XXVI. NEUROPHYSIOLOGY*

Academic and Research Staff

Dr. W. S. McCulloch
Prof. J. Y. Lettvin
Prof. P. D. Wall
Prof. M. Blum
Prof. J. E. Brown
Prof. P. R. Gross
Prof. B. Pomeranz
Dr. T. O. Binford
Dr. S-H. Chung
Dr. J. L. Simões da Fonseca
Dr. H. Hartman
Dr. Dora Jassik-Gerschenfeld
Dr. K. Kornacker
Dr. T. McLardy
Dr. R. Moreno-Diaz
Dr. A. Natapoff
Dr. S. A. Papert
Dr. A. Taub
B. Howland
Diane Major
W. H. Pitts
Sylvia G. Rabin

Graduate Students

G. P. Nelson
J. I. Simpson
W. A. Wright

A. SYNTHESIS AND LINEARIZATION OF NONLINEAR FEEDBACK SHIFT REGISTERS – BASIS OF A MODEL OF MEMORY†

Linear shift registers have been extensively studied since the pioneering work of D. A. Huffman. A general theory was created by Huffman,[1] Golomb, Elspas, Stern, Friedland, Zierler, Hartmannis, Massey, and others. This progress was mainly due to the identification of the operations performed by linear switching circuits with algebraic operations over Galois fields. There was no general theory for nonlinear shift registers, however.

We shall present: (i) a general method of linearization of nonlinear feedback shift registers with or without inputs, which allows us to handle them by the theory of linear switching circuits; (ii) an algorithm to specify a nonlinear net with a minimal number of delays that generates in its autonomous behavior a desired output sequence; (iii) a method of information storage that, for the average of possible cases, gives a logarithmic reduction in space occupancy; (iv) a class of events embedded in a time sequence that extends to the infinite past, and for whose computation any desired degree of reliability can be obtained by using methods that are adequate for definite events having the same length as the readout operation; and (v) some consequences concerning memory and a wide class of learning processes.

*This work was supported by the National Institutes of Health (Grant 5 RO1 NB-04985-04), and in part by the U.S. Air Force (Aerospace Medical Division) under Contract AF33(615)-3885.

†This work was done partly at the Instrumentation Laboratory under the auspices of DSR Project 55-257, sponsored by the Bioscience Division of National Aeronautics and Space Administration through Contract NSR 22-009-138.

1. Method of Linearization

Our approach extends the Fukunga[2] method for the autonomous behavior of nonlinear feedback shift registers to the general case of any feedback shift register with any number of inputs through any logical gate.

THEOREM 1: Any feedback shift register with any number of inputs through any logical gate always admits a linearized form by introducing a finite set of new variables.

Proof: Given a net with input variables $x_1, x_2 \dots x_n$ and variables that define the state of the delay elements $y_1, y_2 \dots y_m$, we can always write the expression that gives the next value of any variable corresponding to a delay element

$f(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m)$, in its normal disjunctive form.

Let the following expression be the normal disjunctive form of this function:

$$f_1(x_1, \dots, x_n, y_1, \dots, y_m) = a_o x_1 \dots x_n y_1 \dots y_m V a_1 \bar{x}_1 x_2 \dots y_1 \dots y_m V \dots V a_K \bar{x}_1 \bar{x}_2 \dots \bar{x}_n \bar{y}_1 \bar{y}_2 \dots \bar{y}_m. \tag{1}$$

As the terms in Eq. 1 correspond to mutually exclusive cases, we can substitute $\oplus$ for V, and as for any variable x,

$$\bar{x} = x \oplus 1 \qquad \text{(modulo 2)} \tag{2}$$

and

$$x \oplus x = 0. \tag{3}$$

By making the corresponding substitutions in (1), we obtain

$$\begin{aligned} f(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = b_o 1 \oplus \\ \oplus b_1 x_1 \oplus \quad \oplus b_n x_n \oplus b_{n+1} y_1 \oplus \dots \oplus a_{n+m} y_m \\ \oplus b_{n+m+1} x_1 x_2 \oplus \dots \oplus b_{n+m+K} y_{m-1} y_m \oplus \\ \oplus b_{n+L+1} x_1 x_2 x_3 \dots \\ \vdots \\ \oplus b_r x_1 x_2 \dots x_n y_1 y_2 \dots y_m, \end{aligned} \tag{4}$$

where any b_i is a constant, either 1 or 0, and because of (3) no term appears more than once. We note that this is the "exclusive or canonical form," and therefore, in general, for a net with n inputs and m delay elements in order to linearize it we shall have to introduce $2^{n+m} - (n+m)$ variables. Since the state of the delay elements of the circuit can be expressed by an equation of the form of (4), the state of the circuit will be completely defined by m linearized equations.

In many cases this method is a very efficient way of obtaining Boolian functions in a very simplified form.

Example 1: Let $f(x, y, z) = \bar{x}yzVx\bar{y}zVxy\bar{z}V\bar{x}\bar{y}\bar{z}$.

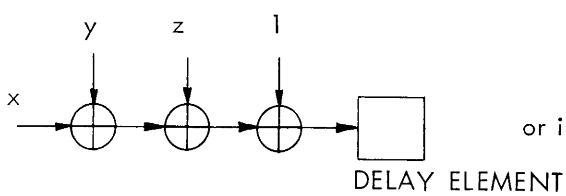
Then by (1) $= \bar{x}yz \oplus x\bar{y}z \oplus xy\bar{z} \oplus \bar{x}\bar{y}\bar{z}$

by (2) $= (x \oplus 1)yz \oplus x(y \oplus 1)\, z \oplus xy(z \oplus 1) \oplus (x \oplus 1)(y \oplus 1)(z \oplus 1)$

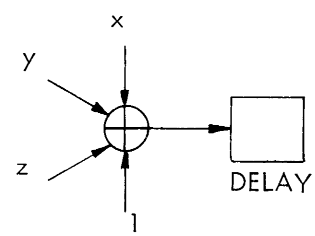
by (3) $= x \oplus y \oplus z \oplus 1;$

that is, simply the constant 1 added to the sum of the original variables; hence, the only "new" variable introduced is a dummy variable whose value is constrained to be 1.

This expression can be implemented in the form



or in abbreviated form



An immediate generalization is the extension of the method of linearization to nonlinear algebraic expressions of n variables where the operations are performed over GF(p), p a prime. In that case in order to linearize the expression it will be necessary to introduce, at most, $p^n - n$ new variables.

a. Application of the Theory of Linear Switching Circuits to the Linearized Feedback Shift Registers with Operations Defined over GF(p), p a prime

The state of a delay element of a linear switching circuit in any given epoch is determined by a linear combination of a subset of the inputs to the net and of the outputs of a specified subset of the delay elements of the circuit

$$V = a_o x_1 \oplus \dots \oplus a_n x_n \oplus a_{n+1} y_1 \oplus \dots \oplus a_{n+m} y_m.$$

The terms that are present after linearization and simplification form two classes, A and B.

$\{A\}$ is a class with terms of the linearized form in which no input variable is present.

$\{B\}$ is a class with terms in which at least one input variable is present.

Let $\vec{S}_j$ be a vector whose n components are the states of the n delay elements of an n^{th}-order shift register in a given epoch, j. This vector defines the state of the net in that epoch.

Let $\vec{I}_j$ be the state of the u external inputs at the end of epoch j, and T be the $n \times n$ matrix that, for every delay element, specifies the subset of delay elements whose state in epoch j determines its state in epoch j + 1.

Let U be the $n \times u$ matrix that, for every delay element, specifies the subset of $\vec{I}_j$ whose state at the end of epoch j determines its state in epoch j + 1.

Then the state of a circuit in any given epoch j + 1 is given by

$$\vec{S}_{j+1} = T \cdot \vec{S}_j \oplus U \cdot \vec{I}_j. \tag{5}$$

Autonomous behavior is defined as the behavior of a net when its input is constantly kept at 0. In that case, Eq. 5 becomes

$$\vec{S}_{j+1} = T \cdot \vec{S}_j.$$

By using the definitions of autonomous behavior and of $\{A\}$ and $\{B\}$, the only terms that contribute to that behavior will belong to $\{A\}$.

The nonlinear n^{th} order net will be embedded in a bigger linear net and the structure of the cycles, transients, and stability states of the nonlinear net will be a subset of the structure of cycles, transients, and stability states corresponding to the linearized net, with the following constraints:

(a) No cycle of the linear net that might correspond to a cycle of the nonlinear net may have a length greater than 2^n, n being the number of the delay elements in the nonlinear net.

(b) All states that belong to transients that lead to the 00...0 state of the linear net may never correspond to a state of the nonlinear net, when, after linearization, a constant 1 remains in the expression that specifies any of the Boolian functions of the linearized net.

(c) The sum of all possible cycles, transients, and stable states of the linear net corresponding to the nonlinear behavior must add to 2^n.

(d) The new terms that belong to $\{A\}$ cannot be supposed to take values incompatible with the values assumed by the initial variables on which they depend.

The structure of cycles to which the preceding constraints apply has been completely handled by Elspas[3] and Friedland[4] for the case of operations modulo p (p any prime) to which all of the preceding arguments may be extended. Elspas obtains that structure from the characteristic polynomial of the matrix that specifies the autonomous behavior.

Consider now the case in which all of the inputs are constantly maintained as 1. It is immediately apparent, for functions over the Galois field (2), that the Boolian expressions that give the state of any of the delays can now be transformed and possibly further simplified by substituting "1" in this expression for all input variables and by applying (3) to it.

The behavior of the net under constant input 1 will be equivalent to the autonomous behavior of the net which is obtained when the substitution and simplification are done.

A new type of behavior can be defined that consists in keeping $n-k$ inputs at value 1, and k inputs at value 0. The adequate substitution will lead, in general, to a different set of expressions for every different assignment of the inputs that are kept constantly at 1. All of those cases can be dealt with as if they corresponded to an autonomous behavior of the net that is obtained after substitution of the corresponding input variables by 1 and after simplification.

As all of the results are valid for operations modulo p (p a prime), any system in which the operations are over a finite set of numbers can be approximated by operations on a Galois field, by mapping from the real numbers that are in one of a set of specified intervals to one element of the field.

When the source statistics are known, Shannon's[5] rate-distortion function specifies the number of p-ary digits per second that are required to represent a given information source within some specified fidelity criterion. A common fidelity criterion is the mean-square error.

2. Minimal Nonlinear Net Specification and Synthesis Algorithm

Let L be the length of an ordered sequence, S, of binary digits $s_L, \ldots, s_2, s_1$.

Let S_n be any subsequence of S, of length n,

$$s_{j+n}, \ldots, s_{j+2}, s_{j+1}.$$

We shall now determine the conditions, given S, to obtain it as the output of a nonlinear feedback shift register with a minimal number of delay elements, in its autonomous behavior.

THEOREM 2: Given any finite sequence S of binary digits, it is always possible to specify a nonlinear feedback shift register with a minimal number of delay elements that produces that sequence as a cycle of its autonomous behavior by means of the following algorithm.

(0) Set $n = 0$. If $s_1 = s_2 = \ldots = s_L$, go to (3); otherwise, set $n = 1$ and go to (1).

(1) Form the L subsequences of length n $S_n(j) = s_{j+n}, \ldots, s_{j+2}, s_{j+1}$ $j = 0, 1, \ldots, L-1$ (where we define $s_{L+i} = s_i$).

(2) For each $i < j$ such that $S_n(i) = S_n(j)$, compare s_{j+n+1} to s_{i+n+1}. If equality always occurs, go to (3); otherwise, increase n by 1 and go to (1).

(3) Stop. The sequence $s_L, \ldots, s_2, s_1$ can be generated from an n^{th}-order nonlinear feedback shift register.

Every $S_n(j)$ can be considered as a state vector of a shift register with digit s_{j+n+1}, as the feedback digit that defines the last component of $S_n(j+1)$, as well as the desired output that corresponds to $S_n(j)$.

It is seen immediately that this interpretation of $S_n(j)$ leads to the definition of the successive states, from $\vec{S}_n(1)$ to $\vec{S}_n(j)$ of a shift register with feedback over the first element that produces S in its autonomous behavior.

For the distinct $S_n(j)$ resulting when the algorithm terminates, the output of the Boolian feedback function is determined by the value of s_{n+j+1}. The value of the Boolian feedback function may be arbitrarily chosen for all other states to completely specify the nonlinear feedback shift register. If $2^n > L$, there exist $2^n - L$ states of the circuit for which the feedback value that corresponds to the more convenient construction can be arbitrarily assigned.

Then, as the behavior of all of the elements of a net is completely specified if all of its states are known, it follows that the circuit that produces S as an output of its autonomous behavior has been completely specified. The procedure further ensures that the circuit has minimal length.

Further steps that lead to an implementation of the net may be outlined as follows. Make a truth table,[6] in which the $\vec{S}_n(j)$ appear in the order of their succession in the sequence, and each $\vec{S}_n(j)$ is put in correspondence with the last digit at the right of $S^{|\dots|}$. We note that the feedback is necessary only to specify the state of the first element of the circuit. All other elements are merely delay elements. If a linearized form is desired, then the normal disjunctive form for any of the elements is obtained from the truth table, and the procedure of linearization and simplification is applied.

A ratio of compression may be defined for any S as

$$R_c = \frac{\log_2 L}{n},$$

where L = length of S.

If $2^n = L$, then

$$R_c = \frac{\log_2 L}{\log_2 L} = 1.$$

THEOREM 3: The upper bounds for the median and average lengths of a nonlinear feedback shift register that has a minimal number of delays and can generate in its autonomous behavior a given purely random binary sequence of length L will be

$$\ell_m < 2 \log_2 L$$

and

$$\ell_{av} < 3 \log_2 L + \frac{1}{2}.$$

Proof: (James L. Massey's proof)

Let $s_1 s_2 \ldots s_\ell$ be the sequence, where the s_i are independent variables, and $\Pr(s_i=1) = \frac{1}{2}$.

Choose any $i < j$, then

$$\Pr\overbrace{(s_i s_{i+1} \ldots s_{i+\ell-2} s_{i+\ell-1} = s_j s_{j+1} \ldots s_{j+\ell-2} \bar{s}_{j+\ell-1})}^{\text{event } E_{ij}} = 2^{-\ell}.$$

That is, the event that the first $\ell - 1$ digits agree, but the last digits differ.

Let

$f(s_1 s_2 \ldots s_L)$ = length of shortest nonlinear FSR that can generate $s_1, s_2 \ldots s_L$.

Then

$\Pr(f > \ell)$ = Pr that $s_1, s_2 \ldots s_L$ contains two identical, except for the last digit in which they differ, length ℓ subsequences, starting in different positions

$$= \Pr(U_j E_{ij}) \leqslant \sum_{i<j} \Pr(E_{ij})$$

$$= \underbrace{(1 + 2 + \ldots + L - 1)}_{\#\text{ of ways of choosing } i,j \text{ with } i<j} 2^{-\ell} = \frac{1}{2}(L-1)\, L \cdot 2^{-\ell}$$

Let ℓ_m = median length required; that is,

$$P(f \geqslant \ell_m) \overset{\Delta}{=} \frac{1}{2}.$$

But

$$P(f \geqslant \ell_m) \leqslant \frac{1}{2}(L-1)\, L\, 2^{-\ell_m}, \qquad \text{by the equation above}$$

$$\therefore \quad \frac{1}{2} \leqslant \frac{1}{2}(L-1)\, L\, 2^{-\ell_m},$$

or

$$2^{+\ell_m} \leqslant (L-1)\, L < L^2,$$

or

$$\ell_m < 2 \log_2 L.$$

Let ℓ_{av} = average length required; that is,

$$\ell_{av} \triangleq \sum_{\ell=0}^{L} \ell \Pr(f=\ell).$$

But for any integer ℓ'

$$\ell_{av} = \sum_{\ell=0}^{\ell'} \ell \Pr(f=\ell) + \sum_{\ell=\ell'}^{L} \ell \Pr(f=\ell)$$

$$\ell_{av} < \ell' \Pr(f<\ell') + L \Pr(f=\ell')$$

$$\leq \ell' + L \Pr(f=\ell').$$

Choosing $\ell' = 3 \log_2 L$ yields

$$\ell_{av} = 3 \log_2 L + L \cdot L \frac{1}{2} L(L-1)\, 2^{-\ell'}$$

$$< 3 \log_2 L + \frac{1}{2} L^3\, 2^{-3 \log_2 L} = 3 \log_2 L + \frac{1}{2}$$

$$\therefore \quad \ell_{av} < 3 \log_2 L + \frac{1}{2}.$$

3. Method of Information Storage

The preceding algorithm gives the possibility of storing a sequence of binary digits not as a stable state but as a pattern of connections and logical relations between delay elements of a shift register.

The basic strategy is this: Rather than designing a neural net to store binary input sequences, one designs a neural net capable of implementing the algorithm of the preceding section and of organizing the minimal length nonlinear feedback shift register. Then when a binary input sequence is processed by the neural net, the resultant nonlinear feedback shift register can be used, upon command, to reproduce the sequence. In other words, a logical structure is stored, rather than the data. We assume that the following conditions are met: (i) The shift register is supposed to be maintained in a 00 . . . 0 state in all the epochs between readout modes, by means of a reset process activated in each epoch, the success of which realizes the state 00 . . . 0 in the shift register. (ii) This reset process has a probability P_1 of being applied in any epoch and probability $Q_1 = 1 - P_1$ of failing to be applied. Furthermore, it is assumed that the reset process acts periodically with period L epochs during the readout mode.

Let p be the probability that the next state of the shift register in the readout mode will be the correct succesor for the shift register specified by the fixation algorithm.

Any such n^{th}-order feedback shift register produces a sequence of binary digits, which together with the digit that corresponds to the first reset operation should constitute a cycle. The probability that the sequence of states of the shift register contains no erroneous successor state is $P_S = p^L$, as follows from the definition of p and from the length of the sequence.

4. Reliability in the Computation of Finite Events Embedded in a Time Sequence That Extends to the Infinite Past

THEOREM 4: Under the specified conditions (i) and (ii), the probability of occurrence of the desired cycle of length L following the reset epoch in the readout mode is at least $P_1 \times p^L$. That is, arbitrary reliability in recalling a sequence from the arbitrarily distant past can be obtained by taking measures to ensure the reliability of the operations that occur during L + 1 epochs.

Proof:

$$\Pr[\vec{S}(t) = \vec{0}] = \Pr[\text{reset is applied at time } t]$$

$$+ \Pr[\text{reset is not applied at time } t] \times \Pr[\text{successor of } \vec{S}(t-1) \text{ is } \vec{0}]$$

$$\geq \Pr[\text{reset is applied at time } t] = P_1.$$

As $P_S = p^L$ a lower bound for the occurrence of S that is completely independent from t - k for $k \geq 1$ is given by $P \geq P_1 \times p^L$. This is a lower bound and not an exact probability because it ignores the possibility that the correct sequence may be produced even when the re-set operation is not successfully applied.

On the other hand, if we look at the output exclusively, then it may have happened that (i) the reset operation was effective but was followed by any combination of correct states and errors such that the correct output has been preserved, although at least one error in the state of the shift register has occurred during the sequence of length L + 1. (ii) The reset operation was not effective, but the previous state $\vec{S}(t-1)$ was a predecessor of $\vec{0}$ and the same succession of conditions as expressed in (i) has happened. (iii) The reset operation was not effective and the previous state $\vec{S}(t-1)$ was not a predecessor of $\vec{0}$, but the same succession of conditions as expressed in (i) or (ii) has happened.

THEOREM 5: When the measures taken to make reliable the re-set operation that immediately precedes the L sequence tend to have absolute reliability without ever attaining it, then, in the limit, the generation of the sequence S will depend on a Markov process of order L + 1.

Proof: This follows from the fact that the probability of cases (ii) and (iii) will vanish when $P_1 \rightarrow 1$, since then case (i) always occurs.

5. Consequences Concerning Memory Models and Learning Processes

The preceding algorithm and the method of section 3 satisfy many of the requirements that can be specified for a formal model of memory when one takes into account current hypotheses about permanent storage in the N. S.

Namely, the contents of memory is not exclusively represented by the actual state of neural elements. The permanent storage of data will be neither as a state vector that shifts cyclically in a reverberatory manner, nor a stable state. The strategy depends on relations between events embodied in neural nets as patterns of connection and modes of operation.

Another feature of the model consists in an active process that periodically drives the state of all the elements to zero. The actual content of the memory is intended to be the zero state, except during recollection and recognition.

From this point of view, the consequences of our model will be that our algorithm or Massey's algorithm[7] for linear shift-register synthesis can be regarded as a model for the process through which structures of the paleocortex, probably with cooperation of the corpora mammillaria and hippocampus, transform a transient input sequence into a permanent structural organization of neural relations.

For that purpose, a highly reliable set of neurons computing an algorithm of the type proposed here would impose weights in a matrix of neuronal connections capable of all of the possible modes of oscillation in such a way that only the specified mode remains.

The main difference, concerning reliability, is that we have supposed previously that we could specify without an error any desired shift register, and this is not the case with real neurons, for which we have to allow a probability of error in applying the algorithm. Our results suggest the extension of the proposal of shift-register models for the synthesis of macromolecules, which is due to Pattee, in such a way that a scheme similar to the synthesis algorithm may be admitted as an explanation for memory storage by RNA molecules.

No supposition about the kind of plastic changes in the neuronal structure is implied by the present model, although it is clear that it could work either if the relevant changes were in the synapses, or in axonal or dendritic trees, or any combination of these possibilities.

This model suggests, furthermore, a possible mechanism for learning that can occur without needing appreciable repetition of experience as it occurs in certain periods in the lives of young animals. Stated in oversimplified form, in those periods, the occurrence of a certain message in the external input, together with information about the commands that have caused certain actions, leads immediately to a double permanent memory storage, so that when later a certain input is recognized as identical to the past one, it triggers the permanent memory where the sequence of commands for action is stored.

As far as classic and other types of conditioning, as well as trial and error, are concerned, this fixation depends upon the congruence of the results of action over a certain span of time, of of congruent associations of external input sequences occurring also in a given span of time.

At this level of learning, it is reasonable to suppose that, whatever the number of possible actions is, there is a limited number of possible modes of action,[8] and many trials are necessary, not only to secure an adequate level of congruence but also, probably, to compare them under an abductive process to find the preferable mode (preferable from the point of view of some hierarchical decision function).

Probably the use of the same concept – conditioning – for many different processes, even those that are classified as classical, Type I, Type II, or Operant, is inadequate from the point of view of the function of the nervous system because the levels of the structures upon whose function their occurrence depends may be very different from one case to another.

In a previous note we have suggested a feedback shift register, which is adequate for the command of lower order neuronlike structures or of effectors. The pattern of connections and of logical operations can be such that its output can go through a cycle of states, irrespective of the input. Nevertheless, it has also information about the input and can use it, together with the state of its elements, to form sequences that constitute commands for entraining action. If these acts belong to a certain class that, from the point of view of an observer, tends to certain ends or goals, it is then legitimate to speak of the shift register as producing intentional acts that tend toward certain aims and attain them, taking into account its purpose and the state of the external environment.

It is reasonable to postulate that in the brain there is a mapping of the past and the present transactions between the system and the external environment.[9,10]

This model, in a given situation, would generate sequences of foreseen outcomes of possible actions, taking into account past experience. If they were judged to be of utility from the point of view of a given decision function, the actions would ensue and the commands for them would be stored in an enduring form, with the information about the transaction with the external environment.

This kind of performance is usually considered as depending on insight. The same approach could also handle other cognitive processes.

A further remark is that the system is self-organizing, in the sense that an internal structure programs neuronal organization capable of adaptive actions, depending only on experience, the rules of permanent storage, and on a decision about their utility from the point of view of a given criterion of utility.

A model for self-replication can also be immediately derived from our present proposal for permanent memory, and reliable computation in finite lengths of time embedded in the indefinite past.

Acknowledgement: The authors wish to thank James L. Massey for his help in proving one of the theorems, as well as for suggestions and criticism of the manuscript. Thanks are also due to David A. Huffman, James Beever and Roberto Moreno-Diaz for their suggestions and criticism.

J. L. Simões da Fonseca, W. S. McCulloch

References

1. D. A. Huffman, "The Synthesis of Linear Sequential Coding Networks," in Information Theory, edited by C. Cherry (Butterworths, London, 1955), pp. 77-95.

2. K. Fukunaga, "A Theory of Nonlinear Autonomous Sequential Nets Using Z Transforms," IEEE Trans., Vol. EC-13, pp. 310-312, 1964.

3. B. Elspas, "The Theory of Autonomous Linear Sequential Networks," IRE Trans., Vol. CT-6, pp. 45-60, 1959.

4. B. Friedland, "Linear Modular Sequential Circuits," IRE Trans., Vol. CT-6, pp. 61-68, 1959.

5. C. E. Shannon, "Coding Theorems for a Discrete Source with a Fidelity Criterion," 1959 IRE National Convention Record, Vol. 7, Part 4, pp. 142-163.

6. K. B. Magleby, "The Synthesis of Nonlinear Feedback Shift Registers," Technical Report No. 6207-1, Stanford Electronics Laboratory, Stanford, California, 1963.

7. J. L. Massey, "Shift-Register Synthesis and Applications," Quarterly Progress Report No. 85, Research Laboratory of Electronics, M.I.T., Cambridge, Mass., April 15, 1967, pp. 239-240.

8. W. L. Kilmer, W. S. McCulloch, and J. Blum, "Some Mechanisms for a Theory of the Reticular Formation," AF-AFOSR-1023-64, Final Scientific Report, Division of Engineering Research, Michigan State University, Lansing, Michigan, February 28, 1967.

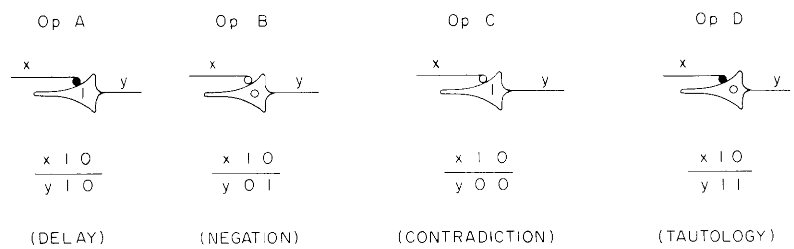
9. W. S. McCulloch, Embodiments of Mind (The M.I.T. Press, Cambridge, Mass., 1965).

10. J. L. Simões da Fonseca, Neuronal Models, Publ. Centro de Estudos Egas Moniz, Lisbon, Portugal, 1966.

B. NEURAL NETS OF CONNECTIVITY ONE THAT FORM CLOSED LOOPS WITHOUT EXTERNAL INPUTS

Def. 1. The connectivity of a neural net is said to be one whenever any neuron of the net has only one input from another neuron of the net and one output that is the input of only one neuron.

Given a set, N, of neurons $\{n_1 n_2 \ldots n_n\}$ with which to form a net that includes all of them and has connectivity one, then (i) any neuron can be connected through its output with any other that does not already have an input; (ii) any neuron can have its input from any other neuron that does not already give its output to another neuron; and (iii) any neuron may perform any one of the four Boolian functions of one variable.



Def. 2. Two nets of neurons, differing only in the number of times that each operation is assigned, and/or, in the order of the assignment, are topologically the same. The pattern of connectivity of any sequence of neurons must have a one-to-one correspondence with the pattern of connectivity of a sequence of neurons in the other. Two nets are independent if they are not connected.

In a former note we reported the establishment of a correspondence between nets of neurons with interaction of afferents, and shift registers. Our present concern is with nets that in some particular cases correspond to linear pure-cycling shift registers, but in others are nonlinear.

Some properties of these nets without inputs and with connectivity one are immediate.

I. The smallest net that can be constructed is formed by two neurons.

II. Any N neurons (for $N \geq 2$) of one output each can be connected to form a net of connectivity one.

III. The smallest number of neurons with which K independent nets can be constructed is $N = 2K$.

IV. Two independent nets of connectivity one, with the same number of neurons can differ only in the number of times each operation is assigned, and/or in the sequence in which these operators are assigned.

V. Given N neurons, the necessary and sufficient condition for obtaining nets with different topology is that one of these nets has at least two independent subnets, and that in any topology having two or more independent subnets, the number of neurons that form each of at least two of these independent subnets differ from the number of neurons in any of the subnets in which the net of the other topology is partitioned.

VI. The number of possible partitions of a set of N neurons such that the topology corresponding to any of the partitions is different from any other is the number of partitions of N that satisfy the following conditions: (i) they differ in accordance with property V; (ii) no independent net has less than two neurons.

VII. In a net of N neurons having at least one neuron that performs functions either C or D, no cycle of states is possible.

VIII. If a net of N neurons is formed exclusively by delay operators, any cycle length $L = \frac{N}{K}$ (K being an integer that divides N without remainder) can be obtained.

Proof:

1. If the first state that we assign to the net is either the 11...1 of the 00...0 state, any shift can only produce the same state again. Therefore, a cycle of length 1 is always possible.

2. Let the first state that we assign be the state 10...0 or any other state with no recurrence sequence of ones and zeros whose length divides N without remainder. In the first case all of the digits except one are zero, therefore the number of shifts of the digits along the net necessary to obtain a one in the same position that it had in the first state is N. Since at that instant all zeros will also occupy the same positions as they did initially, the cycle length L = N is obtained. The same argument applies to any initial state in which there is no recurrence of ones and zeros with length less than N.

3. In the assignment of the initial state we can divide the length N in K recurrent sequences of length $\frac{N}{K}$. Then, if we assign to each of these K partitions the recurrent sequences 10...0 with length $\frac{N}{K}$, we note that in any of those partitions the first neuron receives from the $\frac{N}{K}^{th}$ neuron of the preceding sequence the input that it would receive if it belonged to an $\frac{N}{K}^{th}$-order independent net. From (2) it follows that the same cycle of length $\frac{N}{K}$ occurs for every part of the net of length $\frac{N}{K}$. Therefore, the cycle of the total net would be that of any of its parts, that is $\frac{N}{K}$.

IX. In a net of neurons formed exclusively of negation operators we can obtain any cycle length $L = \frac{N}{K}$ if L is even and $2 \times L$ is odd, except for L = 2, which leads to a cycle of length 1.

Proof:

1. For L = 1, the initial state is either 11...1 or 00...0. Since in the next epoch each neuron will negate its input, the state of the net will be, respectively, 00...0 or 11...1. The initial state will be reproduced in the following epoch. Therefore, for L = 1, the cycle length is $2 \times L$; for L = 2, where the initial state is either 1010...10 or 01...01, each neuron receives as an input the complement of the state in which it is.

Since each neuron performs a negation, it remains in the same state.

2. If L is an even number and $L > 2$, the same argument as in VIII (3) applies because at the end of the L^{th} epoch each of the digits will have shifted an even number of times. At that time the initial state will be reproduced because the division in K different recurrent subsequences is equivalent to dividing the net in K independent nets with the properties that were shown in VIII. Therefore any cycle length L, for L even, $L > 2$, and equal to $\frac{N}{K}$, is possible.

3. If L is an odd number, after L units of time, the state of the parts of the net that correspond to the initial recurrent subsequences will be the complement of the initial subsequence, since each digit will be shifted an odd number of times, L. Only L times later, will the original subsequence repeat. Therefore, for any $L = \frac{N}{K}$ that is an odd number, the cycle length will be $2 \times L$.

COROLLARY: (a) A net formed exclusively by delay operators always has two possible stable states; (b) A net formed exclusively by negation operators has two stable states if N is even and no stable state if N is odd; and (c) A net that contains at least one operator C or D has one stable state, and any state assigned to it can only be either the stable state or belong to a transient that leads to it.

X. If in a net of N neurons there are K negation operators and N-K delay operators, then for $K > 0$: (i) if K is even, at least a cycle of length 1 and a cycle of length N are possible; (ii) if K is odd, then no stable state may occur and at least a cycle of length $2 \times N$ is possible; and (iii) if the distribution of delay and negation operators in the net is such that a given pattern of m neurons with j negations, $j \geq 1$, and $m - j$ delay is recurrently repeated, then for every such disposition a cycle of length m is possible, if j is even, and of length $2 \times m$, if j is odd.

XI. If K independent nets are constructed of N neurons, the length of the longest possible cycle is $2 \times \text{l.c.m.}(l_a, l_b, \dots, l_k)$, where l_j denotes the number of neurons in any independent net j, that is, any of the K partitions of N. In this case, the states of all neurons, N, are used to define a joint state of the K independent oscillators.

XII. The longest cycle that can be obtained out of the elements of a set N of neurons through any arbitrary partition and/or by assignment of one of the four possible operations to each element of the set, is $2 \times \text{l.c.m.}\ (l_a, l_b, \dots, l_u)$, as in property IX, with the following conditions: (i) $l_a, l_b, \dots, l_k$ are prime to one another; (ii) K is the highest number of nets that can be obtained from N such that (i) is satisfied; (iii) if more than one partition gives the highest K, then the partition that gives the longest cycle is that in which the smallest net of that partition is larger than the smallest in any other partition; and (iv) if in more than one partition, conditions (i), (ii) and (iii) are satisfied, then the partition that gives the longest cycle is that in which the difference between the number of neurons in the greatest and smallest nets of the partition is the least.

J. L. Simões da Fonseca, W. S. McCulloch

C. DECIDABILITY IN SHIFT-REGISTER THEORY

Since the theory of shift registers has been constructed upon the identification of the functions performed by these circuits with algebraic operations over Galois fields, it is of interest to know what the solution is of the problem of decidability for these fields.

Given a field F and a set S, of first-order true sentences of F, the problem of decision for that field consists in determining whether S is recursive. If S is recursive, then F is said to be decidable; otherwise, it is said to be undecidable. For a class, C, of fields the problem consists in determining whether the set, R, of true sentences in all fields that belong to C, is recursive.

Our problems concerning shift registers reduce to finding out whether Galois fields, that is, finite fields, are decidable. The answer[1,2] is that any finite field, F, is decidable because any sentence can be reduced to a Boolian combination of sentences of the diagram of F, which is finite.

J. L. Simões da Fonseca, W. S. McCulloch

References

1. J. Robinson, Introduction to Model Theory and to Metamathematics of Algebra (North-Holland Publishing Company, Amsterdam, 1965).

2. J. Robinson, "The Decision Problem for Fields," in The Theory of Models, edited by J. W. Addison, L. Henkin, and A. Tarski (North-Holland Publishing Company, Amsterdam, 1965), pp. 299-311.