Also, the lower limit on the advance current, I_0, is the minimum value that allows intersection of curves 2) and 3).

The reset current must be large enough to make the gate resistive even at low values of gate current. Although lower values can be used, a practical upper limit is that current for which the portion of the gate directly under the grid is completely resistive. This is the saturation value of the gate resistance, and although the reset current could be safely increased much higher, this would serve no purpose since the gate resistance is essentially at its maximum value,[1] and hence the decay time of the stored current is minimized.

High-Speed Operation

The above analysis assumes that the current distribution has reached equilibrium when any of the current sources is switched on or off. It is usually adequate to allow 3τ for a stored current to switch an advance current, where τ is the L/R time constant of the circuit. An equal amount of time is sufficient for a reset current to destroy a stored current. Thus, the advance current only has to be switched on for a time 6τ—3τ to allow the stored current of the previous stage to switch the advance currents and 3τ to allow the reset current of the previous stage to destroy its stored current.

The value of τ calculated from the measured resistance and the calculated inductance of the circuit is 0.33 microsecond. This compares favorably with the accurately observed value[1] of 0.38 microsecond. An order of magnitude value for τ can also be obtained by observing the output signal of the storage cell of Fig. 1, which is shown in Fig. 4. The output signal was amplified by a Tektronix type-121 wideband preamplifier in cascade with a type-541 Tektronix oscilloscope. Amplifier oscillation is responsible for the noise.

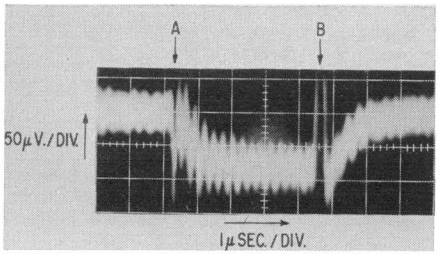

Fig. 4—Output voltage of a CFC. I_0 and I_0' (see Fig. 1) are on at the start of the trace. i_2 is switched on at A, switching I_0, and I_0 is switched off at B, allowing i_2 to destroy the stored current.

If we know τ, the highest speed of operation of the shift register can be calculated. Since the advance pulses can overlap each other in the three loops per bit mode, the minimum cycle time is 9τ, or 3.42 microseconds. This corresponds to an information rate of just under 300 kc.

All the testing was performed using the four-loops-per-bit mode of Fig. 2. Here the advance pulses cannot be overlapped and the minimum cycle time for the reset and advance pulse train is 12τ, or 4.56 microseconds, a rate of over 200 kc. The circuit has functioned perfectly at an advance pulse rate of 140 kc. In this experiment, the advance pulses were 3 microseconds long and the reset pulses were 1.2 microseconds long. All the current sources were 300 ma and the critical gate current was 350 ma. Higher speeds have not been attempted because of instability in the timing of the pulse equipment. The information rate in this mode is, of course, one half the advance pulse rate.

Although the register has only seven gates, longer registers can be formed by connecting the seven gate registers in cascade. This has been done successfully, but because of the unshielded interconnecting wires, the speed of operation is considerably reduced.

Output Waveforms

Fig. 5 shows the output pulses obtained by use of a 500-kc bandwidth, 50-db gain transistor pre-amplifier as an input to a 541 Tektronix oscilloscope.

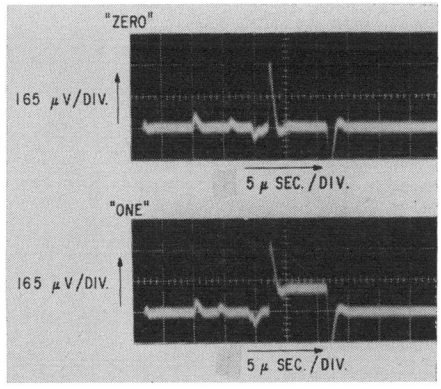

Fig. 5—Output waveforms of a CFC shift register.

The output gate of the register was pulsed with a 10-microsecond, 150-ma output pulse. The "ZERO" shows the output voltage with no circulating current in the last loop of the register. The noise was caused by radiation from the advance and reset pulses and the pulsing of the finite inductance of the output gate.

The "ONE" shows the 130-microvolt output signal caused by a stored current of 150 ma in the last loop. All the advance and reset pulses were 150 ma and the critical gate current was 200 ma.

Acknowledgment

The CFC shift register described in this letter and many, many others were produced by H. E. Tanis and D. MacKellar. Their contribution is gratefully acknowledged.

H. H. Edwards
General Electric
Light Military Electronics Dept.
Ithaca, N. Y.
V. L. Newhouse
General Electric Research Lab.
Schenectady, N. Y.
J. W. Bremer
General Electric Computer Dept.
Palo Alto, Calif.

A Note on Optimum Pattern Recognition Systems*

Chow[1] has shown that, for a given rejection rate, the error rate in a recognition system is minimized if the following decision criterion is used.

Choose class k if

$$p_k F(v \mid a_k) \geq p_j F(v \mid a_j) \qquad \text{for all } j \neq k$$

and

$$p_k F(v \mid a_k) \geq \beta \sum_{i=1}^{c} p_i F(v \mid a_i) \qquad 0 \leq \beta \leq 1;$$

reject the pattern if

$$p_j F(v \mid a_j) < \beta \sum_{i=1}^{c} p_i F(v \mid a_i) \quad \text{for all } 1 \leq j \leq c.$$

Here p_i is the *a priori* probability of the occurrence of class i, $F(v|a_i)$ is the conditional probability of making the measurement v given that a member of class i is present, c is the number of pattern classes and β is a constant chosen to force the system to meet the given rejection rate.

The proper value for β is generally difficult to determine, and an empirical approach may often be necessary. However, there is one important case in which β may be determined analytically, the discussion of which follows.

Let the cost of misrecognizing a pattern, of rejecting a pattern, and of correctly recognizing a pattern be independent of the pattern class. In particular, let

$$w_{ij} = w = \text{cost of misrecognition}$$
$$w_{io} = w_o = \text{cost of rejection}$$
$$w_{ii} = 0 = \text{cost of recognition}$$

where

$$w > w_o > 0.$$

(Since a Bayes criterion is being used, no generality is lost by setting $w_{ii}=0$.[2]) The general loss function is given by[1]

$$R(p, \delta) = \sum_{i=1}^{c} \sum_{j=0}^{c} \int_v \delta(d_j \mid v) p_i w_{ij} F(v \mid a_i) dv$$

where $\delta(d_j|v)$ is the probability that class j will be decided given the measurement v (d_o is the rejection decision), and $R(p, \delta)$ is the loss associated with the *a priori* class distribution p and a given decision function δ.

Using the above cost schedule, the loss function may be written

$$R(p, \delta) = \int_v \sum_{j=0}^{c} \delta(d_j \mid v) \sum_{i=1}^{c} w p_i (Fv \mid a_i) dv - \int_v \sum_{i=1}^{c} \delta(d_i \mid v) w p_i F(v \mid a_i) dv - \int_v \delta(d_o \mid v) \sum_{i=1}^{c} (w - w_o) p_i F(v \mid a_i) dv.$$

* Received by the PGEC, January 23, 1961.

[1] C. K. Chow, "An optimum character recognition system using decision functions," IRE Trans. on Electronic Computers, vol. EC-6, pp. 247–254; December, 1957.

[2] H. Chernoff and L. E. Moses, "Elementary Decision Theory," John Wiley and Sons, Inc., New York, N. Y., ch. 5; 1953.

Noting that

$$\sum_{j=0}^{c} \delta(d_j \mid v) = 1,$$

$$\sum_{i=1}^{c} p_i = 1,$$

$$\int_v F(v \mid a_i)dv = 1,$$

the first integral can be reduced, allowing the cost function to be written as

$$R(p, \delta) = w - w \int_v \sum_{i=1}^{c} \delta(d_i \mid v) p_i F(v \mid a_i)dv$$

$$- (w - w_o) \int_v \delta(d_o \mid v) \sum_{i=1}^{c} p_i F(v \mid a_i)dv.$$

To minimize $R(p, \delta)$, $\delta(d_i \mid v)$ is chosen as follows:

$$\delta(d_k \mid v) = 1, \qquad k \neq 0,$$

if

$$p_k F(v \mid a_k) \geq p_j F(v \mid a_j) \qquad \text{for all } j \neq k$$

and

$$p_k F(v \mid a_k) \geq \left(\frac{w - w_o}{w}\right) \sum_{i=1}^{c} p_i F(v \mid a_i);$$

$$\delta(d_o \mid v) = 1$$

if

$$p_j F(v \mid a_j) < \left(\frac{w - w_o}{w}\right) \sum_{i=1}^{c} p_i F(v \mid a_i)$$

for all $1 \leq j \leq c$.

But this decision criterion is of the same form as that derived by Chow for the case of minimum error rate given a fixed rejection rate, with

$$\beta = \frac{w - w_o}{w}.$$

Therefore, minimizing the cost in the case of constant costs also minimizes the error rate for the rejection rate which corresponds to the above β.

W. H. Highleyman
Bell Telephone Labs., Inc.
Murray Hill, N. J.

Minimal Characterizing Experiments for Finite Memory Automata*

In a recent article by Gill,[1] it has been shown that minimal characterizing experiments for automata of memory (length) M may be derived from a certain *sequential system*. In this note, this idea is carried slightly further so as to make more explicit the relationship between finite memory automata and the corresponding minimal characterizing experiments.

Consider the *maximal* machine of memory M, namely the one that actually contains 2^M states in its minimal Huffman flow table. As an example originally given by Simon,[2] Table I is the flow table of the maximal automaton of memory 2. One may convert this table into the Moore state transition table shown as Table II, by virtue of the following replacements:

$$0, a \rightarrow 0 \qquad 2, e \rightarrow 1$$
$$0, b \rightarrow 4 \qquad 2, f \rightarrow 5$$
$$1, c \rightarrow 2 \qquad 3, g \rightarrow 3$$
$$1, d \rightarrow 6 \qquad 3, h \rightarrow 7.$$

Now the transition diagram of the Moore model may be seen to yield precisely the oriented graph or *sequential system* of eight states illustrated in the Gill article.

TABLE I

	0	1
0	0, a	2, e
1	0, b	2, f
2	1, c	3, g
3	1, d	3, h

TABLE II

	0	1	
0	0	1	a
1	2	3	e
2	4	5	c
3	6	7	g
4	0	1	b
5	2	3	f
6	4	5	d
7	6	7	h

The above example illustrates, then, the following statement. The *cyclic minimal experiments* for the study of automata of memory M are all implicit in the transition diagram of the Moore model for the *maximal* automaton of memory M. As a matter of fact, these experiments may be read off the transition diagram as the sequences of input symbols associated with the arrows of directed, minimal, closed paths that traverse all the nodes or states.

Gill's results, as well as the contents of this note, seem to be direct consequences of Theorem 1 proved by Simon about finite memory automata.

J. E. Mezei
IBM Research
Yorktown Heights, N. Y.

* Received by the PGEC, February 23, 1961.

[1] A. Gill, "Characterizing experiments for finite-memory binary automata," IRE Trans. on Electronic Computers, vol. EC-9, pp. 469–471; December, 1960.

[2] J. M. Simon, "A note on the memory aspect of sequence transducers," IRE Trans. on Circuit Theory, vol. CT-6, pp. 26–29; March, 1959.

On the Size of Weights Required for Linear-Input Switching Functions*

A switching function $f(x_1, x_2, \cdots x_n)$ which takes on its values 0 and 1 according to a linear inequality,

$$f = 1 \Leftrightarrow \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \geq \alpha_0,$$

is said to be a *linear-input* function,[1] also called linearly separated function,[2] threshold function,[3] setting function,[3] and in a more restricted sense, majority decision function.[4] The *weights* $\alpha_i (i = 1, 2, \cdots, n)$ and *threshold* α_0, which we may take to be integers without loss of generality, completely define any particular function of this type.

Several digital devices have been conceived which can realize any linear-input function with a single device. Although only a small fraction of all switching functions are linear-input functions, it is known that an arbitrary combinational switching function can be realized in a network of such devices. The potential simplicity of such networks in comparison with those composed of conventional digital elements justifies a more detailed study of this class of functions.

An important preliminary question concerns the maximum size of the weights α_i which may be required for the realization of arbitrary linear-input functions. In this note it is shown that one must be prepared for rather large weights, approaching a value which grows at a rate no less than $2^n/n$ for an n-input device, as n becomes large. This fact suggests in turn that if the nature of the device imposes any restriction on the maximum size of input weights, this will probably increase substantially the number of such devices required for the realization of arbitrary switching functions.

We lead up to this bound (Theorem 2) through a weaker theorem, but one whose proof is much simpler.

Theorem 1

Some linear-input functions of $n = 2q+1$ variables require weights at least as large as 2^q in any realization with integral weights.

Proof: Let a linear-input function f_1 be defined by

$$f_1 = 1 \Leftrightarrow \sum_{i=1}^{2q+1} \alpha_i x_i \geq \alpha_0,$$

in which the weights are

$$\alpha_1 = 1, \qquad \alpha_{q+2} = 2^q - 1,$$
$$\alpha_2 = 2, \qquad \alpha_{q+3} = 2^q - 2,$$
$$\alpha_3 = 4, \qquad \alpha_{q+4} = 2^q - 4,$$
$$\cdots\cdots \qquad \cdots\cdots\cdots$$
$$\alpha_q = 2^{q-1}, \qquad \alpha_{2q+1} = 2^q - 2^{q-1},$$
$$\alpha_{q+1} = 2^q, \qquad \alpha_0 = 2^q.$$

* Received by the PGEC, November 28, 1960; revised manuscript received, March 13, 1961.

[1] R. C. Minnick, "Linear-input logic," IRE Trans. on Electronic Computers, vol. 10, pp. pp. 6–16; March, 1961.

[2] R. McNaughton, "Unate Truth Functions," Appl. Math. and Stat. Lab., Stanford University, Stanford, Calif., Tech. Rept. No. 4; October, 1957.

[3] M. C. Paull and E. J. McCluskey, Jr., "Boolean functions realizable with single threshold devices," Proc. IRE, vol. 48, pp. 1335–1337; July, 1960.

[4] S. Muroga, "Logical elements on majority decision principal and complexity of their circuit," *Proc. Internatl. Conf. on Information Processing*, Paris, France, June, 1959, UNESCO, Paris, 1960, Paper G.2.10, pp. 400–497; 1960.