Automata of High Complexity and Methods of Increasing Their Reliability by Redundancy

LARS LÖFGREN

Swedish Research Institute of National Defence, Stockholm, Sweden

A natural question regarding the limits of automation is: What are the limits of automata, or how closely can an automaton imitate the human being—what kind of specific processes can an automaton do better, more quickly, and more reliably than the human being? Aspects of these questions may be treated from the standpoint of automata theory. We will give brief summaries of two far-reaching theories in this field: the Turing theory of the possibilities of computing machines, and the von Neumann theory of self-reproducing automata. The question of reliability becomes rather serious when the complexity of the automaton becomes large. An attempt towards a theory of how redundancy should be applied on different levels for reducing the influence of temporary component errors in a universal computing machine is given.

1. INTRODUCTION

The word automation, as related to production, seems to mean automatic processes carrying out the production of specified things. The purpose is to replace more or less completely the participation of the human being by these automatic processes. This usually involves a restatement of processes originally set up for treatment by human beings.

A natural question regarding the limits of automation therefore is: What are the limits of automata, or how closely can an automaton imitate the human being (what kind of specific processes can an automaton do better—more quickly and more reliably—than the human being).

Considering automata as computing machines (the control aspect of automation) the question has been treated by Turing (1936, 1950). A good survey is given by Shannon and McCarthy (1956).

1 This paper covers the contents of two communications given by the author at the International Congress on Automation, held in Paris, June 18-24, 1956, and at the International Congress on Cybernetics, held in Namur, June 26-29, 1956, respectively.
A brief study of automata producing automata has been given by von Neumann (1951). He considers the concept of self-reproduction and approaches in this way the question of human behaviour of automata.

In the present practical stage of automation the question is not so much directed towards an automaton that can imitate a human being, but rather towards one which imitates certain very specific features of man, and furthermore does them more quickly and more reliably.

The question of reliability becomes rather serious at least when the complexity of the automaton becomes large (as for instance for self-reproducing automata). We will try to give a theory of how redundancy should be applied for reducing different types of errors (within the control aspect of automation).

2. SOME RESULTS OF THE TURING THEORY

The question of the possibilities of automatizing the logical processes carried out by the human being is conveniently studied in terms of the propositional calculus.

The set of propositions is closed under the operations of negation, conjunction and disjunction. It can therefore be mapped by a homeomorphism on to the Boolean algebra of the two elements 1 and 0.

Functions of the Boolean algebra can be realized as networks containing only three different kinds of organs, basic organs for an automaton. [Compare also the Sheffer stroke (von Neumann, 1956).]

In his paper, "On Computable Numbers," Turing (1936) investigates the possibilities and limitations of automata. The class of machines he investigates consists of a control element, a reading and writing head, and an infinite tape. The control element is a device with a finite number of internal states (i). The operating characteristic of the machine is the description of how it will change its internal state and what symbol (f) it will write on what place (p) of the tape (p = ±1, 0 which means moving the tape one stage to the right, or to the left, or not moving it at all). Suppose the machine is in state i, and reads the symbol e. This configuration i,e will then determine j, p, and f (where j is the new internal state). The coordination of j, p, f and i,e is the complete definition of a specific automaton, and can be characterized by a description number.

The sequence of symbols which such a machine prints on specified places of the tape may be regarded as a representation of the number computed by the machine (conveniently in binary-point form).

Turing refers the problem of computability to an investigation of cir-
cular and circle-free machines. If a machine never writes down more than a finite number of symbols on the specified places, it is called circular. Otherwise it is circle-free. A sequence is computable if it can be computed by a circle-free machine. A number is computable if it differs by an integer from the number computed by a circle-free machine.

Turing shows that the class of computable numbers is great and in many ways similar to the class of real numbers. It is, however, enumerable. He gives the following example of a number which is not computable. Let $\delta$ be a sequence whose $n$th figure is 1 or 0 if $n$ is or is not the description number of a circle-free machine. This number $\delta$ is not computable.

Turing's main result is that it is possible to give the description number of a single machine which can compute any computable number.

3. von Neumann's concept of self-reproduction

von Neumann (1951) has broadened the theory to deal with automata that produce automata. Here the number of basic organs required is somewhat larger than three (about twelve).

The equivalent of Turing's main result is that it is possible to construct a single automaton which, when furnished with the description of a specific automaton, will construct that automaton. "Constructing" is to be understood here as an appropriate connection of basic organs, which are supposed to exist already.

We will now see that a self-reproducing automaton can be constructed. Suppose we have an automaton $A$ which, when supplied with the instruction $I_F$, will construct the arbitrary automaton $F$ (see Fig. 1). Further we need an automaton $B$ which can make two copies of any instruction $I$ that is furnished to it. Finally we need a control mechanism $C$ which does the following:

Let $A$ and $B$ be furnished with an instruction $I_F$. Then $C$ will first cause $A$ to construct $F$, and cause $B$ to produce two copies of $I_F$. Next, $C$ will supply $F$ with one of the copies $I_F$, and separate $F$ supplied with $I_F$ from the system $A + B + C$ which we call $D$. At the same time the remaining copy $I_F$ is also separated from $D$ and supplied to the $AB$ instruction input.

This system $D$ is now well defined according to the above. (Notice it does not contain a description of $I_F$ but only of $A$, $B$, and $C$.) Thus we can produce its instruction $I_D$ (intelligible relative to $A$), and we may supply $D$ with $I_D$, which entity we call $E$. Clearly this automation $E$
is self-reproducing, since, as is seen from Fig. 1, the entity which will be separated from $E$ is $D$ supplied with $I_D$, which is just $E$.

Small variations of this scheme also permit the construction of automata which can reproduce themselves and, in addition, construct other types of automata.

Probably an automaton of relatively small complexity can only produce automata of still smaller complexity. The concept of self-reproduction is probably a qualitative property, only possible at a certain complexity.

A possible practical limitation of the complexity is the unreliability of the basic organs. In the following, however, we will see how this limitation can be overcome.

---

**Fig. 1. Self-reproduction.**
4. RELIABILITY BY REDUNDANCY

4.1. REDUNDANCY SCHEMES FOR APPLICATION ON DIFFERENT LEVELS

Although the three basic propositions can be realized in three basic organs (or even in one, the Sheffer stroke), we can apply redundancy to different levels to improve the total reliability. (We will use the notation \(ab\) for a conjunction, \(a \lor b\) for a disjunction of \(a\) and \(b\), and \(a'\) for the negation of \(a\).) Redundancy can be applied to (1) basic organs as units or on aggregates of them, and (2) constructional parts of the basic organs.

We will start with (2) since this case will indicate the method for (1). [(1) will be treated in Sec. 4.12.]

4.11. Redundancy of Constructional Parts of the Basic Organs (Micro-level Redundancy)

Let us first consider a constructional part \(C_1\) (of a basic organ) which performs some kind of transmitting operation on the input quantity \(a\), thereby producing the output \(b\) (Fig. 2a). Suppose there is only one type of error, an error which prevents the transmitting operation and thereby makes the output free to assume any value forced upon it by other means. The redundancy scheme obviously is a parallel arrangement according to Fig. 2a. If one of the parts \(C_1\) should fail, the rest will give to the output the correct value \(b\).

Let us next consider the inverse problem. We have here a constructional part \(C_s\) which performs a nontransmitting operation, and the only type of error is a pure transmission. Here the redundancy scheme is the inverse of the parallel arrangement of Fig. 2a, i.e. the series connection of Fig. 2b. Should some of the parts \(C_2\) fail, the rest will perform the nontransmitting operation.

![Fig. 2. Basic redundancy schemes.](image)
If both of the above types are present, the redundancy structure naturally will be a combination of the parallel and series arrangements. Let us take as an example, a diode. Suppose there are only two types of errors: a total open circuit of the element, and a total short circuit. Let the probabilities of these two errors be \( p_o \) and \( p_s \), respectively. Then the probability of error in the redundancy arrangement of Fig. 2c will be (see Creveling, 1956)

\[
p_s^2(2 - p_s^2) + p_o^2(2 - p_o^2).
\]

Naturally the quadrupling according to Fig. 2c can be applied once again on the “quad” and so on.

We will call this type of redundancy, microlevel redundancy.

4.12. Redundancy of Basic Organs or of Aggregates of Basic Organs. Separate and Combined Corrections

Here we have only to consider errors which give outputs contained in the set of allowable input and output quantities. Whatever the error in a single basic element might be, the output of this element will be interpreted as either 1 or 0 by the next organ.

All three basic elements perform different kinds of transmitting operations. We are therefore referred to the parallel redundancy scheme of Fig. 2a. However the connection of the individual outputs to a common output cannot be done as simply as in Fig. 2a. The reason is that an error in a basic element will in general not make it free to accept a value which is forced upon it by the other elements in the parallel arrangement. This calls for a special connecting organ (see Fig. 3). When only a single operation is concerned, the construction of the connecting organ, called by von Neumann (1956) the “majority organ,” is rather straightforward, as we shall see in the following section. After that we shall consider an organ for a redundancy connection for two operations (combined correction) and compare this type of redundancy with two separate connections, one for each operation (separate correction).

Connecting organ (majority organ) for redundancy protection of a single operation. Suppose we want to protect a kind of transmitting operation realized in an element \( T \) by introducing one extra element \( T \) operating on the same input. Suppose one of them operates wrong. They will then produce different outputs. Evidently there is no way to determine, by just observing the outputs in a single experiment, which of the two outputs is correct. Thus we have to introduce a third element \( T \) also operating
Fig. 3. Connecting methods for redundant transmission of (a) a single quantity, (b) a pair of quantities.

on the common input (see Fig. 3a). Now, a single error can be corrected because two of the three outputs will indicate the correct result. Suppose the three outputs are $a$, $b$, and $c$, respectively. Then an organ realizing the majority expression

$$ab \lor ac \lor bc$$

will give the correct result (provided not more than one of the three elements fails at the same time). Such an organ, the majority organ, is easily constructed out of the three basic organs or directly out of diodes only (nine diodes).

Suppose the probability of failure for each $T$-element is $p$. The total probability of failure will then be

$$3p^2(1 - p) + p^3 = 3p^2 - 2p^3.$$  

Evidently we can obtain a still lower probability of error by constructing a majority organ which generates the majority value of the outputs of a larger number of $T$-elements.

**Connecting organs for redundant transmission of a pair of quantities.** We will consider a pure transmission (the identity operator, not an arbitrary transmitting operator as was more generally assumed for the majority organ) of a pair of quantities $A_1$ and $B_1$. Evidently each quantity can be transmitted, as in Fig. 3a, for a correction of all single errors. This method (separate correction) requires six elements $T$. We will now show that this number can be reduced to five by making the two connecting organs operate on some of the five outputs $a, b, c, d,$ and $e$ in common (see Fig. 3b) (combined correction).

Let $T_a$ and $T_b$ operate on $A_1$, $T_d$ and $T_e$ on $B_1$, and finally $T_c$ on
Thus the quantity fed to $T_e$ is 0 if $A_1$ and $B_1$ are equal, and 1 otherwise. We will now design the two connecting organs so as to correct all five single errors.

If $a$ and $b$ are different we know that the single error is in $a$ or $b$; that is, $c$, $d$, and $e$ are correct (for the moment we consider only single errors). Then $d$, for instance, will indicate the correct value of $B_2$, and $(cd' \cup c'd)$ the correct value of $A_2$.

If, on the other hand, $c$ is incorrect, we are aware of this possibility by observing that $a = b$ and $d = e$ (or $ab' \cup a'b = 0$, $de' \cup d'e = 0$). In this case $a$ and $d$ give the correct values of $A_2$ and $B_2$.

Thus, if we construct the two connecting organs according to

$$A_2 = ab \cup (ab' \cup a'b)(cd' \cup c'd)$$
$$B_2 = de \cup (de' \cup d'e)(cb' \cup c'b),$$

all five possible single errors are corrected. If, as before, the probability of error for each of the transmitting elements is $p$, the total probability of error will be

$$9p^2(1 - p)^2 + 9p^3(1 - p)^2 + 9p^4(1 - p) + p^5$$
$$= 9p^2 - 18p^3 + 14p^4 - 4p^5.$$

(One of the ten possible double errors and one of the ten possible triple errors will also be corrected.)

If, on the other hand, we use the scheme of Fig. 3a for $A_1$ and $B_1$ separately, i.e. we use six $T$-elements, the probability of error will be

$$6p^2 - 4p^3 - 9p^4 + 12p^5 - 4p^6.$$
we use combined correction (Fig. 3b) instead of separate correction (Fig. 2a).

von Neumann's multiplexing technique. Suppose that the probability of error in a single basic element (for instance a Sheffer stroke) is $\varepsilon$. An apparent limitation on the final error $\delta$ of the whole automaton evidently is

$$\delta \geq \varepsilon$$

and this for any amount of applied redundancy. The reason is that the output of the automaton is the immediate result of the operation of a single final basic element, and the reliability of the whole system cannot be better than the reliability of this last basic element.

To remove this restriction, von Neumann (1956) suggested to carry a quantity simultaneously on a bundle of $N$ lines ($N$ being a large integer) instead of just on a single line—the output of a majority organ or another type of connecting organ as we have discussed in the above.

A positive number $\Delta(<\frac{1}{2})$ is chosen. The stimulation of $\geq (1 - \Delta)N$ lines of a bundle is interpreted as the number one; the stimulation of $\leq \Delta N$ lines, as the number zero. Any other number of stimulated lines is interpreted as malfunction.

In this way the output of a multiplexed basic element is not the result of a single basic element, and the above restriction is removed. The probability of malfunction of the whole system can now be made arbitrarily low.

However the following remarks can be made about such a method. As soon as a number (for instance the output of the whole automaton) has to be measured, i.e. translated to a zero or to a one (for instance for printing), the restriction $\delta \geq \varepsilon$ appears again.

On the other hand, the critical basic element in a connecting organ in a nonmultiplexed automaton can be made to have an $\varepsilon$ which is sufficiently low by microlevel redundancy (Sec. 4.11). The probability of error $\varepsilon$ for other basic elements just has to be so small as to prevent error growth. This value can be allowed to be rather high for a sufficient amount of redundancy.

A uniform application of multiplexing in a specific type of machine, as for instance in a universal sequence computer, is rather inefficient as compared to the redundancy method, which we describe in Sec. 4.3. However in the field of cerebral mechanisms, for the understanding of
which von Neumann's multiplexing method was developed, its generality evidently is of greatest importance.

4.2. Further Developments of Redundancy as Applied to Several Quantities in Common (Word Redundancy)

By the example given in Sec. 4.12, we have seen that the redundancy for a certain degree of correction may be reduced when making the connecting organs act on combinations of the primary quantities. In the following this will be shown to be even more efficient when going to combinations of higher order than two.

For pure transmission this method of combined correction has been treated by several authors. In the following Sec. 4.21 we will give a short treatment of a work of Slepian (1956). There are, however, some difficulties in a pure application of the method of combined correction (we will call this type of redundancy, "word redundancy") to real operations. These will be investigated in Sec. 4.22.

In the following we will only consider errors due to the operation to be protected by redundancy, i.e. we will consider the connecting organs (or checks) ideal. We will come back to the real situation in Section 4.22.

4.21. Slepian's Method of Automatic Error-Correction for Identity Operations (pure transmission)

For the following study of combined protection of errors in a pure transmission, let us think of the numbers \( a, b, c, d, e \) of Fig. 3b as a single binary number of five positions. Let us then go over to a number

\[
a = q_1, q_2, \cdots, q_n \quad (= 1011 \cdots 10)
\]

of \( n \) positions. We will now study the correcting possibilities of such a number (corresponding to the structure possibilities of the connecting organs of Fig. 3b). To start with, we will make no distinction between information positions and redundancy (check) positions (compare the numbers \( a, d \) and \( b, c, e \) of Fig. 3b).

Let the set of the \( 2^n \) numbers \( a \) be called \( E \):

\[ a \in E. \]

An error in certain positions will be denoted by the error operator \( e \): an \( n \)-place binary number with 1's in the error positions and 0's in the remaining positions. Thus an error \( e \) in the number \( a \) will give the number
ea, where this product means sum, mod 2, digit by digit. Evidently
\[ e \in E. \]

With the above product definition, \( E \) will have group properties, and will be denoted \( G \):
\[ E = G. \]

For if
\[ a \in G \]
\[ b \in G \]
\[ c \in G, \]
then
1. \( ab \in G \)
2. \((ab)c = a(bc)\)
3. \( I \in G \) (\( I = 000 \cdots 0 \), the identity element defined by \( Ia = a \))
4. \( a^{-1} \in G \) (def.: \( a^{-1}a = I \), thus \( a^{-1} = a \))

\( G \) is an abelian group, since \( ab = ba. \)

The intention is now to divide the group in center numbers, representing correct numbers, and surrounding cosets, each representing erroneous center numbers. Intuitively a good possibility of detecting errors will be obtained when the “distances” between neighboring center numbers are equal. The distance \( d(a,b) \) between two numbers \( a \) and \( b \) is defined as the weight of the product (sum, mod 2)
\[ d(a,b) = w(ab) \tag{1} \]
i.e. the minimum number of positions in \( a \) which have to be changed to obtain \( b \) (or vice versa). Evidently
\[ d(a,b) = d(b,a) \tag{2} \]
\[ d(ac,bc) = d(a,b) \tag{3} \]

If a group contains subgroups, as \( G \) does, it can be divided into a subgroup and cosets. Either two cosets are identical or they do not have any element in common. So we can expand:
\[ G = \sum_{i=1}^{I} b_i g \tag{4} \]
(with the notation of Galois). The subgroup is denoted as $g$, and the index of $b(b_i \in G)$ shall indicate that no two cosets $b_i g$ are identical ($j$ is the index of $g$ in $G$). More explicitly, we can write the expansion

$$
G = I \quad a_2 \quad a_3 \quad a_4 \cdots a_\mu
$$

\[ e_1 \quad e_1 a_2 \quad e_1 a_3 \quad e_1 a_4 \cdots e_1 a_\mu \]

\[ e_2 \quad e_2 a_2 \quad e_2 a_3 \quad e_2 a_4 \cdots e_2 a_\mu \]

\[ \vdots \]

\[ e_{i+1} \quad e_{i+1} a_2 \quad e_{i+1} a_3 \quad e_{i+1} a_4 \cdots e_{i+1} a_\mu \]

(5)

Here the elements of the subgroup are denoted by $I$, $a_2$, $a_3$, $\cdots$, $a_\mu$. In order to protect us from equal cosets we must have the restrictions

$$
e_1 \notin g \quad e_2 \notin g \cup e_1 g \quad \vdots \quad (6)
$$

Still the expansion is not unique. If in a coset $e_ig$, the element $e_i$ is replaced by $e_ia_j$, the same coset will result. Since we want to arrange the cosets with growing distances from the subgroup (the center numbers), an element of a coset with the smallest weight should be chosen as an $e$-element

$$
d(e_i, I) \leq d(e_ia_j, I). \quad (7)
$$

With this arrangement the probability that an element $e_ia_m$ should be referred to the center number $a_m$ cannot be exceeded by referring it to some other center number. First we show this with inverse distances instead of probabilities. Then we will indicate that the probability is increasing with decreasing distance.

According to (3) the two distance arguments can be multiplied by an element in $G$. Thus we can multiply the arguments of the left side of (7) by $a_m$, and the arguments of the right side by $e_ia_m$:

$$
d(e_ia_m, a_m) \leq d(e_ia_j e_ia_m, e_ia_m) = d(e_ia_m, a_j a_m)
$$
The element $a_i a_m$ is in $g$ and may be denoted by $a_t$:
\[
d^{-1}(e_i a_m, a_m) \geq d^{-1}(e_i a_m, a_t)
\]
which is the above statement in inverse distances.

Finally if the probability of a single error in any position is $p$, the probability of an error $e$ with $d(e, I) = w$ is
\[
p^w (1 - p)^{n-w} \approx p^w.
\]
Since $p$ is small compared with unity, we see that the probability is increasing with inverse distance.

We will now proceed to the question of determining the subgroup. Suppose we put the number information in the first $k$ positions of the subgroup elements $I, a_2, \ldots, a_n$ ($\mu = 2^k$). The remaining $(n - k)$ positions (redundancy positions) shall carry check information. The simplest way of checking the information part in a single check position, is evidently, to write in this position the sum mod 2 of the information positions. Any single error will then cause the check to fail. This procedure is naturally expanded to
\[
q_i = \sum_{j=1}^{k} \gamma_{ij} q_j \pmod{2} \quad (i = k + 1, \ldots, n).
\]
Here $\gamma_{ij}$ is either zero or one. Evidently, with this definition of the elements $I, a_2, \ldots, a_n$, the group-properties are maintained.

Suppose now that there has been an error in an element $a_j$ giving another element $A_j$. This will cause some of the checks (9) to fail. This can be indicated by the check sequence
\[
S(A_j) = 0100 \cdots 0 \quad (n - k \text{ pos.})
\]
with zeros in the check positions which satisfy (9), and one's in the other check positions. The check sequence of an arbitrary element $e_i a_j$ can be expanded:
\[
S(e_i a_j) = S(e_i) \cdot S(a_j) = S(e_i) I = S(e_i).
\]
[$S(a_j) = I$ because $a_j$ is an error-free subgroup element.] Thus all elements of a certain coset have the same check sequence. Furthermore the sequences of different cosets are different and so a certain $S(e_i a_j)$ uniquely corresponds to $e_i$. Finally the product $e_i (e_i a_j) = a_j$ gives the correct center number.

When all single errors should be detected and corrected this way, the checks (9) are to be determined so that the weights of the subgroup
elements (except 1) are larger than two. Otherwise at least one single error will give a coset element with weight one. This element (regarded as an e-element, a single error) cannot, because of (6), be taken as a coset leader. Therefore not all single errors can be corrected.

Correspondingly all double errors are detectable if (9) is chosen so as to give the subgroup elements weights larger than four, and so on.

A lower bound to the number of check positions is easily determined from (5). Suppose we want to correct all n single errors. Then there must be at least \((1 + n)\) rows in (5):

\[
2^{n/2^k} \geq 1 + n. \tag{12}
\]

Example: \(k = 11\) (eleven information positions)
\(n - k = 4\) (four check positions).

If also all double errors should be corrected, one must add to the right number of (12) the number of all double errors \(n(n - 1)/2\):

\[
2^{n-k} \geq 1 + n + n(n - 1)/2. \tag{13}
\]

Example: \(k = 14\), \(n - k = 8\).

Going still further, we have to add to the right number of (13) the further binomial coefficients. These lower bounds to the number of check positions were first derived by Hamming (1950). It is of some interest to notice that a complete correction of all possible errors requires

\[
2^{n-k} \geq \sum_{r=0}^{n} \binom{n}{r} = (1 + 1)^n = 2^n
\]

i.e. it is only possible when \(k = 0\) (when no information positions are present).

Let us illustrate the method with the same example we used for Fig. 3b. We want here a correction of all the five possible single errors and we can immediately write down the following part of the expansion (5):

<table>
<thead>
<tr>
<th>00000</th>
<th>01xxx</th>
<th>10xxx</th>
<th>11xxx</th>
</tr>
</thead>
<tbody>
<tr>
<td>00001 (= e1)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>00010 (= e2)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>00100 (= e3)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>01000 (= e4)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>10000 (= e5)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>xxxxx (= e6)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>xxxxx (= e7)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
(\(x\) and \(\cdots\) denote numbers to be filled when the check relations (9) have been determined.) In order to have the five single errors \((e_1, \cdots, e_5)\) appear in the first column, (9) should be chosen, so as to give the smallest weight 3 to the subgroup elements (the first row, excluding \(I\)). One such possibility is

\[
\begin{align*}
q_3 &= q_1 \\
q_4 &= q_2 \\
q_5 &= q_1 + q_2 \pmod{2}
\end{align*}
\]

Except for the two last rows, the expansion scheme is immediately completed:

\[
\begin{array}{cccccc}
S & 00000 & 01011 & 10101 & 11110 & 000 \\
& 00001 & 01010 & 10100 & 11111 & 001 \\
& 00010 & 01001 & 10111 & 11100 & 010 \\
& 00100 & 01111 & 10001 & 11010 & 100 \\
& 01000 & 11011 & 00101 & 01110 & 101 \\
& 10000 & 11011 & 00101 & 01110 & 110 \\
& 11000 & 10011 & 01101 & 00110 & 111 \\
& 01100 & 00111 & 11001 & 10010 & 111
\end{array}
\]

Coming down to \(e_6\) we must look for a double error, the representation of which as a binary number, is not present in the above six rows. One soon discovers 11000. In the same way one finds that \(e_7 = 01100\) is not present in the above seven rows. In the last column we have written down the check sequence \(S\) [see (10)].

We will illustrate how a single error is corrected. Suppose the number \(a = 11101\) is received. Its check sequence \(S\) is formed. It is \(S = 011\), and this number corresponds to \(e = 01000\). The correct number can now be determined as the product

\[
e \cdot a = 10101
\]

(compare the above expansion scheme).

We see that beside all single errors, also two of the ten possible double errors will be corrected. Thus the probability of error will be

\[
8p^2(1 - p)^3 + 10p^3(1 - p)^2 + 5p^4(1 - p) + p^5.
\]
In the above we have only said that the check relations (9) should be chosen so as to give the subgroup elements sufficient weights for a complete protection of all single errors, all double errors, and so on, up to a certain degree. This determination may be quite a problem, especially for large $n$ and $k$. In fact it is not always known if check relations of type (9) exist for an arbitrary probability of error in accordance with the Hamming bound.\(^2\) An upper bound on the number of check positions has been given by Gilbert (1952).

Laemmel (1953) gives, however, indications that the Hamming lower bounds may be a good approximation to the true minimum number of check positions. When comparing this method of correction with separate corrections, each according to Fig. 3a, the distinction is so great that it seems adequate to use the Hamming bounds.

Let us take $k = 10$ and $n = 30$. We have seen in Sec. 4.12 that separate correction (each information position is protected by two check positions) can only correct all single errors. Thus the resulting probability of error will start with a $p^2$-term (provided with a factor $30$).

On the other hand the Hamming bounds indicate that a word-redundancy correction ($k = 10$, $n = 30$) might give a complete correction of all single errors, all double errors, and so on up to all sextuple errors. Then the expression for the probability of error would start with a $p^7$-term (provided with a factor of the magnitude $10^6$, which however is rather insignificant since $p$ usually is less than $10^{-3}$).

If we were only interested in protection against single-errors the Hamming bound indicates that with $n = 30$ we could make $k$ close to 25 (compared with 10 for separate corrections).

In this section we have given an outline of a possible way of determining the check position content (compare the unit and the connections to the left of the transmitting elements of Fig. 3b) and of determining the behavior of the connecting organs (to the right in Fig. 3b) [$S(a)$ implies an $e; e; a = \text{correct number}$].

This method is not the only way of handling the problem, since it concerns a special class of codes. The class is, however, sufficiently broad to include the codes of Hamming, the Reed-Muller codes, and all sys-

\(^2\) S. Lloyd (1957) has investigated for which values of $n$ and $k$ there exist strictly $e$-error-correcting codes (codes which correct all $e$-errors with weights $\leq e$ and only these errors). The result is that for $n \leq 150$, the largest nontrivial code with the largest multiple error-correcting capabilities which satisfies the Hamming bound as an equality, is a code due to Golay (1949), which has $n = 23$, $k = 12$ and corrects all triple errors and all errors of less weight.
tematic codes (see Slepian, 1956). Further it is very easy to instrument. We will base on it the following discussion of error correction in real operations.

4.22. Automatic Error Correction for Real Operations

The method of applying redundancy to several quantities in common is sometimes efficient when all the quantities represent a word to be operated upon according to some law. We have already seen in Sec. 4.21 the effectiveness of the method when the operation is a pure transmission.

We will now see that the method cannot be used alone when dealing with real operations. It is sufficient to consider the operation of addition of two words $a$ and $b$ ($a + b = s$). Let all three numbers be of $n$ binary positions, with the information part in the first $k$ positions:

$$a = q_1^1, q_2^1, \ldots, q_k^1, \ldots, q_n^1$$

$$b = q_1^2, q_2^2, \ldots, q_k^2, \ldots, q_n^2$$

$$s = q_1^3, q_2^3, \ldots, q_k^3, \ldots, q_n^3$$

According to (9) we have for $a$ and $b$

$$q_i^1 = \sum_{j=1}^{k} \gamma_{ij}q_j^1 \pmod{2} \quad (i = k + 1, \ldots, n) \quad (17)$$

$$q_i^2 = \sum_{j=1}^{k} \gamma_{ij}q_j^2 \pmod{2} \quad (i = k + 1, \ldots, n) \quad (18)$$

Forming the sum $s$ we obtain

$$q_i^3 = q_i^1 + q_i^2 + c_i \pmod{2} \quad (i = 1, \ldots, k) \quad (19)$$

$$c_i = \text{carry to the } i:\text{th position}$$

$$q_i^3 = \sum_{j=1}^{k} \gamma_{ij}q_j^3 \pmod{2} \quad (i = k + 1, \ldots, n) \quad (20)$$

The sum can be formed as follows:

1a. Transfer $a$ to the accumulator.

1b. Correct eventual errors in the accumulator according to (10).

2a. Transfer and add (19, 20) $b$ to the content of the accumulator.

2b. Correct eventual errors in the accumulator according to (10).

With this procedure, errors in both $a$ and $b$ are corrected (provided the carries are correct). The sources of the errors being corrected may be in $a$ or $b$, or in their transfer into the accumulator.
The addition operations (19) and (20) are also corrected. However the operation of forming the carries is not corrected, since (19) and (20) are formed whether the carries $c_j$ are correct or not. This is a severe limitation, inasmuch as an error in a single carry position may result in a multiple error in the sum.

Since the generation of carries ($c_{i-1} = q_i \cdot q_{i+1} + q_i c_i + q_i c_i \mod 2$) is not easily corrected with checks of type (9), we assume that they are corrected individually by redundancy of constructional parts (diodes, for example) according to Sec. 4.11 (microlevel redundancy).

In the above we have considered the generation of the check sequence (10) and the correction $e_i(e_i a_i) = a_i$ ideal (compare end of Sec. 4.2). We will now assume that these processes are corrected by microlevel redundancy (Sec. 4.11) to a degree which corresponds to the above word-correction.

### 4.3. Applications on a Sequence Computer

Of the four methods

i. Word correction (Sec. 4.2)

ii. Separate correction (Sec. 4.12)

iii. Multiplexing (Sec. 4.12)

iv. Microlevel corrections (Sec. 4.11)

we have seen that i and ii cannot be applied alone but have to be complemented with iv if an arbitrary, low probability of error is to be reached. In i and ii, the required connecting organs must be protected from errors with method iv. Besides, when i is applied to real operations also some complementary circuits (Sec. 4.22) must be protected with iv.

Methods iii and iv may be used alone. They are applied individually to each basic element of an ideal automata structure and are hence applicable to any type of computer. Because of this general applicability, the methods are quite uneconomical for use in a computer of some specific character as, for instance, a sequence computer. In such a computer many pure transmitting operations of words (specific collections of figures) occur. This fact can be taken advantage of by an application of method i in appropriate parts of the computer.

The reason why method iv can be used alone is that redundancy is here supplied on microelements (diodes, relay-contacts) for which a connecting organ is unnecessary (corresponds to a simple wiring-connection of the microelements; this wiring is considered error-free).

We will now see how method i + iv can be applied on a universal sequence computer.
As we have mentioned in Sec. 2, a universal (Turing) machine can be designed with a finite number of internal states, but has to be supplied with an infinite memory capacity (infinitely long tape). In a practical machine an infinite memory naturally is an impossibility. In limiting the memory capacity to some finite value, the possibilities of the machine will also be limited.

Still the large electronic computers of today are termed "universal." This evidently means that they can compute any programmed sequence which is permitted by their memory capacity. A usual value of this capacity is $5 \cdot 10^5$ binary places (usually divided into an internal part of low access time and an external part).

This capacity is large compared to the number of basic elements in the rest (see Fig. 4) of the computer (which generate the number of internal states). The number of these elements usually is of the magnitude of $10^8$.

Because of this difference in magnitude of elements in the memory and in the rest of the computer, and because of the different kinds of operations performed in these two parts, we suggest an automatic error correction according to Fig. 4.

The memory is supplied with a comparatively small number of re-

![Fig. 4. Automatic correction in a universal sequence computer.](image-url)
dundant positions (check-positions) according to I, but with no arrangements for automatic correction.

All words of the memory have to pass the instruction register or the arithmetic element (accumulator). It is therefore sufficient to perform the automatic correction of eventual errors in these two places according to i + iv. Thus errors in the memory content, in transferring the memory content, and in the operations of the accumulator are all corrected, provided the control is error free.

The control is corrected according to iv.

Further, the input and output of the machine have to pass the accumulator, and are thus also corrected, provided the input tape is punched with (the comparatively small amount of) word redundancy. Evidently, errors in punching will also be corrected.

In this way the probability of error can be diminished to an arbitrarily low value. The necessary amount of redundancy is economically applied in the largest part of the machine, the memory. The percentage of redundancy in the remaining part of the machine is larger.

The necessary amount of redundancy for a certain total probability of error and a certain computation depth can be determined according to the above, when the probabilities for temporary errors in the different components are known. This amount of redundancy is by no means claimed to be as small as possible (although it probably is not very far from this limit).

For pure transmission and ideal checking and correction, Shannon (1948) has given a theoretical limit.

Also for pure transmission and ideal checking and correction Slepian (1956) shows that for certain values of n and k there exist no better codes than the best of the above (Sec. 4.21).

In a forthcoming paper we plan to investigate the question of the least amount of redundancy for a certain probability of total error in the general case: Specified computation process of elementary real operations with no assumptions of some operations being error free.3

In a theory of automata with capabilities of correcting not only

3 Moore and Shannon (1956) have treated the relay aspect of the von Neumann (1956) Sheffer stroke problem. Both treatments are, as Moore and Shannon point out, limited in the sense that redundancy is only supplied on elementary cells in a network structure, but this structure in itself is not the subject of any modifications. The application of redundancy made in Sec. 4.3 might be thought of as a structure alteration.
temporary component errors but also stationary errors (errors being permanent after a statistically distributed time), the concept of self-reproduction (Sec. 3) will probably play an important role.

Received: May 10, 1957.

References


