THE NEW SCIENTIST

27 NOVEMBER 1958

VOLCANO ON THE MOON?

by Zdeněk Kopal
Professor of Astronomy, Manchester University

First British Electronic Computer Exhibition

THE DEVELOPMENT AND USES OF DIGITAL COMPUTERS

by Nigel Calder

MACHINES WHICH LEARN

by Dr. A. M. Andrew
National Physical Laboratory, Teddington

A PSYCHOLOGICAL ANALYSIS OF TRANSPORT ACCIDENTS

by Dr. D. Russell Davis
Medical Psychology Laboratory, Cambridge University

ANTAGONISM AND TOLERANCE IN MICROBIAL BEHAVIOUR

by Dr. David Park

VOLUME 4 NUMBER 106 ONE SHILLING WEEKLY

THE NEW SCIENTIST

Cromwell House, Fulwood Place,
High Holborn, London, W.C.1
Telephone: HOLborn 7554.
Telegrams: Newscient, London.

Volume 4 Number 106 27 November 1958

CONTENTS

POSTAL SUBSCRIPTION RATES

Inland and Overseas: 60s. a year. 30s. for six months.
Special Rates. Canada: $10 a year.
U.S.A.: $12 a year.

Registered at the G.P.O. as a newspaper. Authorised as Second Class Mail for transmission by Canadian magazine post. Entered as Second Class Mail matter at New York Post Office.

Machines which learn

The first practical application of "learning machines" will probably be in the control of industrial processes. Experimental work is at a very early stage. Eventually a system may be produced which is a true analogue of some form of biological learning

by Dr. A. M. ANDREW

WHEN machines are required to perform complicated tasks it would often be useful to incorporate devices whose precise mode of operation is not specified initially, but which learn from experience how to do what is required. It would then be possible to produce machines to do jobs which have not been fully analysed because of their complexity. It seems likely that learning machines will play a part in such projects as the mechanical translation of languages and the automatic recognition of speech and of visual patterns. A considerable amount of work has been done by O. G. Selfridge in USA and by W. K. Taylor in London on the application of learning devices to pattern recognition. The suggestion has also been made in some quarters that machines with learning features might be used for economic planning and industrial management.

The application of learning machines which will be discussed at length here is one which may seem mundane by comparison, namely that to process control. Because this is a somewhat simpler application than the others it will probably be the first in which learning machines will prove practically useful.

It is often virtually impossible to analyse an industrial process completely from first principles, whether it is a chemical process or a physical one such as fractional distillation. Such a process would usually be put into operation under manual control, and rules for its manipulation would be developed in a trial-and-error fashion, as the operators gradually "got the feel" of it. A learning machine coupled to the process would similarly develop its control policy by trial and error.

In this type of application the learning machine would make automatic control possible without the need for the tedious mathematical analysis involved in designing a non-learning controller. The learning machine would be particularly advantageous when something about the process was likely to change gradually. It might happen, for instance, that a pipe became clogged or a catalyst exhausted. Under the new conditions a new policy of control might be required to give the best results, and a learning machine could adapt accordingly. There is, of course, a certain danger in this adaptability since it may cause a defect of the apparatus to escape detection when it ought to be remedied. A similar danger exists when a human operator controls something; for example, a driver of a car may not notice, as his brakes deteriorate, that he is continually having to press the brake pedal further down to get effective braking.

A learning machine used to control a process must be *goal-directed*. It will generally be possible to measure a number of variables associated with the process and to compute from these a quantity which may be taken as a measure of how well the process is working. The goal of the learning machine can then be to make this quantity as large as possible. For an industrial process, the quantity will depend mainly on the yield and quality of the product. By analogy with animal behaviour, much of which is directed towards the increase of pleasure, the quantity which is to be maximised has been termed "hedony."

Figure 1 represents a process controlled by a learning machine where the degree of goal-achievement or "hedony" is computed from measurements made on the product and on the raw material. The controller receives information from the process in the form of measures *a*, *b*, and *c*, which may be temperatures, pressures, flow-rates or other variables. It effects control by varying the quantities *d* and *e* which may control valve settings, or heating power applied to parts of the process.

It is unlikely that the measure of hedony, *h*, can be used directly to control the process as there will usually be a considerable delay between the occurrence of a variation in *d* or *e* and the appearance of the resultant change in *h*. In general the only way in which the process can be controlled is by letting *d* and *e* depend, according to some set of rules, on the measures of *a*, *b*, and *c*. This set of rules constitutes a control policy which is altered in a trial-and-error fashion so as to maximise *h*. At any instant, therefore, the effect of a high value of *h* is to help set the seal of approval on the control policy which was operative before it occurred, while the effect of a low value of *h* is the opposite. By such trial-and-error procedures the control policy can be made to approach an optimum form.

Any application of a learning controller to a process of any complexity will certainly involve the use of *sub-goals*. Part of the process will be linked to part of the controller in the general manner represented in Figure 1, but the quantity recognised as "hedony" and maximised by this part of the controller will not be the same as the overall goal of the complete controller. In controlling a distillation column, for instance, the overall goal would be expressed in terms of the yield and quality of the product, but a sub-goal might be to maintain a particular temperature distribution in the column, or to keep the operating conditions as steady as possible.

The sub-goals would probably be modified in a trial-and-error fashion by that part of the controller whose goal is the overall goal of the system. A complicated controller might actually involve a hierarchy of learning machines of which one was the "master," or first-order, controller. This would have, as its goal, the overall goal of the process, and it would set the goals for second-order controllers, some of which would set goals for third-order controllers, and so on. A human controller also makes use of sub-goals; a motorist wishing to keep his engine in good condition sets the sub-goal of keeping the oil-level between the maximum and minimum marks on the dip-stick.

Another property which must certainly be incorporated in a learning machine for process control is some form of *interpolation*. Suppose the machine has learned, for example, that a certain valve should be fully closed when a certain pressure rises to 10 lb/sq in and half-closed when the pressure is 5 lb/sq in. Then, when the pressure becomes 7 lb/sq in, the machine should be capable of deciding that the position of the valve should be somewhere between full- and half-closed.

Much of what has been written in the past about possible schemes for learning machines has not allowed for interpolation. The tendency has been to consider devices whose input and output information is conveyed through a number of connections, any one of which is at any instant either activated or non-activated. That is to say, the signal conveyed in any connection has two possible values. Devices of this kind are interesting as models of human thought processes since a two-valued signal can represent the logical values of "true" or "false."

The measures of temperature, pressure, etc., in a process vary over continuous ranges of values, and it is presumably necessary for efficient operation that the controller should take fairly detailed account of them. It is possible to modify a device which handles two-valued signals so that it accepts signals varying over continuous ranges. This can be done by allocating a sufficient number of two-valued channels to represent each continuous variable with the degree of accuracy required. However, the device would not then incorporate the feature of interpolation, and would need to learn independently how to deal with each situation which could arise. In order that a learning machine applied to process control may learn a suitable control policy in a reasonable time, it is necessary to have interpolation. This is most easily achieved by making the device capable of handling quantitative signals right from the start.

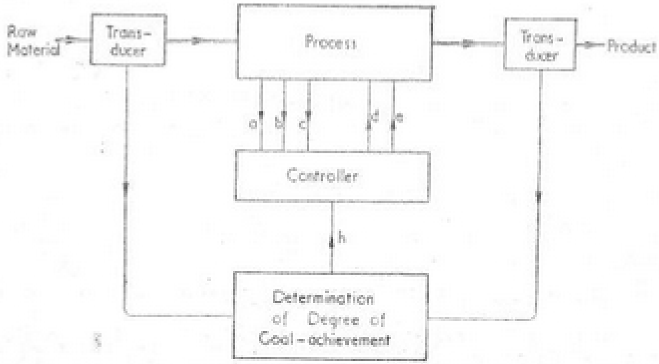

FIGURE 1. *A process controlled by a learning machine. The transducers are sensitive to the quantity and quality of the product and the raw material.*

Referring again to Figure 1, the sets of rules by which the values of *d* and *e* are computed from the measures *a*, *b* and *c* are continuously modified as the machine learns. Part of the learning process consists of finding the best values for "constants" which are incorporated in the sets of rules, but it is also possible to devise learning procedures which modify the actual form of the rules. It is possible that the measured quantities *a*, *b* and *c*, etc., could include some which, initially, are not known to be relevant to the control of any of the quantities *d*, *e*, etc. Then the learning machine may eventually decide that the sets of rules can advantageously be modified to make use of some of this information which initially had no effect on the control action.

If the controller makes use of sub-goals, the degree of achievement of a sub-goal (which might be termed a sub-hedony) is another quantity computed according to a set of rules from measures such as *a*, *b* and *c*. The sets of rules associated with sub-goals would be modified in the course of learning in the same way as were the sets of rules for computing the quantities *d*, *e*, etc.

The learning machine can exhibit flexibility in some other ways. Before arriving at a design for a machine of the general type considered up to now, it is necessary to make rather arbitrary decisions about certain features of it. One of these concerns the amount of its past history the machine will take into account when it decides whether to make a change in its policy of control. Another concerns the level of statistical significance the machine will require of its data before making a change. These decisions may be thought of as determining the "cautiousness" of the machine. It is also possible to assign a number which may be regarded as the "optimism" of the machine, since it determines the relative effectiveness of favourable experience (high hedony) and unfavourable experience on the machine's adaptation.

In designing a learning machine, the quantities determining "cautiousness" and "optimism" would have to be chosen rather arbitrarily, since the mathematical determination of the optimum values is virtually impossible, especially since they depend on the nature of the process to be controlled. The speed and effectiveness of the machine's learning, however, are greatly affected by the values chosen. In fact, the design of the learning machine is precisely the kind of awkward mathematical problem which it is hoped to by-pass by the use of learning machines. It seems likely, then, that there will eventually be no fixed values incorporated for the quantities which determine "cautiousness" and "optimism."

MACHINES WHICH LEARN *continued*

These will be varied in a trial-and-error fashion to find values which maximise the rate of increase of hedony. A machine can thus be visualised in which almost every aspect of the internal organisation is subject to adaptive modification.

Experimental work aimed at the development of a useful learning machine is at a very early stage. One of the most advanced examples which has been demonstrated to date is Uttley's conditional probability computer, illustrated in Figure 2. This device handles only two-valued signals and consequently appears to be unsuitable for process control applications. However, it illustrates important general principles and is a valuable basis for further developments.

The use of the term "learning machine" invites comparison with the learning of people and animals. There are certainly some important differences between the kind of learning discussed here and biological learning. It is, of course, impossible in principle to make a machine which will be universally accepted as learning, thinking, or showing intelligence, since as soon as a

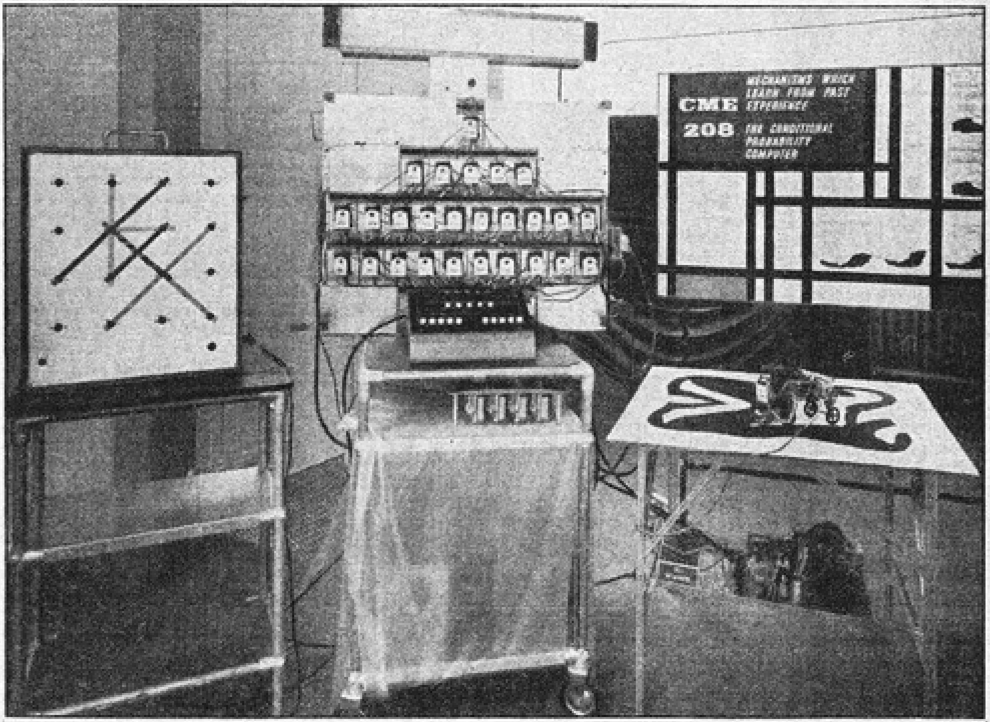

FIGURE 2. In the centre is Dr. Uttley's conditional probability computer. The white screen to the left has photocells behind each of the 16 circular holes and thus forms a simple artificial retina used to demonstrate an application of learning principles to pattern recognition. On the right is the Meccano trolley which is also shown in Figure 3.
The computer has five input channels. It consists of 31 similar units which accumulate statistical information concerning the patterns of activity presented in these channels. When some statistical data have been accumulated, and one or a group of the input channels is activated, the computer determines, on the basis of its stored statistical information, whether this group is usually accompanied by activity in another channel or channels. If it is, the computer makes an *inference* of activity in the other channels.
The 31 units are essentially counters; five of them count the occurrences of activity in the respective input channels, while another ten count the occurrences of simultaneous activity in the ten possible pairs of input channels, and so on for groups of 3, 4 and 5 channels. The counts are represented in the units by amounts of charge on capacitors.
The necessary condition for an inference is that the conditional probability of activity in a channel which is not directly activated exceeds a certain threshold. The relevant conditional probabilities are all computed as ratios of the counts stored on pairs of units of the computer.

machine exhibits some form of interesting behaviour there is an immediate tendency for people to say "It is not what we meant by thinking (or learning, or intelligence), anyway." However, although the kind of machine discussed here may not be a model of any kind of biological learning, it is capable of showing behaviour which would readily have been accepted as "learning" had it been done by an animal.

It has sometimes been suggested that the study of biological learning may be useful for the development of machines, and most workers who have devised learning machines have regarded them as possible models of biological learning. The drawing of analogies between brains and machines requires caution to say the least, but in a general way it is stimulating for workers in either field to know something of what is happening in the other, and it is possible that speculation about machines which learn may eventually produce a system which is a true analogue of some form of biological learning.

I am indebted to the Director of the National Physical Laboratory for permission to publish this article.

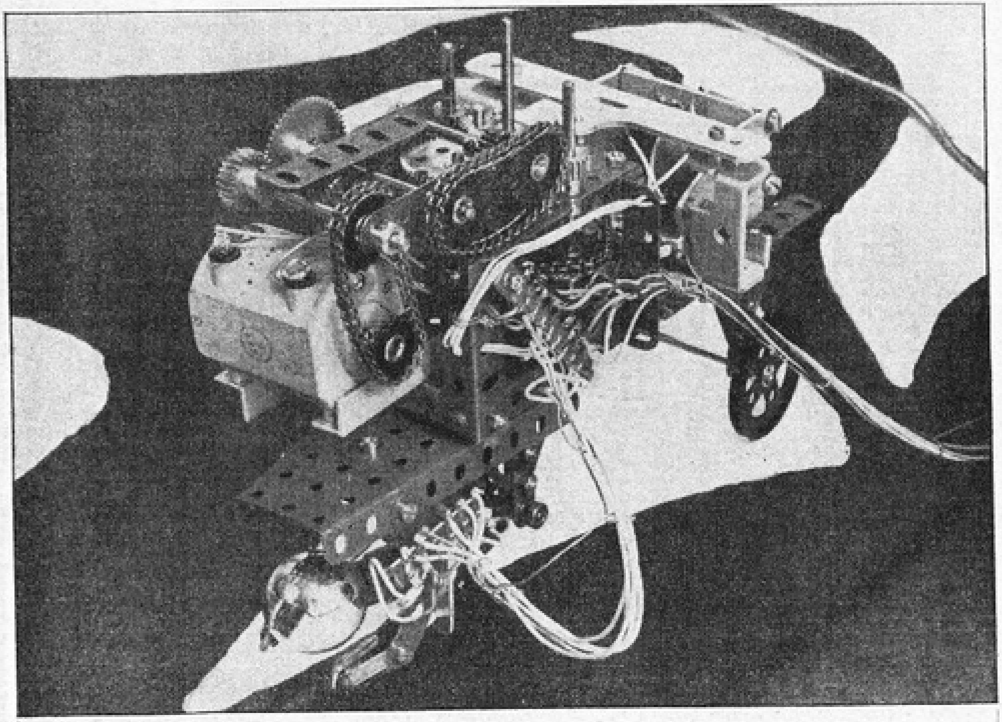

FIGURE 3. A Meccano trolley used to demonstrate control by a learning machine. The trolley follows the boundary between a black area and a white area, irrespective of whether black is on its left and white on its right (as shown) or vice versa. At the front of the trolley are a lamp and two phototransistors: when one transistor is over white and one over black the trolley runs ahead, but if both are over black, or both over white, it stops and hunts from side to side until the phototransistors are again straddling the boundary.
The trolley is connected to the conditional probability computer through four of the input channels, which are labelled for this purpose left, right, black and white. The trolley signals to the computer which direction of searching it has found to be successful after finding both phototransistors on black, or both on white, and the computer learns how to associate left and right with black and white. Thereafter, when the trolley goes on to the white or the black area, the computer makes an inference which indicates to the trolley which way it should turn to regain the boundary. If the trolley is then turned round so that it traverses the boundary in the opposite direction, the inferences made by the computer are all wrong for a time, but it eventually re-learns.
FIGURES 2 and 3 are reproduced by kind permission of the Director of the National Physical Laboratory.