THE UNIVERSITY OF NOTTINGHAM

FACULTY OF APPLIED SCIENCE

Applications of Computers

LECTURE 5

"BASIC PRINCIPLES - INPUT AND OUTPUT"

by

A. S. DOUGLAS, B.Sc., M.A., Ph.D., A.Inst.P.

"BASIC PRINCIPLES - INPUT AND OUTPUT"

INTRODUCTION

In order to load the storage of the machine initially it is necessary to have a device capable of preparing the storage medium or of translating from the medium on which information is prepared into the form of internal computer storage. Similarly the results obtained by the machine must be displayed or made available in some suitable manner.

Regarded from the point of view of the basic computing machine, the input and output mechanisms include any device for communicating with external equipment. From this point of view drum or magnetic tape stores are communicated with by input and output devices. However, it is usual to regard as 'input' devices designed to assimilate information fed to the machine from wholly extraneous sources, and as 'output' results which are not used again by the computer in the immediate future. In so far as it is possible to remove magnetic tape, for instance, from the computer and to attach it to apparatus for printing results, the tape is an output medium; since the information can be read back later it is also a form of backing storage. It is obvious that this duality is fundamental to input/output systems using the same medium for each purpose. However, it is intended to distinguish here the particular use of various media for input or output from their role as a form of storage, even though, in principle, no such distinction exists.

We shall, in practice, distinguish between equipment controlled directly by the computer and equipment operated independently of it. The former we shall designate 'on-line' equipment and the latter, if used in the assimilation of data or production of final results, as 'off-line'. By an 'off-line printer', therefore, is meant a printing apparatus

controlled independently of the computer, but capable of processing an output medium of the computer. By an 'on-line printer' we mean one attached to and controlled by the computer to display results. If the computer forms part of a system designed to control an operation taking place in a short space of time, we speak of a 'real time' application and of the computer as being 'on-line' itself.

As regards results we note that two distinct types of result are normally required. Firstly we may require a form of graphical output. This is taken to include contour mapping, the plotting of graphs and the preparation of nomograms. Secondly we may require the preparation of tabulated information in the form of symbols - numbers, letters or otherwise - an operation we shall call 'printing'. We note that it is possible to draw symbols with a graphical apparatus or, within rather coarse limits, plot graphs with printing apparatus. We shall describe in a later section the apparatus available for output, and give, in Appendix A, a brief comparison of the advantages and disadvantages of the principal media.

INPUT

The functions of the input to a computer may be divided between two applications. We may treat the patterns fed by the mechanism as being in the form suitable for direct storage elsewhere in the computer, or we may accept the patterns and perform a translation of them on the way in. The former situation normally arises when the input/output system is being treated as a backing store, since we may assume that the output was in a form readily reassimilated by the input. However, it can also arise where the preparation apparatus used to put the patterns on the input medium is itself capable of doing any necessary transformation of patterns. The case in which patterns must be inspected and treated before storage is common during true input, particularly of instructions, and where the form of number

representation in the computer is not commonly used outside it. The input is used in either or both of the two roles described above. The connections to the computer should, of course, reflect this.

From a programming point of view the translational role is most conveniently fulfilled if units of information, such as one decimal digit, are read in one at a time to a specified location either in the arithmetic unit or the working store. This implies that each such digit can be treated separately and transformed by the programme. Reading will not, in general, be slowed down by such a process, since the computer is usually fast enough to perform several instructions between each read operation. For pattern assimilation, however, such a system is clearly slower than need be, since whole blocks of information can be assimilated at a time and can be placed directly into backing stores. In this case connection would best be made direct with the backing store, or through the medium of suitable buffer storage. Two types of connection between input and computer are thus indicated, but, apart from magnetic tape systems, this is not always reflected in the actual connections made, these depending in part on the characteristics of the media used.

It has already been pointed out that the fundamental storage system of a computer is not usually in the form in which information is naturally fed to it. Thus it is usual for us to write numbers in the decimal scale and to use various radices depending on convenience and custom. The preparation of data is clearly most suitably done in this form and results will also normally be presented using such conventional notation. The instruments used to prepare tape coupled with the input arrangements to the machine should, therefore, be capable of acting as a translation system to put the data in the form acceptable for operation within the computer. In the same way, reconversion will be required during the output and presentation of results.

The most common, but by no means the only conversion needed, is from decimal to binary and vice-versa. This can be done by the preparation equipment in some instances, or can be done by the computer. If, however, the computer represents numbers in the binary-coded-decimal form, no such conversion is necessary. This internal system is possibly appropriate, therefore, when large amounts of conversion during input and output are otherwise expected. The fundamental speed of operation is slowed or made more costly, however, and conversion is only needed normally where elaborate calculations are to be done on the numbers put into the computer. No conversion is required if the 'number' put in is in fact only to be used as a label, since recognition of the label does not depend upon its numerical value, but upon its properties as a pattern of digits.

There are three principal media used for input to computers, namely punched cards, paper tape and magnetic tape. For engineering purposes it is also usual to provide direct input of words or parts of words by means of hand-operated switches. For programming purposes this input can be ignored unless some integration of the machine with human control is required, when the switches can be used to modify manually the course of a computation. If used, however, it is desirable that some record of the course chosen is printed out or otherwise retained. In some control applications, input may be desirable directly from the apparatus being controlled, and special provision may be made for this. We shall describe below each of the principal media in turn and also consider in outline various other possible methods of input.

PUNCHED CARDS

A punched card consists of a rectangle of cardboard in which holes are punched to indicate information. It is usual to distinguish the information

carried by the position of the holes punched, and the card is usually divided into 80 columns in any one of which a hole can be punched in one or at most two of twelve possible positions (corresponding to the ten decimal digits and two control positions). If two or more columns are linked together for the purpose of conveying information this group of columns is termed a 'field.'

It is usual for a sterling sum, for instance, to be recorded in three fields covering, say, 10 columns altogether; thus £57. 4. 2 might be recorded as 000057/04/02 and this would appear on the card as in Figure 1.

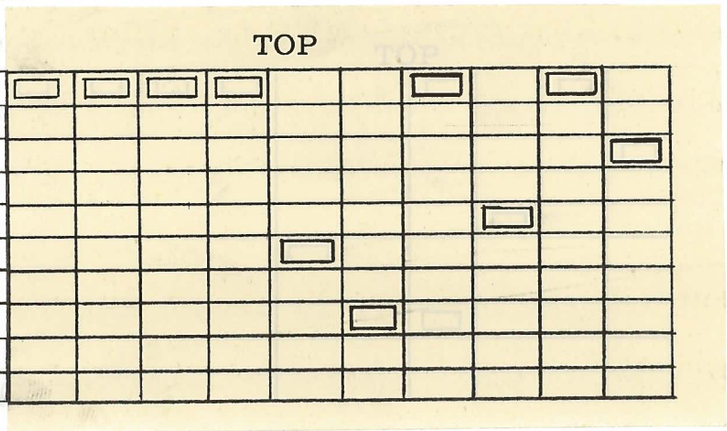

FIG. 1.

An alternative method of punching is to record information on the card as if each column or part of a column is a character (in the sense described in the previous lecture). Yet another alternative is to treat each row as representing a binary number, a hole representing a one, and no hole a zero. By either of these methods or variants of them much more information can be stored on each card. However, this entails removing a substantial amount of the card itself, which, in extreme cases, somewhat resembles a colander, and loses so much of its stiffness that jamming or tearing become a strong possibility in standard card handling apparatus.

It is usual to read the card from the top or bottom (See Fig. 1) by passing it under a set of brushes, one per column, which make electrical contact through holes, thus generating a pulse when a hole

passes them. An alternative reading system recently developed entails reading the card from the side photo-electrically. In this sytem a light is shone on the card face and the beam is stopped by the card but passes through a hole and is detected by an array of photocells (one per row) set below the card. Thus a pulse is generated when a hole passes and can be recorded in the machine storage. In this system a column at a time of holes is recorded, as a rule, in one location. In the electromechanical system a complete card image is formed in the store.

The speed of card reading is from 200 to 900 cards per minute, the former being the speed of standard apparatus used for business purposes. For a card punched in the usual way this means that the machine store is loaded at about 800 bits or 20 words per second. Regarded as a form of auxiliary storage, therefore, cards can be thought of as having access times of the order of 5×10^{-2} seconds. However, the decimal form of the information requires some translation before direct use and the effective access is thus slower. As auxiliary storage it is better to use cards punched in binary form; the reading speed is then three times as fast, and no translation is required. The most rapid access possible in this case is of the order of 4×10^{-3} seconds. Binary punching is not generally suitable for input, since the preparation of the card is made more difficult and this usually dissipates any advantage in reading speed which may accrue. However, if one pack of cards is to be used many times it may prove best to prepare the pack in binary either directly or by using the computer to effect the translation.

PAPER TAPE

The tape normally used for input is the standard Post Office teleprinter tape and consists of a continuous strip of stout paper of 1/2" width. Across the width of this there is room for five holes of standard dimension (1/16"

diameter) and one smaller sprocket hole used for drawing the tape forward mechanically. Other wider tapes capable of having 7 or more holes are in use also, particularly in the United States. The pattern of holes across the tape forms a 'character', and 32 possible such characters are distinguishable on standard tape. In preparing the tape a teleprinter punch is used which has a keyboard not dissimilar from a typewriter. The same character is punched on the tape for more than one key on the board, two different such combinations being distinguished by being preceded by a special symbol denoting letter or figure shift. Thus the character 00101 can represent the letter T or the number 5 depending on which of the two special symbols last preceded it. This duality is usually resolved by the system of input to the computer, which has also to assemble a series of decimal characters which may individually be represented by their binary equivalents into words which can be stored within the computer. The time occupied in this assembly is obviously much reduced in computers using binary coded decimal systems for internal storage.

About 10 characters per inch are stored on paper tape. The tape is read photoelectrically, as a rule, under friction drive, the sprocket hole being used to synchronise the reading system and not for mechanical movement of the tape. Reading speeds are up to 1000 characters a second, the most commonly used at present having a maximum speed of 200/ch. sec. These readers are normally capable of stopping between characters - a useful feature when reading in programmes.

Average reading speed affords an input speed comparable with that achieved by card apparatus used in its normal form, i.e. about 20 words per second, and fast readers working at 1000 or more characters per second give promise of better facilities than are now available. Furthermore, tape reading apparatus is simpler and less

bulky than its card reading counterpart. However, cards have advantages with regard to output and can be read more rapidly in binary form.

MAGNETIC TAPE

In the preceding lecture a physical description of magnetic tape apparatus was given and its use as a backing store described. In the role of an input medium to the computer there are at least two different ways of using it. The method which would appear most simple is to devise an instrument for preparing the tape in exactly the form in which it is used as a backing store. This is quite feasible, and is convenient to do for the input of numerical data in machines, the form of storage in which is binary coded decimal, for instance. However, if the machine stores numbers in fixed binary or in floating point form, or if the instructions are also to be fed in through the magnetic tape any preparation apparatus which is suitable will necessarily be complicated and expensive. The complication can be reduced only by arranging the storage on tape in a form less suited to its role as a backing store, thus slowing or complicating its use in that capacity.

As a general purpose input, therefore, magnetic tape is probably best used in the same way as paper tape, each character being read by a single operation of the computer into some convenient position, such as an accumulator or a part of the working store. The characters thus read in are assembled by the programme into words or otherwise interpreted or translated.

Under these circumstances it is desirable (though not essential) that the tape should be capable of being stopped between two characters. In practice this has proved difficult to achieve in the past, owing to problems in engineering. It has been necessary to move the tape at high speed, in order to ensure adequate signals in the reading system, and a method

of stopping between characters would thus require very low density of packing on the tape. This feature is undesirable both for convenience of preparation and storage and for economic reasons. The difficulties have, however, been recently largely overcome by the development of a new kind of reading head at Manchester University by Dr. Kilburn. Reading units are now under construction capable of reading tape packed at up to 100 characters to the inch at suitable speeds for stopping within the space of one or two characters. Practical operation will be at about 1000 characters/second, when the behaviour of the unit will be exactly analogous with that of a paper tape reader.

SOME OTHER INPUT POSSIBILITIES

In a few instances direct input of information from a variety of sources has been achieved. Apart from direct input from special apparatus, such as radar sets, telemetering devices and so forth, it is possible to connect manual devices, such as punches, directly to the computer or to feed their information in via buffer stores. Thus data can be fed directly to the machine. This will normally necessitate checking of the validity of the data received by the computer itself, which is not necessarily a simple problem. Even if corrupt information is acceptable for the purposes of the answer required, suitable allowances must be made for operator's error.

Apparatus which has been recently developed achieves the translation of printed documents into electrical pulses directly, and thus enables the computer to be fed with information from the printed page. Whilst this is satisfactory if the documents to be read are printed and are in good condition, corrupt reading can arise from heavily folded documents or badly defined letters or numbers.

In view of the possibility of corruption of information fed to the computer by direct communication, it is not usually advisable to employ this method of input. It is for this reason primarily that indirect methods via cards or tape are usually employed. Furthermore, as has been pointed out above, these media are suited themselves for use as storage available to extend that incorporated in the computer, and thus can serve an additional purpose. Whilst it might be possible to use printed paper in this role with the print scanner described above, manual methods such as direct typing are not suited to this.

OUTPUT

It is possible to control with a computer any wholly automatic piece of apparatus. However, in practice, direct control of complex apparatus is not widely used, and, in order to conserve computer time, it is usual to operate directly only rapid forms of presentation, such as cathode ray oscillographs or feeds onto media such as tape or cards which can later be processed off the computer to present results. In a recent article, Dr. Gill has discussed how direct control of slow apparatus might be achieved without wasting computer time, but little practical work on these lines has yet been attempted. It is to be hoped that experiment will proceed, particularly with the control of print setting apparatus, such as the linotype.

For 'off-line' operation there are once again the three principal media, cards, paper tape and magnetic tape, the physical form being unchanged from previous descriptions, but the speed characteristics being substantially altered. On-line apparatus includes the teleprinter, the line printer, the cathode ray oscilloscope, and the mechanical plotter, the two former being used for tabulating numbers (or printing letters) and the latter two for plotting graphs. However, as we pointed out above, it

is possible to plot a crude graph using a printer, or to output numbers or letters using a plotter by drawing them on the plotting surface.

PLOTTERS

Mechanical plotters are of two types, those with a fixed plot and those with a continuously moving plot. In the former case marks are made on the plot by moving a marking arm or arms in two dimensions. In these circumstances it is necessary to supply the two co-ordinates of each position to be plotted, and a graph or contour is drawn by supplying successive co-ordinates of points lying on the line being plotted. Lines can be made to appear continuous by plotting points sufficiently closely. The accuracy achievable in positioning the arm can be, at best, of the order of 1 part in 10,000. This apparatus has the advantage that a series of graphs or contours can readily be plotted on a single sheet of paper, even when these cross frequently. Such a mechanical plotter driven by paper tape and suitable for plotting from computer output is available commercially. It is capable of plotting 2-3 points per second.

Continous plotters cause the graph paper to move at a steady rate round fixed rollers. Parallel to the axes of the rollers is an arm on which are mounted one or more markers capable of moving along the arm and making marks at specified positions. Graphs are plotted by feeding to the apparatus the co-ordinates of successive points on the graph in a continuous stream starting from one end. This apparatus is particularly useful with analogue equipment, the output voltage of which constitutes the required co-ordinate. However, unless the paper is joined as a continuous band and passed under the marker many times or several arms are provided, it is not suitable for multiple graph drawing or for contour plotting. Plotters of this type are fairly

rapid in operation, the paper being moved at up to 2-3 inches/second.

Cathode ray tube output is usually operated in a way analogous to the fixed plot mechanical system. The beam is switched on to act as a marker, and its position is controlled by co-ordinates supplied from within the computer. A permanent record is usually made available by photographing the face of the tube. Such an arrangement suffers from four drawbacks. Firstly, the accuracy achievable is not high due to the diffuseness of the spot, secondly the area covered is not large relative to the spot size, thirdly, the definition of lines, especially those near the vertical or horizontal is poor, and fourthly, variations in intensity of spots is limited. A typical system can produce spots at the points of a grid 1024 x 1024 with side about 1/4 of a spot diameter. Accuracy is thus about 1 part in 1000. Overlap of the spots, implicit in this system, allows good lines to be drawn horizontally, vertically or at 45^{o}, but lines within 5^{o} either side of these directions are very poorly represented. Two to three intensity levels are recognisable (depending on how many times a spot is repeatedly marked), but this is not satisfactory for contouring or multiple graphing as a rule. Nevertheless for rough graphical work this is a useful instrument and gives very rapid output, each point being plotted in a few milliseconds. By drawing numbers on the face of the tube, each being defined by a matrix of about 16 x 8 spots, an output rate equivalent to or faster than punched cards and considerably faster than paper tape can be obtained.

PAPER TAPE

Standard punching apparatus for paper tape operates, at present, at 25-30 characters per second. Apparatus available in the United States includes punches working at 60 characters/second, and also at 240 ch/sec. Equipment under development in the United Kingdom includes a punch working at over 300 ch/sec. It is not thought that speeds higher than this

can readily be obtained with this medium in the near future.

The tape can be used to feed various types of output equipment such as the plotter mentioned above. However, the most commonly used is the Post Office type teleprinter. The standard model has a travelling carriage similar to that of a typewriter, and a similar selection of printing symbols (with minor differences). Speed of printing is 6-10 ch/sec.; up to 66 symbols can be printed on each line. A newer model with a fixed carriage and travelling printing head will operate at 12 ch/sec.; but is otherwise somewhat similar. A faster printer is under development. Apart from its slow speed the standard teleprinter has two other drawbacks. These are that all layout of tabulated information must be specified on the paper tape drive, and that the moving carriage makes the feeding of continuous strip stationery awkward. The latter difficulty is overcome on the newer model, but the former remains and reduces the effective speed of the apparatus for output purposes, unless it is possible to tabulate all information in a single column or a compact set of closely spaced columns.

PUNCHED CARDS

Standard punched card output speed is 100 cards per minute but apparatus exists working up to a speed of 600 cards/minute. However, fast output speeds are sometimes misleading to the extent that, in some cases, less than a full 80 columns can be punched at that speed. It is clear that even the slowest card punching speed is considerably more rapid than standard paper tape output. The advantage in speed is emphasised also by the printing apparatus available, since this does not demand the inclusion of layout characters on the cards.

Printing from cards is possible through standard apparatus such as the tabulator, or through newly developed fast line printers. All

tabulators have facilities for controlling the layout of output through a plug-board. The number of symbols which can be printed across a sheet of paper exceeds 70 as a rule, and a separate type bar containing a full range of symbols is provided to each possible position. Type bars are selected through the plugboard and are set by the symbols sent to them from the card. On some tabulators it is also possible to process the information before setting the type bars. Printing is carried out a line at a time, the paper being moved into position after the bars are set up, and the system reset after each line. Speeds vary on United Kingdom machines from 70 up to 300 lines per minute. Faster printers are under development using various different techniques. A wide range of equipment is, of course, available for handling cards output from the machine and for arranging paper feeds through the line printer.

MAGNETIC TAPE

It is possible to print information directly from magnetic tape, and at least one British apparatus to do this is in operation in the United Kingdom. Further apparatus is under development. The systems employed involve driving a line printer of the type suitable for card operation from magnetic tape. This necessitates a form of buffer storage being available owing to the reading characteristics of the tape described above under backing storage. The situation should be improved by the introduction of the new heads, but some buffer storage will still be required. Since the tape has only to be moved from the computer to the off-line printing system (or switched to this, if the printer is computer-controlled), the speed of output may be taken to be the speed of writing on the tape - i.e. about 10,000 ch./sec. However, in order to simplify the printing system some preparation of the information into character form may be necessary and this reduces the effective output speed slightly. Clearly, however, this is

the fastest form of output available.

Using the new type of reading head it is possible to allow the tape reading speed to follow the speed of the printing mechanism, and it is quite feasible to treat the magnetic tape output in the same way as paper tape, driving a teleprinter from it. Whilst the final printing speed is slower than can be achieved with a line printer, this is useful for small amounts of output, and may even be economic for large amounts due to the relatively low cost of teleprinters. Operated as an off-line system any desired speed of printing can, of course, be obtained by using a sufficient batch of teleprinters and making suitable arrangements to feed them.

APPENDIX 'A'

COMPARISON OF CHARACTERISTICS

The advantages of cards can be said to comprise the following:-

(1) Speed of output (and input when used as a backing store).

(2) Physical durability.

(3) Flexibility in use (possible to sort or rearrange readily).

(4) Availability of equipment for handling and printing.

The disadvantages are:-

(1) Awkward to store - bulky and liable to rapid deterioration in humid atmosphere.

(2) Vulnerability to getting out of sequence.

(3) Expense.

(4) Expense of associated equipment.

The advantages of tape are:-

(1) Compactness of storage.

(2) Speed of input.

(3) Sequencing assured.

(4) Cheapness.

(5) Cheapness of associated equipment.

The disadvantages are:-

(1) Speed of output.

(2) Lack of flexibility (cannot be sorted, etc.).

(3) Lack of physical durability (not serious, however).

(4) Temporary lack of fast printing and suitable preparation equipment.

Magnetic tape has the following advantages:-

(1) Speed of input and output.

(2) Physical durability (relatively good).

(3) Sequencing assured.

(4) Reusable.

(5) Compactness of storage.

The disadvantages are:-

(1) Expense (but less than cards since reusable).

(2) Expense of and temporary lack of associated equipment.

(3) Storage may require refrigeration.

(4) Lack of flexibility (sorting, etc.).