic unit operates and a number appears in the middle register. If we wish to use this number subsequently, it is written on a sheet of paper which represents our memory device. Again, if this number were the end point of the computation the middle register would give the output. The person operating the machine represents the control unit.

Consider addition and subtraction on a desk machine assuming a number in the middle register and one set on the keyboard register. When we press the add key, we can imagine that an add gate is opened and numbers flow from the keyboard (source) register and are added into the middle (destination) register. Similarly for subtraction, the subtract key is actuated and this would open a different gate so that numbers would flow from the keyboard (source) register and

subtract from the number in the middle (destination) register.

In the case of the desk computer, when an operation is not being performed, the digits in the register remain static. However with the electronic computer the digits are no longer engraved on mechanical wheels but are represented by electrical pulses.

A number is in fact represented by a train of electrical pulses equally spaced in time (see fig. 1).

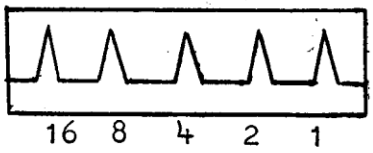

Fig. 1

Generally the pulses represent a binary number - the above train of 5 binary digits represents:

$16+8+4+2+1 = 1x2^4+1x2^3+1x2^2+1x2^1+1x2^0$ = the decimal number 31 and can be written 11111. Moreover a digit train may have some digits absent as in fig. 2 which represents:

$10110 = 1x2^4 + 0x2^3 + 1x2^2 + 1x2^1 + 0x2^0$ = the decimal number 22.

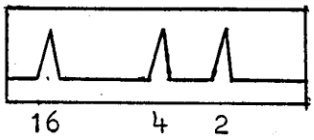

Fig. 2

Thus the presence of an electrical pulse represents a certain power of two, the least significant as usual being to the right in the above diagrams, whilst the absence of an electrical pulse means that the certain power of two is absent.

Very often a register of digits is contained in a delay unit which is connected to form a closed circuit by a wire, as shown in fig. 3. The pulses spend a negligible amount of time circulating in the wire and by looking at the contents of the delay unit at the correct time in any cycle the number in the register can be inspected and furthermore it remains there circulating until it is changed.

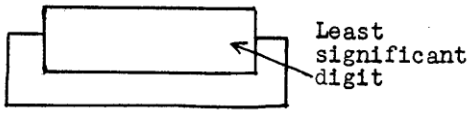

Fig. 3

Consider the transfer of a digit train from one register (B) to another (A) as shown in fig. 4.

Each register has one or more source gates and one or more destination gates associated with it. To transfer a number from B to A overwriting what is in A, we open source gate "B" which is numbered 11 and destination gate "Transfer to A" numbered 4. A copy of what is circulating in B is transferred at the correct time along wire D.T. (digit trunk) to A

.

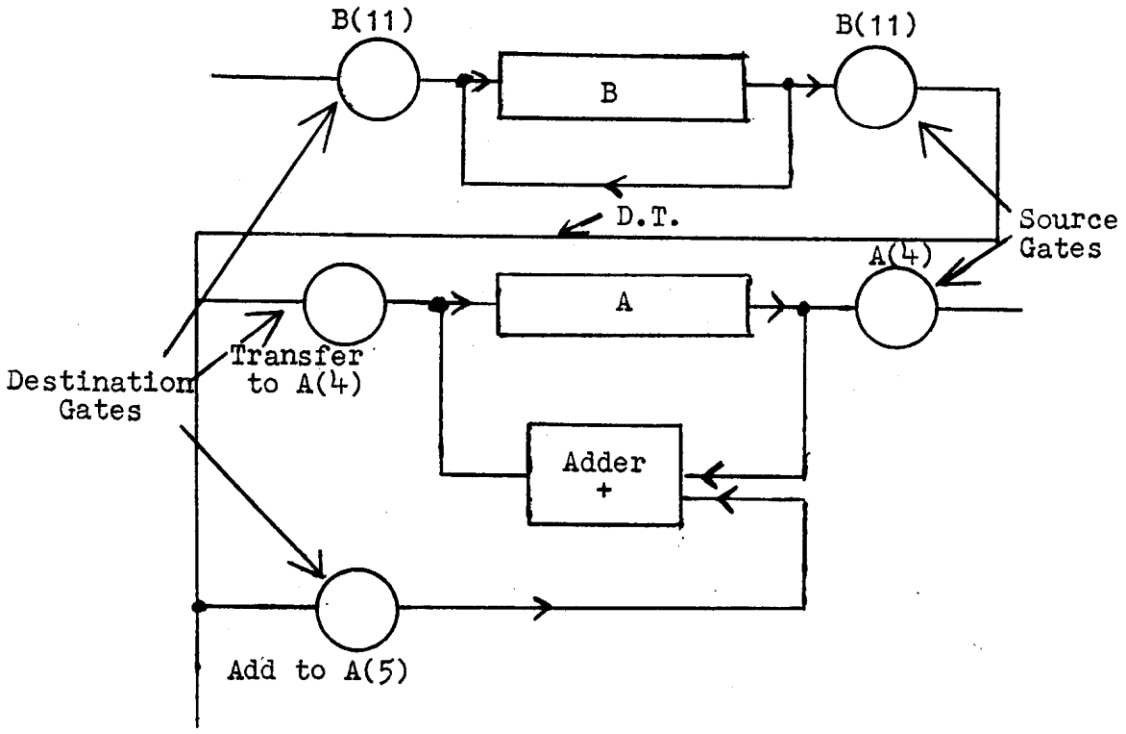

Fig. 4

If it is necessary to add the contents of B to A, the source gate 11 and destination gate 5 are opened so that a copy of the digits from B proceeds along the digit trunk and through the +A gate leading to an electronic adding device (+) which combines the incoming digits from B with those circulating in A in such a manner that the sum is formed which now circulates in A.

Thus to perform an operation, that is, carry out a command with an electronic machine, we merely open the appropriate source and destination gates so that the number flows from source to destination. In general there will be several registers in the machine with associated gates. The command or instruction to be performed defines which source gate and destination gate is to be opened to execute an operation. Many operations are required for a computation, so we arrange to do a programme of commands.

In particular the memory source gate leads from the memory or memory destination leads to the memory. The latter is a multiplex register which may have (say) 1024 cells (see fig. 5).

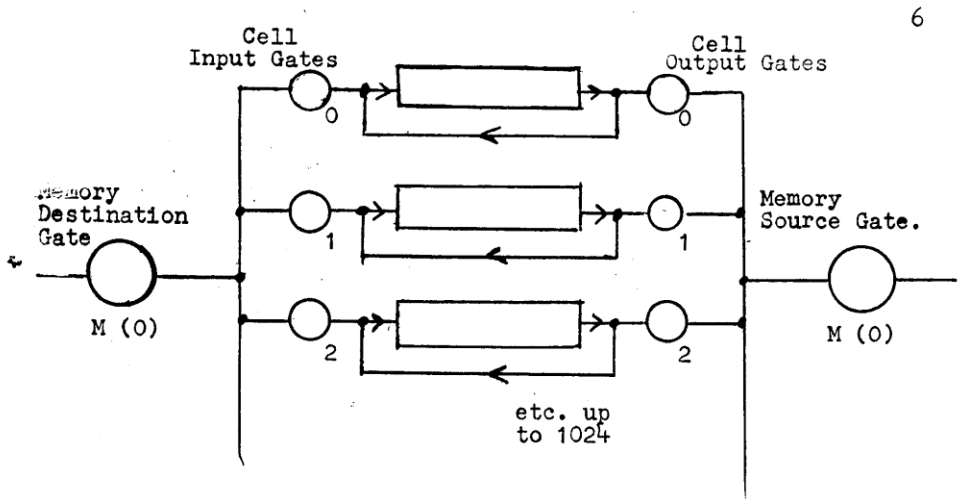

Fig. 5

If we are transferring to a certain memory cell or copying from a certain memory cell we must define not only the main memory source or destination gate, but also the individual cell gate (0,1 2, etc.). Thus a command must define also the individual cell gate to be opened and this constitutes the memory "address". It is to be noted that both the cell input gate and the output gate are opened simultaneously when the "address" in the command defines the individual cell.

4. 0rganization of an Elementary Digital Computer

CSIRAC, as mentioned previously, has registers capable of storing 20 binary digits and a fast memory of 1024 cells. An array of 20 digits which can either represent a command or a number to be operated upon is designated a "word". The following discussion describes the principles of an elementary digital computer and the mode of operation follows along the lines of CSIRAC.

A programme of commands is taken from punched paper tape by means of an input reader and is stored in the memory. Starting from the beginning of a programme, the commands are performed in a particular sequence. The 20 binary digit commands represent sufficient information for the computer to open the correct gates to perform the calculation, the results of which may be printed out by a teleprinter output or punched on to paper tape by an output punch.

Each command is made up of three parts (see fig. 6).

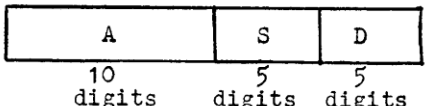

Fig. 6

The top 10 digits (A) represent the address and with 10 binary digits 1024 cells can be specified ($2^{10} = 1024$) from 00000 00000 to 11111, 11111 (i.e. 0 - 1023). For example, if we wish to transfer to or from memory cell 23 the 10 address digits would be 0000010111.

The 5 digits, S in fig. 6, define one of 32 source gates (since $2^5 = 32$) from 00000 to 11111 (i.e. 0 - 31). The 5 digits, D in fig. 6, define similarly one of 32 destination gates.

Assuming that in cells 5 and 6 are stored data x and y respectively, consider the following programme of commands:

Cell number where stored		Written			Coded as Binary	Coded as Decimal
(0)	0	5	M	A	0000000101,00000,00100	5, 0, 4
(1)	0	6	M	PA	0000000110,00000,00101	6, 0, 5
(2)	0	7	A	M	0000000111,00100,00000	7, 4, 0

The first command (stored in cell 0) says, take the contents of memory cell 5 (viz. x) and transmit a copy to the A register. The second (stored in cell 1) says, take the contents of cell 6 (viz. y) and add into the A register which will then contain x + y. The last command (stored in cell 2) says, take the contents of A and store in cell 7, which then will contain the sum x + y, and the programme of commands has been completed.

We can understand the operations in more detail by reference to fig. 7.

The sequence register holds a number specifying the cell which contains the command to be performed. Since the first command to be performed is in cell 0, the sequence register is initially cleared to zero.

The computer performs a command by carrying out a 4 phase cycle being under the control of the sequence control unit. In phase I the sequence number (0) flows via a gate (open during phase I) to the cell selector which receives a pulse pattern representing 0 and so opens cell gates 0. The digits forming the first command can now flow in phase II through the output cell gate 0, the main memory source gate (open as indicated during phase II of the computer cycle) along the digit trunk as indicated by arrow and into the interpreter since the gate to the interpreter is open during phase II.

It is the function of the interpreter to interpret the command and arrange that the appropriate address gates, source gate and destination gate are opened.

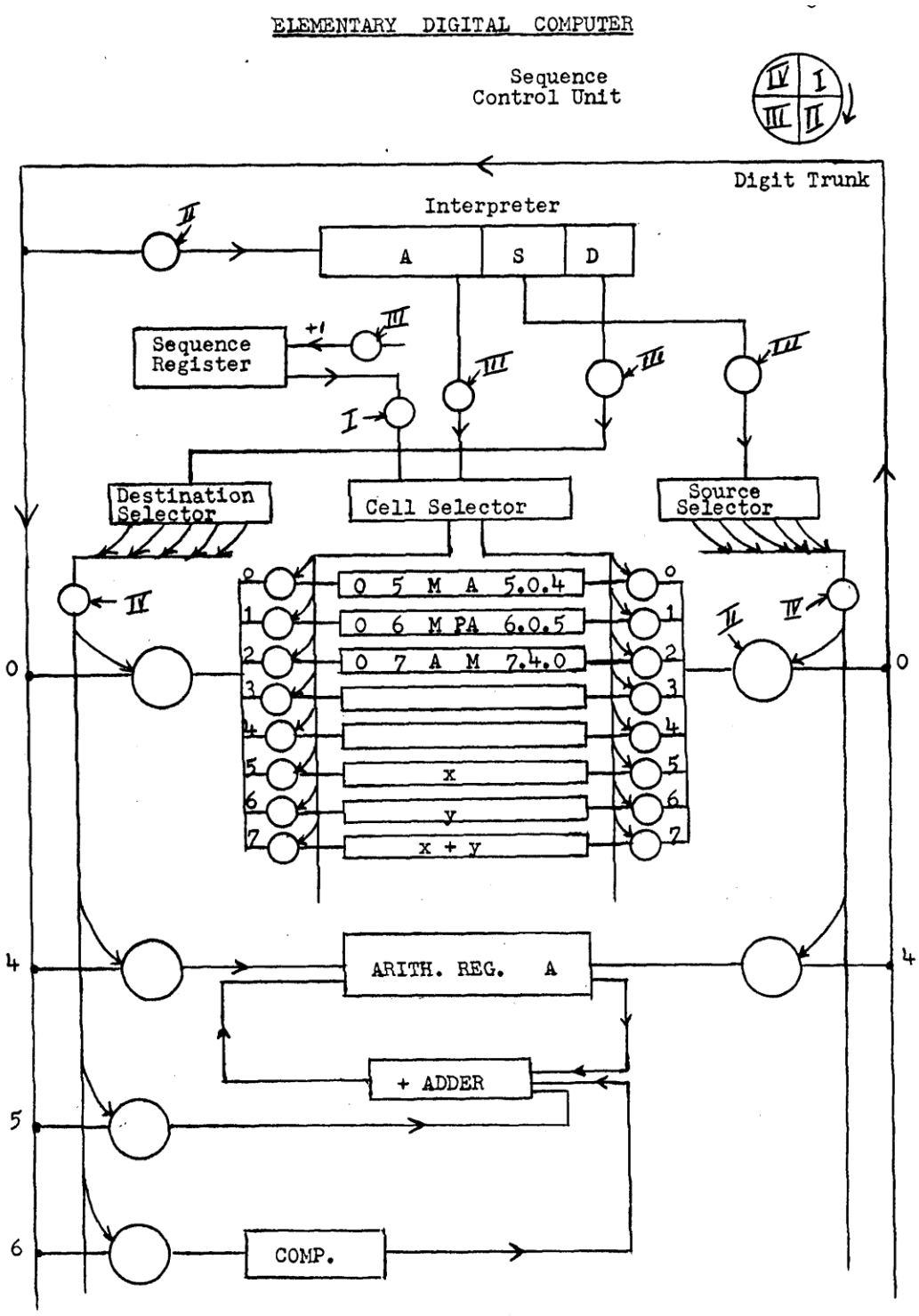

Fig. 7

It accomplishes this by partitioning the command into the 3 groups of digits A, S and D. Then in phase III, the A group is transmitted to the cell selector to select the appropriate address gate, the S group is sent to the source selector to select the proper source and the D group to the destination selector to select the proper destination. The 3 gate selectors are now all set up and in phase IV the gates are opened by the selectors so that the number flows from source to

destination. In the case of the first command the address is 5 so that the cell gates 5 are opened, together with the memory source gate 0 and the A register destination gate 4. Thus the number x flows out of cell 5 through source gate 0 along the digit trunk into the direction indicated and enters the A register via the destination gate 4.

During phase III the sequence number is increased by one so that when a new cycle is commenced, during phase I cell gates 1 are opened and the command in that cell is transmitted to the interpreter during phase II. This command is interpreted and the 3 selectors are set up in phase III and the transfer from source to destination takes place in phase IV. The command in cell 1 opens cell gate 6, source gate 0 and destination 5 so that digits forming y leave cell 6, flow out of the memory along the digit trunk through the destination gate 5 and enter A via the adder. Thus the digits forming y are added to those of x which are circulating in A.

Again 1 is added to the sequence register in phase III and the last command is selected. The cell gate 7, the source gate 4 and destination gate 0, are opened. Thus in phase IV, digits forming x+y leave A via the source gate 4 proceed along the digit trunk, enter the memory (since this is the one and only destination gate open) and proceed into cell 7, the sum x+y being retained there.

In this manner a programme which may consist of several hundred commands is performed one command after another. Because the pulse trains move at electronic speeds a programme can be performed very rapidly which enables the electronic machine to complete complicated calculations in a reasonable amount of time.

It is the task of the programmer with the 32 sources and destinations available, to encode a series of commands which will instruct the machine to perform the necessary computation.

5. Memory Systems

The CSIRAC machine uses a mercury delay line working memory consisting of 102~ cells and in practice 16 cells are accommodated in each loop, there being two loops per line. The digit pulses are spaced 3 μsec apart and therefore a 20 binary digit word is 60 μsec long. Since there are 16 words per loop the circulating time or "access" time of the memory is 960 μsec.

The machine is also equipped with a magnetic disc backing store capable of storing 4096 words. The average "access" time of the disc is 15 msec.

Other memory systems are as follows:

(1) Magnetostrictive delay line store.
(2) Electrostatic store.
(3) Magnetic matrix store.
(4) Magnetic tape store.

Of the above, the first three are used for fast access storage, the latter as a backing or auxiliary store.