The next clock pulse will produce b_1 at the serial output terminal, and so on until the conversion is complete.

We turn our attention now to some details. The indices (of u_i, v_i) adopted in Fig. 6 indicate that the binary point is assumed to be to the right of the LSB. In a base 2 device, one could arbitrarily assign the position of the binary point, since shifting it would only affect the scaling. In the present case, one could only shift the binary point an even number of bits, as an odd shift would introduce a sign change. One therefore has to decide before converting whether $(D - 1)$, the index of the MSB, is even or odd. Fig. 6 is drawn for the even case. The same device, however, could handle the odd case if the reference voltages are interchanged, and if the y flip-flop is initialized by applying the start pulse to the CLEAR rather than the PRESET terminal. The simple switching involved could be implemented either as a manual switch set by the user or through a control signal.

The theory and design of positive radix A/D converters seem to be highly developed. Therefore, our presentation here has concentrated on that which is different in the negative radix case. This means that various details common to both systems have been ignored (Example: synchronization of the start pulse with the clock sequence.) Another point to bear in mind in this context is that Fig. 6 is just one possible implementation of the general conversion strategy indicated in Fig. 1. Reviewing the various established decoder designs will yield different implementations of the negative radix A/D converter.

REFERENCES

[1] S. Zohar, "New hardware realizations of nonrecursive digital filters," *IEEE Trans. Comput.*, vol. C-22, pp. 328–338, Apr. 1973.
[2] ——, "Negative radix conversion," *IEEE Trans. Comput.*, vol. C-19, pp. 222–226, Mar. 1970.
[3] D. F. Hoeschele, Jr., *Analog-to-Digital/Digital-to-Analog Conversion Techniques*. New York: Wiley, 1968.

Shalhav Zohar (A'54–M'60), for a photograph and biography, see page 338 of the April 1973 issue of this TRANSACTIONS.

Correspondence

Representation of Associated Data by Matrix Operators

TEUVO KOHONEN AND MATTI RUOHONEN

Abstract—**It is shown that associated pairs of vectoral items ($Q^{(r)}$, $X^{(r)}$) can be recorded by transforming them into a matrix operator M so that a particular stored vector $X^{(r)}$ can be reproduced by multiplying an associated cue vector $Q^{(r)}$ by M. If the number of pairs does not exceed the dimension of the cue and all cue vectors are linearly independent, then the recollections are perfect replicas of the recorded items and there will be no crosstalk from the other recorded items. If these conditions are not valid, the recollections are still linear least square approximations of the $X^{(r)}$. The relationship of these mappings to linear estimators is discussed. These transforms can be readily implemented by linear analog systems.**

Index Terms—**Associative memory, associative recall, correlation matrix memory, feature filter, least square estimator, linear estimator, regression analysis.**

I. INTRODUCTION

In this correspondence we point out that there are linear analog systems, e.g., electrical networks that can act as selective filters for parallel signal patterns. As such, they perform the same function as the associative memories do in which a set of input signals selectively evokes another set of associated output signals. Selective filters can be used, e.g., for: 1) scale transformation, correction, and separation of multivariate measuring values; and 2) selective detection of patterned information.

In a few recent papers [1]–[5], associative recall from so-called associative nets was reported. In [1] and [3] pairs of vectoral items $Q^{(p)}$, $X^{(p)}$, indexed by the elements p of an index set P, were recorded by networks described by a matrix operator M, and it was shown that when a cue vector $Q^{(r)}, r \in P$ is multiplied by M, an approximate recollection of the associated item $X^{(r)}$ can be obtained. A common feature of such systems is that a recollection is not perfect but a mixture of all recorded items in which the searched item dominates. In so-called correlation matrix memories [3], a measure of crosstalk is the inner product between cue vectors, and only for mutually orthogonal cues is the recollection a perfect replica of the stored item. On the other hand, approximate recall using fragmentary cues is possible.

In this correspondence we introduce an associative matrix transformation *in which orthogonality of cue vectors is not required*. In fact, the only restriction imposed on the cue vectors for perfect recall is that they be linearly independent; this is a rather mild condition if the number of pairs does not exceed the dimension of the cue vectors, although the cues were selected without prior checking. We can then show that there exists a record matrix M, equivalent to an input–output transfer relation of a signal transforming system, which represents all pairs of items and by which any of the recorded data items $X^{(r)}$ can be reproduced by multiplying an associated cue vector $Q^{(r)}$ by M. There will be no crosstalk from other items.

II. PERFECT RECALL

We are looking for a matrix operator M that shall represent pairs of recorded vectors ($Q^{(p)}$, $X^{(p)}$, $p \in P = \{1, 2, \cdots, s\}$ and by which any of the recorded items $X^{(r)}$, $r \in P$ will be reproduced by a linear operation

$$X^{(r)} = MQ^{(r)}, \quad r \in P; \quad X^{(r)} \in \mathcal{R}^n, \; Q^{(r)} \in \mathcal{R}^m. \tag{1}$$

Equation (1) for all $r \in P$ can be rewritten in a compact form by introducing the matrices

Manuscript received October 5, 1972; revised January 19, 1973.
The authors are with the Department of Technical Physics, Helsinki University of Technology, Otaniemi, Finland.

$$X = [X^{(1)}, \cdots, X^{(s)}]$$
$$Q = [Q^{(1)}, \cdots, Q^{(s)}]$$

whereby

$$X = MQ. \tag{2}$$

We shall now use a result derived by Penrose [6].

Lemma: A necessary and sufficient condition for the equation $AMB = C$ to have a solution for M, where A, B, and C are arbitrary matrices, is that

$$AA^+CB^+B = C$$

in which case the general solution is

$$M = A^+CB^+ + Y - A^+AYBB^+$$

where Y is an arbitrary matrix of the same dimensions as M, and A^+ and B^+ are the *pseudoinverses* of A and B, respectively.

Therefore, the general solution to (2) is

$$M = XQ^+ + Y(I_m - QQ^+) \tag{3}$$

where Q^+ is the pseudoinverse of Q, I_m is the $m \times m$ identity matrix, and Y is an arbitrary matrix of the same dimension as M, provided that the following condition is satisfied:

$$XQ^+Q = X. \tag{4}$$

If this condition is to be valid for arbitrary X, then it is reduced to

$$Q^+Q = I_s. \tag{5}$$

Equation (5) is equivalent to the fact that the rank of the matrix Q is s, i.e., the columns of Q are linearly independent.

III. Approximate Recall

If the rank of the matrix Q introduced before is smaller than s, which is the case if the vectors $Q^{(p)}$, $p \in P$, are linearly dependent, e.g., if there are more than m cue vectors, no exact solution of (2) for M generally exists. The best approximate solution in the sense of least squares is now obtained according to Penrose [7].

Let us minimize the norm of the matrix $X - MQ$, i.e., the square root of the squares of its elements. This is equivalent to minimizing the trace of the matrix

$$E = (X - MQ)(X - MQ)^T. \tag{6}$$

This leads to the so-called *normal equation*

$$MQQ^T = XQ^T. \tag{7}$$

We shall now show that the least square solution of the normal equation also minimizes the trace of E in (6). According to [7], the least square solution of (7) is

$$M = XQ^T(QQ^T)^+ = XQ^+. \tag{8}$$

By a substitution of this value for M in (6) we obtain

$$E_1 = X(I_s - Q^+Q)X^T. \tag{9}$$

Now, after a few steps of algebraic manipulation, $E - E_1$ can be put in the form

$$E - E_1 = (M - XQ^+)QQ^T(M - XQ^+)^T. \tag{10}$$

A matrix of the form BB^T is always positive semidefinite, and so $E - E_1$ is also positive semidefinite. The value for M obtained from (8) has thereby been shown to minimize the trace of E.

Now putting $Y = 0$ in (3) we may collect the results obtained in Sections II and III in the following theorem.

Theorem: When a linear signal-transforming system with the matrix transfer operator $M = XQ^+$ is excited by an input signal vector $Q^{(r)}$, $r \in P$, then the output signal vector $\hat{X}^{(r)}$ is the best approximation of the recorded data vector $X^{(r)}$ in the sense of least squares. Moreover, if the columns of Q are linearly independent, the recollections are perfect replicas of the recorded data vectors $X^{(r)}$.

IV. Relationship to Linear Estimator

Finally we derive an interesting dualism between matrix operators discussed before, and regression analysis. If we regard the vectors $X^{(p)}$, $Q^{(p)}$, $p = 1, \cdots s$ as observations of stochastic variables x and q, respectively, then the linear least square regression of x on q is given by the matrix M which minimizes the expression

$$\sum_{p=1}^{s} [X^{(p)} - MQ^{(p)}]^T [X^{(p)} - MQ^{(p)}].$$

This is equivalent to minimizing the diagonal elements of

$$\sum_{p=1}^{s} [X^{(p)} - MQ^{(p)}] [X^{(p)} - MQ^{(p)}]^T = (X - MQ)(X - MQ)^T.$$

The value of M obtained from these expressions is obviously the same as the solution given in (8), so that

$$x = XQ^T(QQ^T)^+q = XQ^+q \tag{11}$$

gives the least square estimate for x when q is known.

References

[1] D. J. Willshaw, O. P. Buneman, and H. C. Longuet-Higgins, "Non-holographic associative memory," *Nature*, vol. 222, pp. 960–962, 1969.

[2] T. Kohonen, "A class of randomly organized associative memories," *Acta Polytech. Scandinavica*, El 25, 1971.

[3] —, "Correlation matrix memories," *IEEE Trans. Comput.*, vol. C-21, pp. 353–359, Apr. 1972.

[4] K. Nakano and J. Nagumo, in *Advanced Papers Conf., 2nd Int. Joint Conf. Artificial Intelligence*, British Comput. Soc. (London), Sept. 1971, pp. 101–110.

[5] K. Nakano, "Associatron—A model of associative memory," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-2, pp. 380–388, July 1972.

[6] R. Penrose, "A generalized inverse for matrices," *Proc. Cambridge Phil. Soc.*, vol. 52, pp. 406–413, 1955.

[7] —, "On best approximate solutions of linear matrix equations," *Proc. Cambridge Phil. Soc.*, vol. 52, pp. 17–19, 1956.

Comments on "On the Definition and Generation of Walsh Functions"

J. W. J. VAN TILL

***Abstract*—A correction to the Gray code-to-decimal conversion of Davies is given, which follows from Davies' own proof if performed on the level of operations on wave number indices. An alternative proof for Walsh functions as combinations of Rademacher functions is also given, starting from the definition of Walsh–Kaczmarz functions.**

In his short note,[1] Davies states an equation, taken from Sobel, to obtain the decimal number of a Gray code number $g_m g_{m-1} \cdots g_1 g_0$:

Manuscript received September 26, 1972; revised January 10, 1973.

The author is with the Laboratory for Information Theory, Department of Electrical Engineering, Delft University of Technology, Delft, Holland.

[1] A. C. Davies, *IEEE Trans. Comput.* (Short Notes), vol. C-21, pp. 187–189, Feb. 1972.