Christopher Strachey

The "Thinking" Machine

WHEN people discover that my job is concerned with electronic computers, there are two questions they usually ask. The insecure or politically-minded, and those who have read one of Norbert Wiener's books, usually want to know how long it will be before these machines take over their jobs. The others want to know how far machines can actually think, either because they want to marvel at them or because they need to be reassured that in fact no machine can ever really think in the way a human being does. Before we can discuss either of these questions without being actively misleading, it is essential to get a clear idea of what sort of things these computers can do and the way in which they do them.

Electronic computers by themselves are not capable of doing anything at all. They are merely devices which can carry out a long sequence of instructions very rapidly. These instructions, which are known as the program, must specify exactly what the machine is to do in every possible situation it can encounter, for the machine can do nothing which is not laid down for it in its program. Once the program has been written—and ultimately it must be written by a human being—the machine is able to apply the instructions in the program to the data of the problem in hand, even if the process seems fantastically complicated and laborious. It is rather as if all the thought required to solve the problem is concentrated into the program, while the machine is left to do the donkey work. It behaves as if it were a completely obedient and accurate slave who works with incredible speed, but wholly unintelligently.

It is important to realise that the machine can exercise no judgement at all on its own initiative, so that it is perfectly happy to put out quite ludicrous answers. If, for example, it is calculating weekly wages and there is a mistake of some sort in the figure it reads for the number of hours worked, the result may be an obviously absurd wage—such as fourpence or four million pounds. A human being getting this sort of answer would immediately suspect an error and would stop to investigate, but the machine will be quite content with any answer at all unless the program specially instructs it to reject answers which are obviously silly. And, moreover, in order to do this the program has to specify in detail what makes an answer obviously silly. In this particular case, the sort of test which might be included would be to verify that the wage lies between certain limits, say one pound to twenty pounds; another possible test for a "reasonable" wage is that it does not differ too much from last week's wage for the same man, but in this case a special provision would have to be made for the first time a man is paid, as there would otherwise be nothing to compare with.

It is impossible to overemphasise the importance of the program; without it a computer is like a typewriter without a typist or a piano without a pianist. With this in mind it becomes clear that all our questions about what a computer can do need to be rephrased. The proper question to ask is not "can a computer do this?", but "can we write a

program to make a computer do this?", and this is really a very different sort of question because the writing of programs (most unfortunately known as "programming" in the trade) is still an essentially human activity.

WRITING a program to make a computer solve a particular problem falls into two stages. The first and most difficult of these involves no technical knowledge of the functioning of a computer beyond the fact that it must have a completely detailed program. In this stage we have to produce a set of rules which we know will lead to the required answer in every possible case. The second stage consists of transforming these rules into a precise set of instructions in a form suitable for the computer. This stage is much less difficult (and much less interesting) than the first and it obviously involves a considerable technical knowledge of the machine to be used.

It might appear from all this that the need to have a detailed program meant that the machine could never produce any results which were unexpected or new, or indeed anything which was not rather obviously trivial. This however is not the case. One of the most interesting facts brought out by the attempts to make computers imitate human methods of thought is that a great deal of what is usually known as thinking can in fact be reduced to a relatively simple set of rules of the type which can be incorporated into a program. Simple games such as noughts and crosses present no difficulty and programs exist which make the computer invincible in such games. Chess, so far as I know, has never actually been programmed for a computer, but a program exists for the machine at Manchester University which will make it play quite a tolerable game of draughts, and the same machine has also written quite a number of unmistakable love letters. Several programs have been written which allow a computer to learn in a limited sense, and it should be quite possible to get one to do a simple type of intelligence test and to do fairly well at it.

Put baldly like this, with no explanation of the way in which they work, these programs can very easily give the impression that computers can "think." They are, of course, the most spectacular examples and ones which are easily understood by laymen. As a consequence they get much more publicity—and generally very inaccurate publicity at that—than perhaps they deserve. In fact they are unlike most of the problems put to computers as they are primarily non-mathematical in aim. They are, however, relevant to the problem of whether machines can think, so I shall describe two of them in outline to give some idea of the way in which they work.

IN SPITE of a certain impression of rather Victorian Babu, I think there is very little doubt of the intention of these letters:

Darling Sweetheart

You are my avid fellow feeling. My affection curiously clings to your passionate wish. My liking yearns for your heart. You are my wistful sympathy: my tender liking.

Yours beautifully
M. U. C.

Honey Dear

My sympathetic affection beautifully attracts your affectionate enthusiasm. You are my loving adoration: my breathless adoration. My fellow feeling breathlessly hopes for your dear eagerness. My lovesick adoration cherishes your avid ardour.

Yours wistfully
M. U. C.

The Manchester University Computer (hence the irreverent signature) can type out letters like this at the rate of about one a minute for hours without ever repeating itself. The scheme on which it works, however, is almost childishly simple. Apart from the beginning and the ending of the letters, there are only two basic types of sentence. The first is "*My*—— (adj.) —— (noun) —— (adv.) —— (verb) *your* —— (adj.) —— (noun)." There are lists of appropriate adjectives, nouns, adverbs, and verbs from which the blanks are filled in at random. There is also a further random choice as to whether or not the adjectives and adverb are included at all. The second type is simply "*You are my* —— (adj.) —— (noun)," and in

this case the adjective is always present. There is a random choice of which type of sentence is to be used, but if there are two consecutive sentences of the second type, the first ends with a colon (unfortunately the teleprinter of the computer had no comma) and the initial "*You are*" of the second is omitted. The letter starts with two words chosen from special lists; there are then five sentences of one of the two basic types, and the letter ends "*Yours* —— (adv.) M. U. C."

There are many obvious imperfections in this scheme (indeed very little thought went into its devising) and the fact that the vocabulary was largely based on Roget's Thesaurus lends a very peculiar flavour to the results. The chief point of interest, however, is not the obvious crudity of the scheme, nor even in ways in which it might be improved, but in the remarkable simplicity of the plan when compared with the diversity of the letters it produces.

It is clear that these letters are produced by a rather simple trick and that the computer is not really "thinking" at all. This is true of all programs which make the computer appear to think; on analysis they are nothing more than rather complicated tricks. However, sometimes these tricks can lead to quite unexpected and interesting results.

AN EXAMPLE of this occurred in the program for playing draughts. Although draughts is a relatively simple game, it has no complete mathematical analysis and there is no way of prescribing the best move in any given situation. It is quite simple to determine all the possible legal moves; the problem is to devise a way of choosing one of these which will allow the machine to play at least a tolerable game. The simplest possible method would be to choose a move at random, but this does not produce a realistic game at all, so a more sophisticated approach is necessary. In the scheme actually used the machine "looks ahead" for a few moves on each side. That is to say that it selects one of its own possible moves, discovers all the legal replies for its opponent, and tries them out one by one. For each combination of its own move and its opponent's reply, it then finds all its own possible second moves and so on. As there are, on an average, about ten legal moves at each stage of the game, the number of moves it has to consider gets very large indeed if it is to look ahead more than a very few steps. After a certain number of these stages (the number being determined by time taken) the machine evaluates the resulting position using a very simple scheme. It notes the value of this position and ultimately chooses the move which leads to the best possible position, always assuming that its opponent makes the best of the possible moves open to him.

The chief limitation of this scheme is the time it takes to examine all the possible moves. The machine can examine about ten moves a second, but the number of moves to be considered increases so rapidly that it is only possible to examine three stages in all (two of its own moves and one of its opponent's) if the process is to be completed in two or three minutes, and this is not enough to allow the machine to play a good game.

This basic move-finding scheme is embedded in a general game-playing program which accepts the opponent's moves, displays the positions, prints the moves, and generally organises the sequence of operations in the game. This program, though basically quite simple, gives a rather spectacular example of anthropomorphic behaviour on the part of a computer. For example it starts by printing "*Please read the instruction card.*" When its human opponent has assimilated the instructions, he presses a key and the machine prints "*Shall we toss for the first move? Will you spin a coin?*" It then calls, in a random manner, and asks "*Have I won?*" When it has been given this information (by means of another key) it either makes the first move itself, or waits for its opponent to do so.

Its opponent has to indicate his moves to the machine by setting certain keys, and if he makes a mistake in doing this, so that the move he indicates is meaningless, the machine will reject it and point out his mistake. If he makes too many mistakes of this kind the remarks printed by the machine will get increasingly

uncomplimentary and finally it will refuse to waste any more time with him.

There are various other situations in which the machine prints comments (for example if its opponent spends too long in making a move) and the whole behaviour of the machine, though perfectly determinate and simple if you know the rules, appears to the uninitiated as complex and unpredictable. All this, however, though fascinating to watch, is only a quite trivial form of trickery and has no real importance.

There is, however, one feature of the machine's game which was quite unexpected and rather interesting. The way in which the machine values the positions it reaches when looking ahead is extremely crude. It counts each man as one point and each king (being obviously worth more than an ordinary man) as three points; the value of any position is the difference between its own points and its opponent's points. A large number of the positions it examines will, of course, have the same value, and it chooses between these at random.

Suppose now its opponent has a man in the seventh rank so that he is about to make a king in his next move, and the machine is unable to stop him. The machine will effectively lose two points at its opponent's next move, and a human being would realise that this was inevitable and accept this fact. The machine, however, will notice that if it can sacrifice a single piece its opponent must take this at once. This leads to an immediate loss of only one point and, as it is not looking far enough ahead, the machine cannot see that it has not prevented its opponent from kinging but only postponed the evil day. At its next move it will be faced with the same difficulty, which it will try to solve in the same way, so that it will make every possible sacrifice of a single man before it accepts as inevitable the creation of an opponent's king.

This is obviously a fatal flaw in the strategy the machine is using. An opponent who detected this behaviour (and it is extremely conspicuous in play) would only have to leave his man on the point of kinging indefinitely for the machine to sacrifice all its remaining men as soon as the opportunity offered.

The interesting feature of the behaviour is that it was wholly unexpected. It is quite true that it follows inevitably as a logical consequence of the rules adopted, and that once it has appeared it is easily understood. But this is true of a great many discoveries in the realm of thought and at first sight it appears as if this is a case where the machine has made a genuine, if minor, discovery. But on closer inspection I do not think this is an accurate description. It is true that the machine demonstrated this unexpected behaviour, but it did not make the next and vital step of recognising that the behaviour was either unexpected or interesting. This helps to make clear both the possibilities and the limitations of the way we use computers at present.

It is entirely possible that machines can help us to discover a great many new and interesting ideas. They can work out the consequences of the assumptions built into their programs, and from time to time the results they get will be both unexpected and interesting. But just because they are unexpected we cannot make the computer recognise them when they turn up—for essentially the machine can only look for the sort of thing it has been told to look for.

One of the most difficult processes of thought is the recognition of a relationship between apparently unconnected events. It is a very difficult process even if the relationship is a fairly simple one and its general type is already known—indeed a large number of the questions in intelligence tests are of just this type. I know of no program for a computer which allows it to recognise any general relationship even of a limited type; it always seems to be necessary to specify exactly what is to be looked for. This means that it is wholly beyond us at present to write a program which will allow the computer to recognise an unexpected relationship—one whose existence we had not envisaged.

The problem of deciding if a relationship is interesting or not, once it has been discovered,

appears to present very little difficulty to a human being, but the mental steps we go through in reaching this sort of value judgement are most obscure. As a result there seems, at the moment, to be no prospect of being able to write a program to make a machine exercise judgement in this sort of way.

At present, then, computers can at the most only provide us with the raw material for new ideas. The final step of recognising the idea itself and realising that it is worth considering at all, has still to be carried out by a human being.

WE ARE now in a rather better position to consider whether or not computers can think. In the first place the question must clearly be recast in the form "Can we write a program to make a computer 'think'?" and in this form it becomes obvious that the answer depends almost entirely on the meaning we attach to the term "think."

A. M. Turing, who was discussing the mathematical limitations of these machines as early as 1936 has suggested an interesting way round the difficulty of defining this question. He points out that most people asking this question are in fact not interested in the answer. They really want to ask the rather different question, "Can a machine think as well as a human being?" He then describes an imitation game which might provide a satisfactory answer to this question. Basically the game is for the computer to try to make its human opponent believe that it is a human being. In order to remove irrelevant considerations, such as its shape, all communication with the machine is to be by a teleprinter. The player is then at liberty to ask the machine any question he likes and, after a suitable interval, a reply comes back over the teleprinter. The problem is for him to decide, from the answers, whether the teleprinter is being operated at the far end by a man or by a machine. The game might go as follows:

Question: Add 17 and 49.
Answer: 66.
Question: Multiply 314159 by itself.
Answer: (after about five minutes) 98696034011 (this is incorrect).
Question: Let's start a game of Chess. 1. P–K4.
Answer: P–K4.
Question: Write me a sonnet.
Answer: Come off it, old man, you know I can't do that sort of thing.
Question: What is your favourite food?
Answer: Strawberries and cream.

If the humanity of the operator was still in doubt after half an hour of this sort of question and answer, it would be impossible to deny that the machine was capable of imitating human thought processes to a very large extent.

No machine at present can play this game at all, but the type of program required is relatively simple (though immensely long) and Turing suggests that in the next ten years or so we shall see machines which can keep half of their opponents in doubt for five minutes. I think this may be a rather optimistic estimate, but I have no doubt that ultimately it will be possible to produce such a machine.

WHEN discussing the effect that computers and similar machines are likely to have on our lives, it is important to distinguish between those limited applications of existing techniques which we can foresee now, whose effects will be apparent in a few years, and the more remote and possibly more far-reaching results of future developments. The American mathematician Norbert Wiener has discussed the latter in two books—*Cybernetics* and *The Human Use of Human Beings*—and he draws an alarming, but I think probably justified picture of the social and economic upheaval that these are likely to produce. In effect he predicts a second Industrial Revolution, with changes at least as sweeping as those caused by the first. I do not propose to discuss this wider and more speculative aspect of the subject in this article, and shall confine myself to those applications of computers which I expect to see in the next few years.

There are two main fields in which these machines are used at present. The first, and the one for which they were originally designed, is in performing long and complicated mathematical calculations for scientific and engineering purposes. This is still probably the principal use of the machines and it is a very important

one. The effects it produces on the world at large, however, are only indirect. They take the form of bigger and better (or cheaper) aeroplanes, rockets, dams, buildings, and suchlike, or a better understanding of the complexities of nuclear physics or crystal structure. In this field the computer makes it possible to do things which could not be done without them, so that they actually provide more work for human beings.

Since almost all clerical work is done either partly or wholly by hand, and since most of it is either of a completely routine nature or only requires very simple judgements based on fixed rules, this is just the type of work for which it is possible to write a program for a computer. It is not surprising, therefore, to find that the second main field for these machines is in business and commerce.

The total volume of accountancy and clerical work done is quite startling. The 1951 census showed that 2,100,000 people in Great Britain were engaged in clerical work of some kind. This is about a tenth of the working population. A great many of these work in very large offices; the Ministry of National Insurance at Newcastle employs over 8,000 in a single building and there are several large firms with over 1,000 clerks. In a small office there is room for a certain amount of individuality and imagination, but if many of the 8,000 clerks in the Ministry of National Insurance started using their imagination they would wreck the whole system completely in a very short time. The operations of a large office have to be planned in as much detail as the mass production lines in a factory. This inevitably means that virtually no freedom of choice is left to the individual clerks, so that they are working more or less automatically according to a plan which has been laid down in advance. But this is precisely what a computer does except that a computer works several hundred times as fast, makes fewer mistakes, doesn't go on strike or stop for a cup of tea or a chat, and generally behaves in a much more satisfactory way. Mechanical and repetitive jobs are done much better and more economically by machines than by human beings—even if they are clerical ones.

There are, of course, a number of technical limitations on what a computer can do. There is an interesting difference between the sort of arithmetic done in business and commerce, and the sort of arithmetic done in science and mathematics. Scientific work, in general, involves very complex manipulation of a small amount of data. The machine reads in a few numbers and then proceeds to work away on them for a long time—several seconds, or possibly several minutes. This may not sound a long time, but these machines perform about one thousand operations a second, so that they are working about a thousand times as fast as a man using a small desk machine and perhaps fifty thousand times as fast as a man with only a pencil and paper.

Business arithmetic on the other hand consists almost entirely of doing very few and simple operations with a vast number of separate figures. The chief problem is to assemble all the material required for the calculation in hand, and it often costs more to bring all the relevant figures together than it does to perform the required calculation. This means that if a computer is to do office work efficiently, it must be able to absorb the figures it needs at a very high rate, and this in turn brings up the question of communication between the computer and the outside world.

The basic difficulty is that computers cannot read written or printed information, and, moreover, there is very little prospect of any machine being able to do this for a long time. The whole process of reading is fantastically complicated and we know distressingly little about it. For example we don't even know how the mind is able to recognise an individual letter, which can appear in literally hundreds of shapes and sizes. The whole process of learning to see is a very slow and difficult one, and the visual memory occupies a very large part of the brain.

In order to get information into a computer it is therefore essential to put it in some form that the computer can read. There are several possible forms (such as punched cards, punched paper tape, magnetic tape) each of which has various advantages and disadvantages. The one thing they have in common is that they

are not our usual method of writing, so that some form of translation is necessary; and this translation still has to be done by a human being.

It might seem that the fact that this translation is necessary nullifies most of the advantages of using a computer for office work, but in fact it is not so serious a disadvantage as it appears at first sight. The reason for this is that nearly all the writing which is done in offices is never used outside the office. Most of the figures that are written by one clerk are merely used at a later stage by another clerk, so that they form a series of intermediate results. If a computer is doing the work of the whole office all these intermediate results can be kept in the form suitable for the computer; there is no need ever to write them in ordinary figures or to translate them back again for the machine. Nearly all the files and records in the office can be treated in this way, and on the other rare occasions when it is essential to translate this information into ordinary writing, the translation can be done automatically by a quite simple machine.

It is therefore entirely technically possible for computers to take over a very large proportion of the work in offices, particularly in large ones. This is the first application of computers which will have an appreciable effect outside scientific laboratories. Already in this country there is one computer doing this sort of work, and in America the process is considerably further advanced. It will, of course, be several years before the full effect of computers is felt, but the process has started, and I think it is quite certain that in five or ten years' time a large proportion of office work will be done by machines.

There are, of course, a great many difficulties of a non-technical kind. The introduction of a computer of this sort involves a complete reorganisation of the office routine, and this is a slow and difficult process. Then again, to get the full benefit of mechanisation it is essential to standardise procedures to a far greater extent than is the case today. This sort of standardisation involves the cooperation of a large number of people with different interests, most of whom have more training in competition than cooperation, so that progress will probably be slow and difficult. But the economic advantages to be gained are so great that this sort of human cussedness can only delay the introduction of computers slightly.

I do not think that the introduction of computers into offices in this way will produce any very marked disturbances. It is true that a very large number of people will ultimately be replaced by these machines—possibly as many as a million in the next ten years—but the process will be a gradual one. Perhaps more important is the fact that there is, in any case, a very large turnover in the clerks they will replace. The majority of these are young women who leave after a few years to get married, so that the introduction of a computer will probably not throw very many people out of work. It will merely stop the intake of new clerks, who will presumably have to seek other occupations. What these will be, is an interesting matter for speculation.

Arnold J. Toynbee

The Elusiveness of History

HISTORY is elusive. The word itself is as elusive as the things that it means. The original Greek word "historia" meant "an inquiry." It could mean an inquiry about anything in the world, but it came to have the particular meaning of an inquiry into human affairs, and this in a limited usage of these two words. Human nature has a physical aspect, but the word history has never been used to mean the study of the human body. Sciences like anatomy, neurology, physiology, and biology have been excluded from the word "history's" empire. "History" has been restricted to meaning the study of the experiences and actions of human personalities. One might have thought that the word had now been pinned down, but it had not been, after all; for it had no sooner been confined within these limits than it burst its bounds by acquiring an alternative usage. It now came to mean human actions and experiences themselves, besides meaning the study of them, and it went on to widen its meaning still further.

One feature of human experiences and actions is that they are events on the move down a one-way stream of time. But human affairs are not the only things, known to human observers, that move through time along an irreversible course. So history came to mean all movement of this irreversible kind. The Earth's non-human fauna could have a history, its flora could have a history, the solar system could have a history, the whole cosmos could have a history; history need not be something that happened just to mankind or to individual human beings. A thing is in history when it is moving along a time-track on which it cannot turn back; and our modern men of science seem to think that most things in the universe move in this "historical" way, in contrast to Aristotle's view that all the heavenly bodies—from the moon inclusive, outwards, measuring from the Earth as the centre point—move in circular orbits in which each circumgyration is an exact repetition of every one that has preceded it.

So the word "history" has a whole gamut of meanings, and the two extremes seem far apart. At one end history means the study of human affairs, at the other end it means, not a study, but a movement which, so long as it is an irreversible movement down a time-stream, may be a movement of anything in the world.

IS there anything in common between the "subjective" history which is a historian's observation and record and the "objective" history which is the movement that the historian is trying to track down? Well, yes, there is. In the first place, "objective history" and "subjective history" are inseparable. Without an object there can be no inquiry, and, without an inquirer, there can be no object—or, at least, no object can be known to human minds except through some inquirer's observation of it. In the second place, the subject of an historical inquiry is, at the same time, part and parcel of the object that he is studying; for the historian himself, as well as the people or the things that he is observing, is afloat on the stream of time and is all the time being carried, like them, down time's irreversible current.