XII. PROCESSING AND TRANSMISSION OF INFORMATION

Prof. P. Elias
Prof. R. M. Fano
Prof. D. A. Huffman
Dr. M. V. Cerrillo
W. G. Breckenridge
E. Ferretti
K. Joannou
J. B. O'Loughlin
L. S. Onyshkevych
A. J. Osborne
E. T. Schoen
F. F. Tung
S. H. Unger
M. L. Wetherell

RESEARCH OBJECTIVES

The fundamental problem of the transmission of information is divided in two parts. The first can be described loosely as the task of eliminating redundancy in natural messages – for instance, in speech and in pictures – in order to achieve an economical representation. This has been called the problem of bandwidth compression. The second can be described loosely as the operation of putting redundancy into an economically represented message in order to match it to a noisy channel and to improve the reliability of communication in the presence of random disturbances. These two aspects of the problem are intimately related to the two fundamental results of information theory.

In both of these tasks, switching theory is needed to implement the required coding and decoding devices; however, it is our feeling that there are more fundamental relationships between switching theory and information theory. On the one hand, there are questions arising from tracing information flow in switching networks, and defining such networks as information-lossless or information-lossy. On the other hand, there are strong connections between the problem of sending correct information over a noisy channel and the problem of building reliable switching devices from unreliable components. These are the reasons that led to the merging last year of the switching theory and information theory groups of this Laboratory.

Of necessity, a full investigation of the properties of switching circuits must be undertaken before we shall know which of these properties should be exploited to make coders for the transmission of information. During the past year a study of the special properties of linear switching circuits began, and some means for their orderly realization were found. These linear circuits are especially promising because they are information-lossless and because their force-free responses form a set of multiple-error correcting codes and are nearly noiselike in their correlation properties.

A way of designing switching circuits from elements having one particular kind of unideal behavior was found last year and described in a paper, "The Design and Use of Hazard-Free Switching Networks" (presented at the A.C.M. meeting, September 1955). In this paper it was shown that lack of equality in the reaction times of several switching elements that are controlled by a single switching signal could be compensated for by additional paralleled elements with proper control. Success in solving this problem leads us to hope for the achievement of ways of compensating for other kinds of unideal behavior.

Research on the measurement of probability distributions of picture intensities was completed last year and will be reported in two forthcoming technical reports (No. 296, by J. Capon, and No. 302, by J. C. Stoddard). Further measurements of the statistics of pictures and efforts at bandwidth compression continue. A new program of research on picture processing has been undertaken. The objective of this program is to treat pictures in such a way that detail is reduced without displeasing the observer (for example, by eliminating photographic grain and still leaving sharp edges). This work will include the examination of the effect on two-dimensional signals of some operations, such as filtering, clipping, and noise-suppression, which are already familiar in their application to one-dimensional time functions.

Two devices for speech compression are under investigation. One is a modified vocoder, designed by Müller (Quarterly Progress Report, April 15, 1954, p. 46) the construction and testing of which are being completed. The other is a device that transmits only one pitch-period of a speech waveform and repeats it several times at the receiver. Some results from the second of these devices, and perhaps from the

first, should be forthcoming.

Work on coding for noisy channels continues. During the past year there was an appreciable amount of progress in the theory of this field, in relating the transmission rate, the channel capacity, the error probability of received messages, and the permissible time delay in coding and decoding. The task of constructing coders and decoders, generally speaking, remains difficult, but for one kind of channel, the erasure channel, it is becoming manageable. The possibility of making use of the erasure channel to combat the fading problem in scatter transmission is being investigated.

P. Elias, D. A. Huffman

A. TIME-VARYING LINEAR BINARY CIRCUITS

Binary circuits synthesized from modulo-two adders and unit delays that may contain feedback loops have been shown to be information-lossless networks and to exhibit response properties similar to those of conventional linear electric networks (1). Perhaps their most interesting characteristic is the possibility of exciting in them force-free, steady-state periodic outputs. These outputs correspond to the nondecaying steady-state responses of an electrical filter composed only of inductances and capacitances.

A more general class of linear binary circuits utilizes not only delays and modulo-two adders but also a "commutator" element. The latter element is represented in Fig. XII-1. Its input is connected to each one of N outputs once every cycle of N units of time. We define alternately N corresponding multiplicative operators $C_o, C_1, C_2, \ldots, C_j, \ldots, C_{N-2}, C_{N-1}$ so that only one at a time is equal to unity and so that $C_{j+1} = 1$ just one unit of time after $C_j = 1$. (Note also that $C_o = 1$ one time unit after $C_{N-1} = 0$.) Thus, $C_j C_k = \begin{cases} 0 \text{ if } j \neq k \\ 1 \text{ if } j = k \end{cases}$.

The product of operators C_j and D^q (D^q is a delay of q time units) represents a cascading of the corresponding circuit elements. It may be reasoned physically (see Fig. XII-2) that $C_j D^q \neq D^q C_j$ because the binary signal $D^q C_j X$ cannot possibly equal unity unless the commutator wiping arm is in the $(j+q)^{th}$ position. (Since the commutator arm starts another cycle after the (N-1) position, the number j + q must be considered to have been taken modulo N.)

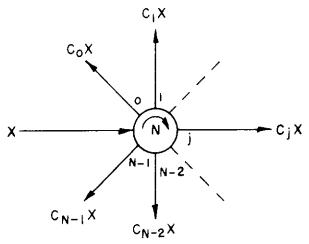

Fig. XII-1. A "commutator element."

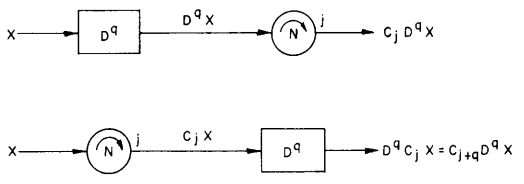

Fig. XII-2. Illustrating that $C_j D^q \neq D^q C_j$.

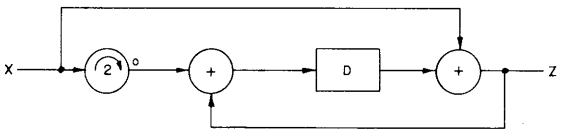

Fig. XII-3. Circuit realization of $Z = X + D(Z + C_0X)$.

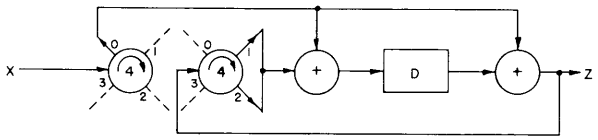

Fig. XII-4. Realization of $\frac{Z}{X} = C_0 + C_1D + C_2D^2 + C_3D^3$.

The terminal description of a time-varying linear binary circuit can be accomplished by means of a transfer function written as the ratio of two polynomials in the delay operator D and whose coefficients are C operators. Particular care must be taken in the algebraic synthesis procedure, since the C operators form a "ring" (the cancellation law does not always hold) and the order of forming products is important.

The equations below represent steps in the complete synthesis of a typical circuit. The derived circuit diagram is given in Fig. XII-3. The circuit has a "behavior-period" of two units of time; thus there are only two C operators: C_0 and C_1. The transfer ratio is to be

$$\frac{Z}{X} = \frac{C_0D^2 + D + C_1}{C_0D^3 + C_1D^2 + I}$$

which may be reduced to

$$\frac{Z}{X} = \frac{\left(C_1D^2 + D + I\right)\left(C_0D^2 + D + C_1\right)}{\left(C_1D^2 + D + I\right)\left(C_0D^3 + C_1D^2 + I\right)} = \frac{C_1D + C_1}{D + I} = \frac{C_1(D+I)}{D + I}$$

(the common right-factor (D+I) may not be cancelled!) and

$$Z = DZ + C_1DX + X = X + D(Z + C_0X)$$

Time-varying circuits of this class can be information-lossy, and for this reason may not have inverses. Their responses to input impulses need not be periodic (as in the circuits of ref. 1) but may actually terminate after a finite interval of time.

The algebra of time-varying binary circuits allows us to express in compact form procedures that are likely to be useful in the coding of information to be transmitted through noisy binary channels. For instance, a circuit that takes every fourth input and repeats it four times has a transfer ratio

$$\frac{Z}{X} = C_0 + C_1D + C_2D^2 + C_3D^3$$

We can show algebraically that only one delay element is necessary for synthesis. Thus,

$$\frac{Z}{X} = C_o + C_1 D + C_2 D^2 + C_3 D^3 = \frac{C_o}{(C_1 + C_2 + C_3) D + I} = \frac{C_1 D + C_o}{(C_2 + C_3) D + I}$$

The last ratio can be realized in the form given in Fig. XII-4.

D. A. Huffman

References

1. D. A. Huffman, The synthesis of linear binary coding networks, Third Symposium on Information Theory, London, September 1955.

B. DESIGN OF LINEAR SEQUENCE FILTERS FOR ARBITRARY IMPULSE RESPONSE

The discussion that follows demonstrates that any single-input, single-output, non-time-varying linear sequence filter can be synthesized from delay elements and modulo-two adders only. A similar conclusion holds for (possibly) time-varying circuits, including those that have more than one input and more than one output.

We consider that a linear sequence filter is completely specified if its response to an input sequence that contains but a single "1" symbol (the "impulse") is given. In order for the filter to be physically realizable the 1's in the response sequence cannot occur before the input "impulse" occurs. Furthermore, if the response is to be that of a finite filter the impulse response must ultimately become zero or become periodic with a finite period.

As an example, note that the arbitrarily specified impulse response, Z^*, of Fig. XII-5 can be resolved into two parts. One of these (Z_p^*) is the ultimately periodic response behavior extended back to the time of the input impulse. The other component (Z_t^*) is the transient part of the response and is chosen so that $Z_p^* + Z_t^* = Z^*$.

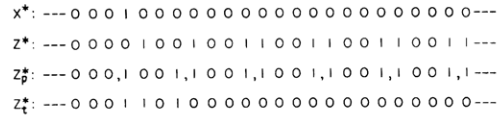

Fig. XII-5. Resolution of an impulse response into periodic and transient components.

The transient impulse-response (Z_t^*) can be considered a consequence of a subfilter that has feed-forward paths only and of the form given in Fig. XII-6a. The periodic response (Z_p^*) can be considered a consequence of a subfilter (Fig. XII-6b) that is resolved into two parts: the first has a single feedback path and produces from the input impulse a train of 1's at equally spaced intervals equal to the ultimate repetition period of the response; the second

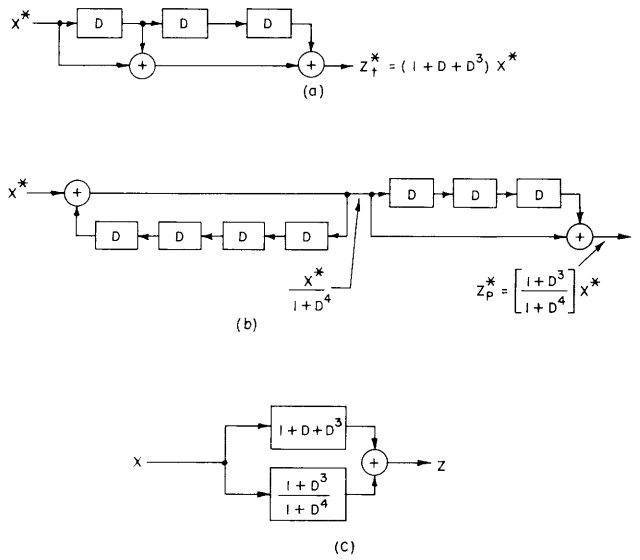

Fig. XII-6. Synthesis of a binary filter to meet impulse response specifications.

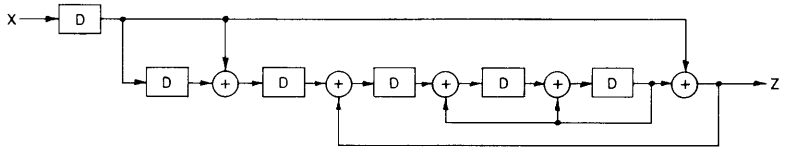

Fig. XII-7. Chain realization of a binary sequence filter.

produces, from each 1 symbol put into it, a finite response equal to the (repeated) typical cycle of the ultimate response.

The composite filter that has the desired impulse response is given in Fig. XII-6c. This filter will not, in general, be economical in its use of delay elements. The following algebraic simplification procedure and the final circuit of Fig. XII-7 are one further example of the usefulness of an algebra of polynomials in a delay operator:

$$\frac{Z}{X} = I + D + D^3 + \frac{I + D^3}{I + D^4} \qquad \frac{Z}{X} = \frac{D(I + D^3 + D^4 + D^6)}{I + D^4}$$

$$\frac{Z}{X} = \frac{D(I + D + D^2 + D^4 + D^5)}{I + D + D^2 + D^3}$$

$$Z + DX = DZ + D \cdot DX + D^2Z + D^2 \cdot DX + D^3Z + D^4 \cdot DX + D^5 \cdot DX$$

$$Z + DZ = D\left\{Z + DX + D\left\{Z + DX + D\left\{Z + D\left\{DX + D \cdot DX\right\}\right\}\right\}\right\}$$

The design of the more general multi-input, multi-output, time-varying circuit will be the subject of a forthcoming paper, "An Algebraic Synthesis Procedure for Linear Binary Sequence Filters." The effect of a single input digit at one of the inputs of such a filter may be "smeared" over the various output sequences in an arbitrarily specified manner, hence it will be foreseen that such filters (especially when the number of outputs exceeds the number of inputs) may be particularly useful in the redundant coding of information so that it will be less vulnerable when transmitted through noisy channels.

D. A. Huffman

C. TIMING CONSIDERATIONS IN CLOCKED SWITCHING CIRCUITS

In a sequential switching circuit, either clocked or synchronous, the problem of race conditions is solved by permitting changes of state to occur only during certain limited, and generally periodic, time intervals defined by pulses generated for this purpose. To secure reliable system operation certain constraints must be satisfied by relevant circuit parameters, the duration and period of the clock pulses, and the tolerances involved. In the discussion of these relations an idealized model of a clocked system will be used and a procedure for obtaining the maximum operating speed will be developed.

Figure XII-8 shows the kind of system that is being studied. The box outlined by dotted lines can be considered one of several binary memory elements that can be realized physically in many different ways.

The system operates as follows. Between clock pulses there is no output and there is no connection between the combinational circuit and the inputs to the memory elements. The local feedback loop around each amplifier holds the output of the corresponding D_1 constant. When a clock pulse (CP) arrives, the combinational circuit operates on both the system input (assumed to be at the proper level when the CP goes on) and the output of the memory to produce both the system output and a set of input states fed to the memory elements through the multiplier gates controlled by the clock pulses. The delays D_1 and D_2 keep the memory output from changing while the CP is on. (D_2 alone would suffice; D_1 is the unavoidable delay associated with the amplifier.) Notice that in this mode of operation the delays serve a secondary purpose, the principal storage function being accomplished by the feedback loop.

The symbols used are as follows: T_{D1}, T_{D2}, and T_c represent the average values of the delays, D_1 and D_2, and the CP width, respectively. The minimum allowable time between leading edges of adjacent clock pulses is T_p. The maximum normalized deviations of D_1, D_2 and the pulsewidth are designated E_{D1}, E_{D2}, and E_c, respectively. For example, the delays D_1 are assumed to be within the range from $T_{D1}(1 - E_{D1})$ to

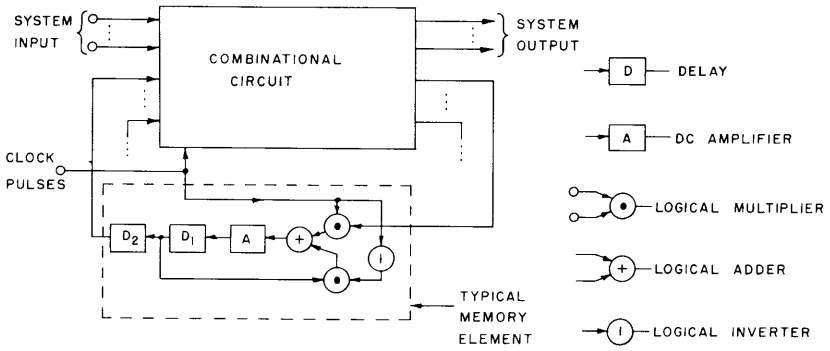

Fig. XII-8. Clocked sequential circuit.

$T_{D1}(1 + E_{D1})$ under all conditions. Obviously, the E's must lie between zero and one.

It is appropriate to inquire what relations must exist among these quantities if the system is to operate properly regardless of the number of memory elements that change their states simultaneously. First of all, the inputs to the combinational circuit from the memory must not change while the CP is on:

$$T_{D1}(1 - E_{D1}) + T_{D2}(1 - E_{D2}) \geq T_c(1 + E_c) \tag{1}$$

Furthermore, the outputs of the D_1 elements must change before the CP ends (if they are to change at all):

$$T_{D1}(1 + E_{D1}) \leq T_c(1 - E_c) \tag{2}$$

Finally, enough time must be allowed between the leading edges of successive CP's to permit the effects of changes of state while the first CP is on to emerge from the appropriate D_2's before the next CP starts:

$$T_{D1}(1 + E_{D1}) + T_{D2}(1 + E_{D2}) \leq T_p \tag{3}$$

In designing this kind of system it is necessary to find T_{D2}, T_c, and T_p as functions of T_{D1} and to determine the estimated tolerances. Generally, it is desirable to minimize T_p so as to maximize the operating speed, which would indicate that the equality sign should be used in Eq. 3. Next, solve Eq. 1 for T_{D2} to obtain

$$T_{D2} = \frac{T_c(1 + E_c) - T_{D1}(1 - E_{D1})}{1 - E_{D2}} \tag{4}$$

The equality sign in Eq. 4 minimizes T_{D2}, which is a necessary step in minimizing T_p, as shown by Eq. 3. Substituting Eq. 4 in Eq. 3 and simplifying the new expression yields

$$T_p = \frac{2(E_{D1} - E_{D2})}{1 - E_{D2}} T_{D1} + \frac{(1 + E_{D2})(1 + E_c)}{1 - E_{D2}} T_c \qquad (5)$$

This in turn shows that T_c should also be made as small as the constraint of Eq. 2 will allow. Substituting that minimum value in Eq. 5 reveals that the smallest allowable T_p is equal to an unpleasant-looking, positive constant multiplied by T_{D1}. Thus T_{D1} should be minimized; and T_c, T_{D2}, and T_p can be found by successive substitutions in Eq. 2 (using the equality sign), Eq. 4, and Eq. 3 (using the equality sign), respectively.

To take an example, suppose that D_1 is conservatively estimated to lie between 0.06 μsec and 0.14 μsec. Then $T_{D1} = (0.06 + 0.14)/2 = 0.1$ μsec, and $E_{D1} = (0.1 - 0.6)/0.1 = 0.4$ μsec. Furthermore, let us assume that D_2 can be constructed to operate reliably with a tolerance of ±20 per cent and that deviations in the clock pulsewidth can be held within ±15 per cent. Then $E_{D2} = 0.2$, and $E_c = 0.15$. From Eq. 2 we find $T_c = \left[(1 + E_{D1})T_{D1}\right]/1 - E_c = 0.165$ μsec. Equation 4 and the previous result indicate that T_{D2} should equal 0.163 μsec; combining these results with Eq. 3 yields $T_p = 0.34$ μsec.

Thus we find that the maximum operating rate of this system would be obtained by using a value of 0.163 μsec for D_2 and a clock source that produced pulses 0.165 μsec in width at a rate never exceeding one every 0.34 μsec.

S. H. Unger