Theoretical Considerations of Routine Maintenance

by E. S. Page

Summary: A model of the value of computer operation is proposed, and its consequences are evaluated far enough to indicate what it is necessary to believe about the benefits of routine or preventative maintenance, in order to make such maintenance valuable.

1 INTRODUCTION

Many manufacturers of automatic digital computers recommend that regular periods be set aside for engineers to perform what is called routine or preventative maintenance. During these periods more or less extensive tests are performed, under either normal or marginal conditions or both, components are changed, and engineering work is carried out according to a general scheme. Whatever the details of the scheme its adoption is advised because the manufacturers recommending it believe-and indeed may have extensive evidence to support their belief-that it promotes a better standard of performance, in some sense, from their product. In his turn, the owner of the computer adopts the manufacturer's advice (if he does) not merely because he trusts the manufacturer implicitly, but also because he believes that it is to his advantage to do so; in other words, he orders routine maintenance because he hopes to gain a better standard of performance from his computer. It is necessary to realize that the criteria for assessing a "better standard of performance" are not necessarily the same for the owner and the manufacturer-though they may well be. Whatever the criteria, it is important to analyse the conditions under which routine maintenance is worthwhile. This paper is not intended to challenge the wisdom of instituting routine maintenance on computers of today where it is performed, or conversely the wisdom of dispensing with it where this course is adopted; instead this paper gives an attempt to provide a basis for analysing the rather vague impressions that routine maintenance is a good or a bad thing. The arguments that are presented here apply with only minor changes to the routine maintenance of many other types of machine, and even perhaps to certain organic and human situations.

In Section 2 the model is proposed and explained, while in the following section a few cases covered by the theory are examined. The mathematical derivations are given in the Appendix which may be omitted by readers who are interested solely in the results.

2 MODEL OF COMPUTER OPERATION

The principal concern of an individual user, when he comes to the computer, is that there should be a high probability that he will be able to complete his calculation without machine failure. Mayne (1959) has proposed a model of computer operation, and has shown how to compute this probability for different lengths of calculation and different computer reliabilities. In this paper we are concerned not with these individual probabilities but with the behaviour over the whole pattern of problems presented to the computer by all its users. This situation has a parallel in the sampling inspection of batches of manufactured products for acceptance or rejection. The inspection scheme specifies a rule according to which the sampling is performed, and the batch is accepted or rejected; it is then usually possible to calculate the probability that an individual batch will be accepted given that it contains a stated proportion of defective items. Thus the behaviour of the inspection scheme on individual batches is known. If the total cost of applying such an inspection scheme is to be calculated, it is necessary to know not only the behaviour of the scheme on individual batches, but also the pattern of the standards of the batches presented for test, and some basic costs, for example, sampling. In the sampling application the distribution of the quality of the batches is described by the process curve; in the model of computer operation the corresponding description will be given by a specification of the pattern of problems to be done.

It has been mentioned that the criteria for assessing the value of a given standard of performance of a computer may vary according to the person doing the assessment. More specifically, the value of the performance of a computer to its owner or user depends on the problems that are to be presented to the machine. Manufacturers naturally try to consider the requirements of a wide range of possible users, and they therefore attempt to ensure that their computers would regularly satisfy performance standards such as: k error-free runs of *n* hours within a period of *N* hours, less than *n* hours lost by unscheduled maintenance in N hours, less than m machine failures in M hours, and so on. The value and the relevance of such standards to users vary. For example, a user whose problems require error-free runs longer than n hours would attach little value to runs of just n hours; again, it would be possible for a user with a large number of short problems to obtain results from a computer which failed frequently. This latter situation was often experienced in the early days of computers, when their reliability was much less than it is today. To express these ideas mathematically we define a function v(x) which is the average value to a particular user of an error-free run of exactly x hours, and which is zero for negative x. The function v(x) is intended to express both the pattern of the lengths of computer runs

required by the problems and the relative values of the problems when completed. One of the simplest functions is v(x) = x; for this function error-free runs have a value proportionate to their length. This interpretation is not usually applicable exactly to the operation of automatic computers, since very short runs terminated by a failure are rarely of any value. It may be more appropriate to a manufacturing machine which produces output at a uniform rate and which stops at an error. A function which is small for small error-free runs and which increases linearly for long runs is v(x) = $x(1 - e^{-kx})$. For a set of jobs of length *l*, a step function is useful for v(x); v(x) = [x/l] where square brackets denote "integral part of." Another possibility is $v(x) = e^{kx} - 1$, which attaches great weight to the longer runs and perhaps may include an element for the freedom from irritation and loss of confidence provoked by machine failures.

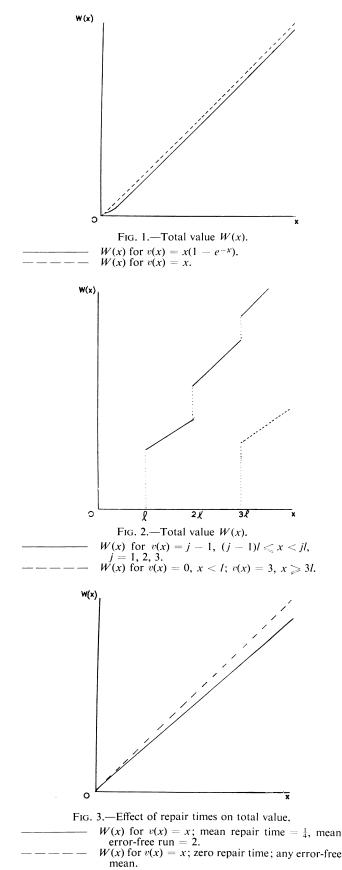
The model supposes that the computer is operated for a period of T hours, the first t of which are spent in routine maintenance. If further periods of routine maintenance are performed before the computer is switched off, T is the time between the start of two consecutive periods of routine maintenance. We suppose that the lengths of error-free periods are independent random variables with frequency function f(x). The characteristics of this probability distribution may depend on the time allowed for maintenance, but it is assumed that the distribution is unchanged by the occurrence and repair of a fault. This assumption seems a reasonable one for the common types of fault in computers, which are rectified by the repair or replacement of just a few components; the overall standard of reliability of the machine is not changed by the repair of a single fault. Similarly we assume that the times required to identify and repair faults are independent random variables with frequency function g(x) which again may depend on the routine maintenance time.

From this model we shall calculate the total expected value to the user for periods of T hours under the different assessments of value and with distributions f(x), g(x) which are affected differently by the amount t of routine maintenance. The mathematical derivations are given in the Appendix and interpreted in the next section.

3 CONSEQUENCES OF THE MODEL

3.1 General

For the model described, we define a function W(Z), the expected total value of the computer's operation in a period Z before the next routine maintenance period. The calculation of W(Z) from the other characteristics of the model is given in the Appendix; here we consider the results of particular evaluations. First, we have neglected any repair time and have considered the particular case when the distribution of error-free time has exponential form. The results for four of the value functions described above are shown in Figs. 1 and 2.



In all cases the form of the value function v(x) exerts considerable influence on the total value W(x). Presentday computers are not normally subject to long delays for repairs, and the mean repair time is likely to be small compared with the basic period over which the evaluation is considered. It can be expected that when repair times of the orders experienced now are taken into account, the form of W(x) will still depend considerably upon the form of v(x). This expectation is confirmed by the example shown in Fig. 3. Other examples may be evaluated from the results in the Appendix. Clearly the repair times have a smaller effect the smaller the probability of an error in the period considered.

The effects of assuming different distributions of errorfree time can be studied using the formulae in the Appendix. Exponential and rectangular distributions yield (A.2.7) (case k = 1) and (A.2.10) and it is seen that, even with an *a priori* unlikely distribution such as the rectangular, the effect of v(x) on W(x) is great and similar in substance to that for the exponential; in particular for v(x) = x, W(x) = x for all distributions when repair time is neglected. We thus have reason to expect that, in conditions applicable to computers, the model is not over-sensitive to the form of the error-free time distribution; in what follows we shall study principally an exponentially distributed error-free time and shall expect the results obtained to apply qualitatively for some other distributions.

An immediate consequence of the model shown by Fig. 2 is that it is more advantageous to program a long computing job so that it may be split into shorter ones. The appropriate value functions to be compared are step functions, the first with a single step and the second with several steps finally reaching the height of the single step. The total value W(x) for the shorter runs is never less than that for the single run. It has long been the practice of many programmers to construct long programs so that checking information is printed periodically, and so that the programs may be restarted if necessary after the last satisfied check in the calculation. Particularly in the early days of computers, it was foolish to waste whatever error-free machine time had been obtained; even today it is wise to avoid too much reliance on a normally high standard of reliability. This general conclusion is accepted by many programmers and to them it is certainly no revelation of the model; instead its emergence from the theory will support a belief in the adequacy of the model. As it is formulated above, the model goes into little detail and it may be that those disagreeing with the above conclusion require additions and amendments to the assumptions.

3.2 Routine Maintenance

A period will not be set aside for routine or preventative maintenance unless it is believed that some benefit is obtained by doing so; in our terms the benefit is gauged by the increase in the total value of the machine runs during the period considered. The time devoted to routine maintenance may or may not affect the time theoretically available for computing; the engineers may start work early and complete their tests before the users arrive. There is thus usually some flexibility in the operation of the system, and this flexibility is limited by what the human beings concerned—engineers, operators, programmers—will accept. However, it is usually the situation that time spent in maintenance reduces the time for computing, and this is certainly the case during continuous twenty-four-hour operation. Accordingly, if a total time T is available and the initial period of length t is devoted to routine maintenance, we are interested in the total value in time T - t, i.e. W(T - t).

Since W(Z) is a monotonic increasing function of Z it is clear that, if there is no change in the distributions of error-free runs and of repair time, the value when any routine maintenance is performed, W(T-t), is less than that with no maintenance, W(T). If we do routine maintenance, we must therefore believe that, as a consequence, the distributions of error-free run and repair time are changed, or, more drastically, that an entirely new model is necessary to describe the computer performance. We shall not pursue the last alternative, and in the examples that we have worked we have only considered cases where the mean of the distributions changes but the functional form does not; the formulae quoted are, however, applicable to any form of the distributions. In computer applications the mean repair time of a fault is likely to be small whether or not routine maintenance is performed, and a general indication can therefore be obtained by neglecting any change in repair time.

Suppose that the effect of t hours of routine maintenance regularly causes a linear increase in the mean error-free run; thus for an exponential distribution $f(x) = \lambda e^{-\lambda x}$, $\lambda = (a + bt)^{-1}$, where a, b > 0. It may be that the ultra-sceptics believe that too much maintenance causes a greater frequency of faults; if this is true the form assumed for λ will hold only as an approximation for the smaller values of t.

It has been remarked that the method of assessment of the value of an error-free run affects the total value in a period considerably, and will therefore affect the decision whether or not to do routine maintenance. For example, if there is just one long continuous job to be performed in the given period, it seems on general grounds that it would be advantageous to do maintenance beforehand in order to increase as much as possible the probability that the first run would be error free. We have thus to compare the W(Z) given by (A.2.7) for Z = T - t, k = 1, $\lambda = (a + bt)^{-1}$ for different values of the maintenance period, t. If we just compare the two extremes in this case, no maintenance and as much maintenance as we can do and still leave enough time for the problem, i.e. a period t = T - l, we have that the maintenance is worthwhile if

$$\exp -l/(a + bt) > \{1 + (T - l)/a\} \exp -l/a.$$
 (3.2.1)

LEAST	VALUES	OF	b	FOR	Economic	MAINTENANCE
-------	--------	----	---	-----	----------	-------------

Т	а	т	b
8	6	0.25	68
8	8	0.25	86
8	6	0 · 50	33
8	8	1.00	21

T = time between successive maintenance periods.

a = mean error-free time without maintenance. m = mean repair time.

m = mean repair time.

b = rate of increase of error-free time with maintenance.

This inequality is certainly satisfied if b > a/(2l - T). In this case quite a small increase in the mean error-free time can justify the period of maintenance; for example, if T is an eight-hour day and l a six-hour run for a computer with mean error-free run without maintenance of four hours, the maintenance is worthwhile if b > 1. In some cases a shorter period of maintenance will be profitable. In another case where all error-free time is useful, v(x) = x, however large an increase in the length of a mean error-free run is obtained by routine maintenance, it is not worthwhile to do it if the repair time is negligible, but it may be if repair time is appreciable.

If the probability is high that the period considered will be error free even when no routine maintenance is performed, it is reasonable to believe that it will not be economic to institute maintenance. Suppose that the repair times are approximately constant at *m* hours, and that the mean error-free run is a + bt, a, b > 0, when *t* hours are regularly assigned for maintenance. If v(x) = x, the total value W(Z) is given approximately by (A.4.4) with Z = T - t and $\lambda = (a + bt)^{-1}$. Then some maintenance is worth while if dW/dt > 0 at t = 0. Accordingly, we have the condition for economic maintenance that

REFERENCES

MAYNE, A. J. (1959). To be published. SMITH, W. L. (1958). "Renewal Theory and its ramifications," J. Roy. Stat. Soc., Vol. 20, Part B, p. 243.

APPENDIX

MATHEMATICAL DERIVATIONS

A.1 General Solution

Under the model described, the time available for computation before the start of the second routine maintenance period is T - t. Let W(Z) be the expected total value of the computer's operation in a period Z before the next routine maintenance. If the computer is error free in T - t then the value is v(T - t), and the probability of this is $\int_{T-t}^{\infty} f(x) dx$. If an error occurs after

time x, and this error takes time y to rectify, there is still a period T - t - x - y before the next maintenance, and the expected additional value in that time is f(x)dxg(y)dy. Thus we have

$$W(T-t) = v(T-t) \int_{T-t}^{\infty} f(x)dx + \int_{0}^{T-t} \left\{ v(x) + \int_{0}^{T-t-x} W(T-t-x-y)g(y)dy \right\} f(x)dx$$
(A.1.1)

$$b > \frac{a}{T} \left[\frac{a}{m} e^{T/a} - 1 \right]. \tag{3.2.2}$$

Some values are shown in Table 1. These values show how large the gain must be from maintenance for it to be economic when the performance is good without it. It must be remarked that the value function used in this example is frequently appropriate to computer operation when performance is high; its failings are its disregard of users' true feelings about very short error-free runs, and such runs are less frequent when reliability is good.

3.3 Conclusion

A model of the value of computer operation has been proposed which has as its consequences some widely held beliefs and which, in addition, sheds some light on what it is necessary to believe about the effects of routine maintenance if it is to be carried out. It appears reasonable in some important cases to perform maintenance immediately before a long machine run which it is necessary to complete in one piece; this merely corresponds to the prudent motorist whose usual practice is to overhaul his car before a long journey. If the performance of the computer is such that errors in the given period are infrequent, then it is probably not worth while trying to make them still less frequent. This indicates that a measure of computer efficiency that is commonly used-the ratio of error-free time to the time available for computing-is not necessarily a useful measure of the value of the computer operation: if efficiency is near unity it may not be worth while sacrificing computer time to maintenance to make the efficiency still nearer unity. The results indicate that it may be valuable to lengthen the period between successive spells of maintenance, and this is indeed the practice with some computers operating continuously for more than twenty-four hours.

Routine Maintenance

Let Z = T - t. Hence

$$W(Z) = v(Z) \int_{Z}^{\infty} f(x) dx + \int_{0}^{Z} v(x) f(x) dx + \int_{0}^{Z} f(x) dx \int_{0}^{Z-x} W(Z - x - y) g(y) dy.$$
 (A.1.2)

This integral equation for W(Z) yields the function we require. Let an asterisk denote the Laplace Transform of a function so that

$$f^*(p) = \int_0^\infty e^{-px} f(x) dx.$$
 (A.1.3)

Then (A.1.2) is of the form

$$W(Z) = F(Z) + \int_{0}^{Z} f(x)h(Z - x)dx. \quad (A.1.4)$$

where
$$F(Z) = v(Z) \int_{Z}^{\infty} f(x) dx + \int_{0}^{Z} v(x) f(x) dx$$
, (A.1.5)

$$h(Z) = \int_0^Z W(Z - y)g(y)dy.$$
 (A.1.6)

Hence

Hence
$$h^*(p) = W^*(p)g^*(p)$$
, (A.1.7)
and $W^*(p) = F^*(p) + W^*(p)f^*(p)g^*(p)$. (A.1.8)

Thus
$$W^{*}(p) = F^{*}(p) / \{1 - f^{*}(p)g^{*}(p)\}$$
 (A.1.9)

The problem of determining W(Z) is therefore reduced to the not always trivial one of inverting (A.1.9). Even if an explicit mathematical solution is difficult to obtain, a numerical one can be calculated. In the remaining sections of the Appendix special cases are evaluated: (i) repair time neglected, two functions v(x) for the same distribution of error-free time, f(x); (ii) repair time neglected, one function v(x) for two distributions f(x); (iii) appreciable repair time, one function v(x) and one distribution f(x).

A.2 Negligible Repair Time

If the repair time is negligible, equation (A.1.2) simplifies to

$$W(Z) = V(Z) \int_{Z}^{\infty} f(x) dx + \int_{0}^{Z} f(x) f(x) dx + \int_{0}^{Z} f(x) W(Z - x) dx, \quad (A.2.1)$$

so that

$$W^{*}(p) = F^{*}(p)/\{1 - f^{*}(p)\}.$$
 (A.2.2)

The case V(x) = x degenerates if the repair time is ignored with the solution W(Z) = Z; this merely expresses the fact that the total error-free time in a period Z is just Z if no time is lost. More interesting results can be obtained for other value functions.

It is necessary to specify the form of the distribution of error-free time. One of the simplest mathematically that is likely to occur in practice is $f(x) = \lambda e^{-\lambda x}, x > 0$; this distribution arises if faults occur at random at rate λ , so that the probability of a fault in a short interval δt is $\lambda \delta t$. We suppose that λ may be a function of t.

 $f^*(p) = \lambda/(\lambda + p),$

Since

$$W^{*}(p) = (1 + \lambda/p)F^{*}(p).$$
 (A.2.3)

Hence
$$W(Z) = F(Z) + \lambda \int_0^Z F(x) dx.$$
 (A.2.4)

In particular, for $v(x) = x(1 - e^{-kx})$, (Fig. 1)

$$W(Z) = \{(k^2 + 2\lambda k) - k^2 e^{-(\lambda + k)Z}\}Z/(\lambda + k)^2.$$
 (A.2.5)

For the value function treating the situation, where kjobs each of length *l* need to be done, take v(x) = j - 1, $(j-1)l \leq x < jl$ for $j \leq k$.

Then
$$F(Z) = \sum_{i=1}^{i-j} e^{-i\lambda l}, (j-1)l \leq x < jl,$$
 (A.2.6)

and W(7)

(A.1.7)

$$\begin{split} W(Z) &= 0, \quad Z < l \\ &= e^{-\lambda l} \{1 + \lambda (Z - l)\}, \quad l \leq Z < 2l \\ &= e^{-\lambda l} \{1 + \lambda (Z - l)\} + e^{-\lambda l} \{1 + \lambda (Z - 2l)\}, \\ &\qquad 2l \leq Z < 3l. \quad (A.2.7) \end{split}$$

and so on, ending with the interval $kl \leq Z < \infty$ (Fig. 2).

In order to compare the effects of different distributions of error-free time, x, we take an example for which the distribution of x is uniform in (0, d). We consider the case where d is greater than times for which the average value of operation is to be evaluated, so that there is a non-zero probability of a complete error-free period of the length we are considering. Direct solution of (A.2.1) for this distribution and for value functions such that F'(Z) exists for $Z \leq d$, gives

$$W'(Z) - W(Z)/d = F'(Z), \quad 0 \le Z \le d, \quad (A.2.8)$$

so that

$$W(Z) = e^{zId} \int_{0}^{Z} e^{-x/d} F'(x) dx, \quad 0 < Z < d, \quad (A.2.9)$$

F'(x) = (1 - Z/d)v'(x).where

For the value function v(x) = 0, x < l; v(x) = 1, $x \ge l$, direct solution of (A.2.1) gives

$$W(Z) = 0, \quad Z < l$$

= $\left(1 - \frac{l}{d}\right)e^{(Z - D)/d}, \quad l \leq Z \leq d.$ (A.2.10)

A.3 Exponential Repair Time

If the repair time is exponentially distributed so that $g(x) = \mu e^{-\mu x}$, $x \ge 0$, and the error-free time has a similar distribution $f(x) = \lambda e^{-\lambda x}$, $W^*(p)$ given by (A.1.9) is easily inverted. We have

$$W^{*}(p) = \frac{(p+\lambda)(p+\mu)F^{*}(p)}{p(p+\lambda+\mu)}.$$
 (A.3.1)

For v(Z) = x we have (Fig. 3)

$$W(Z) = \frac{\mu Z}{\lambda + \mu} + \frac{\lambda}{(\lambda + \mu)^2} (1 - e^{-(\lambda + \mu)Z}). \quad (A.3.2)$$

The solutions for $v(x) = x(1 - e^{-kx})$ and for v(x) = 0, x < l; v(x) = 1, $x \ge l$ can be obtained straightforwardly.

A.4 Approximate Solutions

If the probability of more than one fault in the period to be considered is negligible, an approximation to W(Z) may be found explicitly. Since second and successive errors are to be neglected, the average value in the period is the sum of the value in the first error-free run and the value of the time to the end of the period after the repair of the first fault. Hence

$$W(Z) \simeq v(Z) \int_{Z}^{\infty} f(x) dx + \int_{0}^{Z} v(x) f(x) dx + \int_{0}^{Z} f(x) \left\{ \int_{0}^{Z-x} v(Z-x-y) g(y) dy \right\} dx. \quad (A.4.1)$$

Thus the transform of the exact solution (A.1.9) is approximated by

$$W^{*}(p) \cong F^{*}(p) + v^{*}(p)f^{*}(p)g^{*}(p).$$
 (A.4.2)

In particular, for the case v(x) = x, $f(x) = \lambda e^{-\lambda x}$ and negligible repair time, (A.4.1) gives the exact solution W(Z) = Z. Again, for the same frequency functions but v(x) = 0, x < l; v(x) = 1, x > l, (A.4.1) gives

$$W(Z) \simeq e^{-\lambda l} + 1 - e^{-\lambda(Z-l)},$$
 (A.4.3)

which differs from the exact solution by $0(\lambda^2 Z^2)$. If the repair time for all faults is equal to a constant *m* we have an approximation valid for *m* small compared with *Z*

$$W(Z) \simeq Z - m(1 - e^{-\lambda Z}),$$
 (A.4.4)

for v(x) = x and $f(x) = \lambda e^{-\lambda x}$. This result follows directly since there is no fault in time Z with probability $e^{-\lambda x}$, and this approximation ignores the effect of the second and subsequent faults; hence time m is lost with probability $1 - e^{-\lambda x}$, where here we neglect the reduction in time lost when the fault occurs in (Z - m, Z). An approximation ignoring only the third and subsequent faults may be easily obtained.

A.5 General

The model of computer value that has been discussed is a stochastic process of renewal type. Instead of being given one sequence of identically distributed non-negative random variables $x_1, x_2 \dots$ and considering problems concerning their partial sums, we have two such sequences $\{x\}, \{y\}$ with different distributions which represent the error-free runs and the repair times respectively. We are interested in the random variable $w(Z) = \sum_{i=1}^{n} v(x_i)$, and in particular its average value W(Z), where *n* is the smallest integer such that $\sum_{i=1}^{n} (x_i + y_i) > Z$; in the sum for W(Z), the last term $v(x_n)$ is replaced by

$$v\left\{Z - \sum_{i=1}^{n-1} (x_i + y_i)\right\}$$
 if $x_n > Z - \sum_{i=1}^{n-1} (x_i + y_i)$.

Alternatively we can regard the process as a random walk, towards an absorbing barrier at x = Z, in which the steps are taken alternately from the distributions f(x), g(x). Only the steps from the x distribution contribute to the sum w(Z). Again, the process can be regarded as arising from a counter with dead time (e.g. Smith, W. L., 1958, Section 3). Most of the previous work is concerned with asymptotic results which, in our model, would deal with the case of a large number of machine faults—and we hope that this no longer has relevance to automatic computer operation.

The Mechanization of Thought Processes

The Proceedings of the Symposium on "The Mechanization of Thought Processes" held at the National Physical Laboratory on 24th-27th November 1958 have now been published. They are, as far as possible, a complete record of the Symposium, including the 32 papers in full, together with a number of appendices and a full report of the discussion. The discussion was recorded and all contributors and authors were asked to edit their contributions. In addition, the Proceedings contain a description of apparatus demonstrated at the Symposium. The field covered by the Symposium was very wide: the main session headings were:—

- 1. General principles-artificial intelligence, intellect, habituation, conditional probability.
- 2. Automatic programming—in Russia, America and U.K. Mechanical language translation.
- 3. Speech recognition. Learning in machines.
- 4A. Implications for Biology-medical diagnosis, animal learning, sensory mechanisms, redundancy, nervous pets.
- 4B. Implications for Industry-legal world, information retrieval, learning processes.

The Proceedings are published by Her Majesty's Stationery Office, price £2 10s. 0d. (two volumes).