XX. PROCESSING AND TRANSMISSION OF INFORMATION*

Prof. D. N. Arden
Prof. E. Arthurs
Prof. W. B. Davenport, Jr.
Prof. P. Elias
Prof. R. M. Fano
Prof. R. G. Gallager
Prof. F. C. Hennie III
Prof. E. M. Hofstetter
Prof. R. A. Howard
Prof. D. A. Huffman
Prof. E. F. Moore
Prof. W. F. Schreiber
Prof. C. E. Shannon
Prof. J. M. Wozencraft
J. H. Babcock
G. Cheadle
J. E. Cunningham
P. M. Ebert
G. Einarsson
J. Epstein
H. A. Ernst
E. F. Ferretti
D. H. Friedman
T. J. Goblick, Jr.
J. P. Green, Jr.
U. F. Gronemann
T. S. Huang
F. Jelinek
R. S. Kennedy
L. Kleinrock
C. L. Liu
H. H. Loomis, Jr.
J. L. Massey
C. Moraga
R. C. Norris
E. L. O'Neill
J. W. Pan
R. R. Pfeiffer
L. G. Roberts
J. R. Sklar
I. G. Stiglitz
I. E. Sutherland
O. J. Tretiak
J. Ziv

RESEARCH OBJECTIVES

This group continues its investigations of sources that generate information, channels that transmit it, and machines that process it.

In the field of image transmission systems, work is proceeding along several lines of investigation. A flexible high-quality system has been built that prepares computer tapes from original pictures and produces pictures from computer-generated tapes. With this apparatus, studies are being made by computer simulation of bandwidth reduction systems and color transmission methods. Test images are being prepared to study optimum filtering in image systems and properties of the observer. In general, our aim is to elucidate the fundamental properties of vision as they apply to the image-transmission process. This knowledge will then be exploited to design efficient systems and to build systems capable of performing some "human" operations such as noise reduction, image detection, and quality improvement.

During the past year, significant results have been reported[1] on the fundamental limitations of communication over discrete memoryless channels. This research has also been extended to include Gaussian memoryless channels, and an upper bound has been established on the performance attainable with realizable coders and decoders.[2] A unified characterization of optimum receivers for time-variant dispersive Gaussian channels[3] has provided additional insight into the problem of communicating through actual propagation media.

The effort to relate theory to practical communication problems continues. The effect on decoders of discrete noise with Markov memory is under investigation. In addition, research is directed toward determining maximally efficient combined error-correction and error-detection feedback strategies for noisy two-way channels, subject to constraints on equipment complexity. Also, telephone lines and deep-space data links are being investigated to determine the advantages of coding in specific communication environments.

Similarly motivated work has been done in the area of data processing. A number of basic properties of logical cells have been established,[4] including the recursive undecidability of many steady-state and transient problems. Work is continuing on the development of useful synthesis techniques and canonical forms for stable networks.

*This research was supported in part by Purchase Order DDL B-00306 with Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the joint support of the U.S. Army, Navy, and Air Force under Air Force Contract AF19(604)-7400; in part by the National Science Foundation (Grant B-16526); and in part by the National Institutes of Health (Grant MP-4737).

Also, an investigation of information flow in communication networks, and the queues and delays therein, is nearing completion.

Various models that are useful in establishing the theoretical limitations of digital computing circuits are being studied. Studies indicate that there are invariants that interrelate the complexity of logic, the number of feedback loops, and the number of storage elements to the reaction time of the resulting system and the complexity of the data-processing job.

Current digital systems are designed to carry out a sequence of operations serially, or one at a time. Research continues on how to exploit "parallel" logical organization, so that computation rate need not be limited by propagation or component-response times.

Two simple digital circuits that generate finite sequences of pulses which are ideal for radar-ranging have been discovered; the autocorrelation functions achieve the theoretical optimum for time-limited waveforms. It is conjectured that there exists a sequence of such circuits, with waveforms that are not only ideal (in the sense mentioned above) but also achieve arbitrarily large ratios of total energy to peak power.

E. Arthurs, D. A. Huffman, W. F. Schreiber, J. M. Wozencraft

References

1. R. M. Fano, Transmission of Information (The M. I. T. Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1961).

2. J. Ziv, Coding and Decoding for Time-Discrete, Amplitude-Continuous Memoryless Channels, Sc.D. Thesis, to be submitted to the Department of Electrical Engineering, M. I. T., January 1962.

3. T. Kailath, Communication via Randomly Varying Channels, Sc.D. Thesis, Department of Electrical Engineering, M. I. T., June 1961.

4. F. C. Hennie, Iterative Arrays of Logical Circuits (The M. I. T. Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1961).

A. PICTURE PROCESSING

1. LABORATORY EQUIPMENT

Further improvements and modifications in the digital picture tape recording and playback equipment have been undertaken. One was the redesign of the digital-to-analog converter circuits in the scanner so as to obtain a better long-term stability of the position of the spots and lines. Another modification was in the circuitry for monitoring and photographing the digitized video signal during recording.

The compatibility of tapes written by the digital video tape recorder and the tapes used by the IBM 709 computer was tested by having the computer read and rewrite tapes containing known video signals. The IBM 709 computer was able to read tapes written by the digital tape recorder, but there were some problems concerning signal strength, record lengths, and record gaps. During the next quarter attempts will be made to process the video information by using the IBM 7090 computer that is being installed.

During the next quarter the system will undergo further modification to accommodate color pictures. If the position stability of the scanner will permit it, the total picture will be scanned three times successively – once for each color – through an appropriate filter. In this manner, the present optical system with the single phototube will suffice.

J. E. Cunningham, U. F. Gronemann, T. S. Huang,
J. W. Pan, O. J. Tretiak, W. F. Schreiber

2. COLOR-SYSTEM INVESTIGATIONS

The following projects will be undertaken in connection with color pictures:

1) An estimate of the entropies of typical color pictures will be computed and compared with those of the corresponding monochromatic pictures.

2) The psychovisual effects of quantization of color information will be investigated. Valensi's division of the chromaticity diagram into constant-color zones will serve as the starting point.

3) Combined quantization of color and brightness information will be investigated.

4) The visual response to various color components as a function of spatial frequency will be evaluated in terms of further reduction in information requirements.

U. F. Gronemann

3. THRESHOLD MEASUREMENTS OF GRAIN VISIBILITY

The brightness distribution in a piece of grainy photographic material may be regarded as a sample function of a two-dimensional random process. Thus far, there is no model that uses the parameters of this random function to predict whether the grain is above or below the threshold of visibility.

We propose to use a digital computer to generate sample functions of a two-dimensional random process of controlled probability distribution. The digital television equipment will then be used to obtain pictures whose brightness distribution is a sample function of a two-dimensional random process of known statistics. Grain visibility thresholds will be found for pictures of this type.

The primary question of this research is whether the threshold is related in some way to the statistics of the random brightness function.

O. J. Tretiak

B. THE R(d) FUNCTION FOR A DISCRETE SOURCE WITH A DISTORTION MEASURE

This report is concerned with work done on the encoding of a discrete information source with a distortion measure.[1] The information source selects letters independently from an alphabet X according to the probability distribution $P(x)$, $x \in X$. There is another alphabet Y, called the output alphabet, which is used to encode or represent the source output as follows. The encoder maps sequences of source letters of length n into output sequences $v_1, v_2, \ldots, v_M$, which consist of n letters of the output alphabet Y. This mapping may be regarded as a partitioning of the input or source space U, of all possible sequences u of n letters of the X alphabet, into M disjoint subsets $w_1, w_2, \ldots, w_M$. Each w_i consists of all of the u sequences that are mapped into the sequence v_i.

The distortion measure $d(xy) \geq 0$ gives the amount of distortion when source letter x is mapped by the encoder into output letter y. When the sequence

$$u = \xi_1\xi_2 \ldots \xi_n, \qquad \xi_i \in X$$

is mapped by the encoder into the sequence

$$v = \eta_1\eta_2 \cdots \eta_n, \qquad \eta_i \in Y$$

the distortion is defined by

$$d(uv) = \frac{1}{n} \sum_{i=1}^{n} d(\xi_i\eta_i).$$

For any block-encoding scheme defined by a set of M output sequences v_i of length n and a partitioning of the source space U into encoding subsets w_i, the average distortion is given by

$$\overline{d} = \sum_{i=1}^{M} \sum_{w_i} d(uv_i)\, P(u).$$

Define the distance function

$$D(xy) = d(xy) - \ln f(x)$$

where f(x) is an arbitrary probability distribution on the X alphabet. The distance between two sequences is

$$D(uv) = \frac{1}{n} \sum_{i=1}^{n} D(\xi_i\eta_i)$$

$$= \frac{1}{n} \sum_{i=1}^{n} (d(\xi_i\eta_i) - \ln f(\xi_i)) = d(uv) - \frac{1}{n} \sum_{i=1}^{n} \ln f(\xi_i).$$

THEOREM 1. For a given $f(x)$, consider a particular set of output sequences v_i, and a partitioning of the input or source space U into disjoint subsets w_i for which the following conditions are satisfied for some constant D_o: Each subset w_i contains all sequences u for which $D(uv_i) < D_o$ and no sequence u for which $D(uv_i) > D_o$. For the given set of output sequences and for all partitionings of the input space U, the average distance satisfies the following inequality.

$$\bar{D} \geq \sum_{i=1}^{M} \sum_{w_i} D(uv_i)\, P(uv_i) = \sum_{i=1}^{M} \sum_{w_i} D(uv_i)\, P(u)$$

Note that the average distortion for such partitionings then satisfies

$$\bar{d}(f) = \sum_{i=1}^{M} \sum_{w_i} d(uv_i)\, P(u)$$

$$\geq \sum_{i=1}^{M} \sum_{w_i} D(uv_i)\, P(u) + \sum_{X} P(x) \ln f(x)$$

for a particular $f(x)$ because $\frac{1}{n} \sum_{i=1}^{M} \sum_{w_i} P(u) \ln f(u)$ is independent of the output space and the encoding scheme.

THEOREM 2. Suppose that the set of output sequences in Theorem 1 was composed of sequences all of which have a prescribed composition $n(y)$; that is, the number of times each letter y of the Y alphabet appears in a sequence is $n(y)$ and is the same for any output sequence. Let U_o be the set of source sequences y for which $D(uv_o) < D_o$ for any particular sequence v_o having the prescribed composition $n(y)$. Then, if

$$M \leq \frac{1}{\sum_{U_o} P(u)}$$

all encoders that use output words only of the prescribed composition yield

$$\bar{D} \geq M \sum_{U_o} D(uv_o)\, P(u).$$

The information rate for the encoder described above is given by

$$R = \frac{1}{n} \log M.$$

In the limiting case of $n \to \infty$, it can be shown that for a fixed D_o the limiting value of R is given by

$$R_L = sD_o - \gamma(s), \qquad s \leq 0$$

where

$$\gamma(s) = \sum_Y \frac{n(y)}{n} \gamma_y(s)$$

and

$$\gamma_y(s) = \ln \sum_X e^{s(d(xy) - \ln f(x))} P(x).$$

The variable s is chosen so that

$$\frac{\partial}{\partial s} \gamma(s) = D_o.$$

The limiting value of $\bar{\bar{D}}$ as $n \to \infty$ for fixed D_o is simply D_o.

Now suppose that R_L is optimized for constant D_o by use of the functions $f_o(x)$ and $n_o(y)$. Then we have a block code for which

$$\bar{d} \geq D_o + \sum_X P(x) \ln f_o(x) = d_L(D_o)$$

in the limit of $n \to \infty$. We can then make the statement that for any block code for which

$$R \leq R_L(D_o)$$

we must have

$$\bar{d} \geq d_L(D_o).$$

We have then defined a rate-distortion function $R_L(d)$ which has the following significance. All encoding schemes that have rates equal to or less than $R_L(d)$ give average distortion equal to or greater than d.

By means of a random coding argument, a corresponding limit function $R_U(d)$ can be found which implies that there exist encoding schemes with rates equal to or greater than $R_U(d)$, for which the average distortion is less than or equal to d. By proper choice of functions analogous to f(x) and n(y), the $R_U(d)$ function can be made equal to the $R_L(d)$

function. This defines uniquely a limit function R(d) that gives the equivalent information rate of a source with a distortion measure.

Further investigations along these lines will be concerned with the study of systematic encoding schemes and their relative complexities.

T. J. Goblick, Jr.

References

1. This is an elaboration of work presented by C. E. Shannon, Coding Theorems for a Discrete Source with a Fidelity Criterion, IRE National Convention Record, Part 4, 1959, pp. 142-163.

2. R. M. Fano, Transmission of Information (The M. I. T. Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1961), see pp. 285, 286.

C. INFINITE-MEMORY BINARY SYMMETRIC CHANNELS

This investigation is concerned with the use of coding to achieve highly reliable communication by means of time discrete binary symmetric channels that possess a particular form of infinite memory. The specific objectives are to determine: (a) upper and lower bounds on the minimum attainable probability of a decoding error Pe as a function of the information transmission rate R; and (b) relatively simple encoding and decoding procedures that achieve, or nearly achieve, this minimum. In a broader sense, the objectives are to determine: (a) the degree to which channel memory can be exploited to improve the reliability of communication; and (b) the degree to which the reliability deteriorates when this memory is not fully exploited – either because of incomplete knowledge of the channel, or because of a desire to reduce equipment requirements.

1. Channel Model

The class of channels chosen as a basis for this study are mathematically tractable and still possess the basic characteristics of some physical communication channels. Specifically, binary input-output channels with additive (modulo 2) binary noise are considered. That is, the channel input consists of sequences, u, of binary symbols 0 and 1, as does the channel output, v, and

$$v = u \oplus \eta,$$

where η is the channel noise. We further restrict ourselves to channels for which η is a projection of a finite-state irreducible Markov process.

A channel of this class is specified by the properties of the noise sequence η. These properties are, in turn, determined by the initial state probabilities and transition

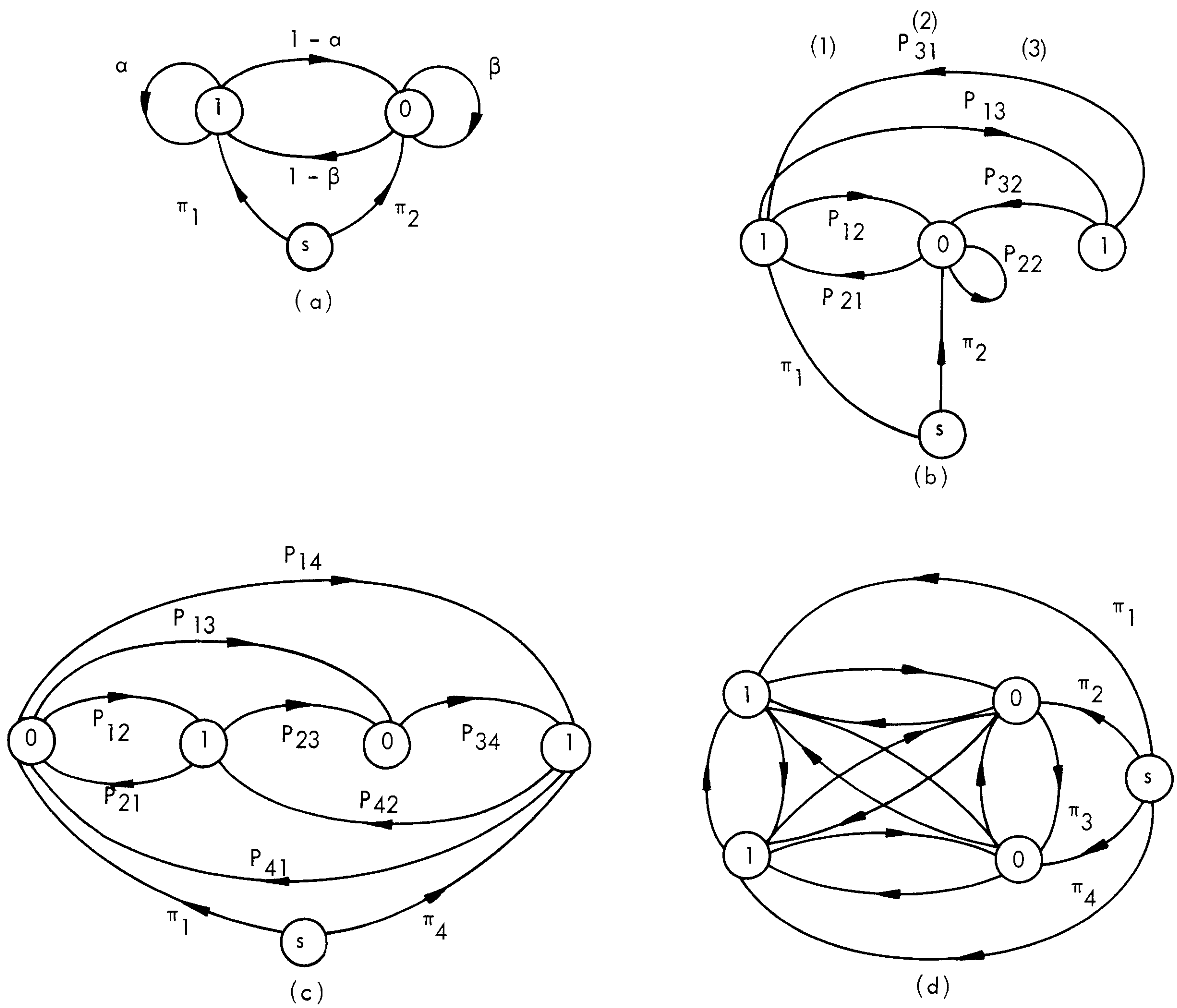

Fig. XX-1. Channel state diagrams.

matrix of the underlying Markov process, and the Markov states that are projected into noise symbols 1 and 0. This information can be presented in a state diagram of the form illustrated in Fig. XX-1.

In these diagrams s is a starting state included to indicate the initial state probabilities π_i, while the other branch labels indicate the state transition probabilities Pij. Finally, the encircled state symbols 1 and 0 specify the projection that generates the noise sequence η. For example, in Fig. XX-1b a noise symbol 1 is generated whenever the Markov process is in states 1 and 3, whereas a noise symbol 0 is generated in state 2.

2. Summary of Results

Figure XX-1 illustrates three types of behavior exhibited by channels of the class investigated.

In Fig. XX-1a the channel noise is itself a Markov process. For these channels, upper and lower bounds to the minimum attainable probability of a decoding error, Pe, for block codes of length N have been found. These bounds were obtained by the use of random coding and sphere-packing arguments, in the former a uniform distribution over the ensemble of code words and maximum likelihood decoding are assumed. These bounds are

$$\frac{K_2}{N} \epsilon^{-(N-1)\alpha_L} \leq \mathrm{Pe} \leq K_1 \epsilon^{-(N-1)\alpha_U} \tag{1}$$

where K_1 and K_2 are constants, and α_L and α_U are functions of the state transition probabilities and the information rate R. The dependence of α_L and α_U upon R is of the form illustrated in Fig. XX-2. When R equals channel capacity C, α_L and α_U are both equal to zero, as are their derivatives with respect to R. (We assume equiprobable messages so that $R = (\log_2 M)/N$ bits, where M is the number of messages.) As R is reduced α_L and α_U increase, but remain equal, and their derivatives with respect to R decrease monotonically and attain a value of −1 at a value of R which is defined as R_c. For R less than R_c the α_U versus R curve has slope −1, while the slope of the α_L curve continues to decrease, approaching minus infinity as R approaches zero.

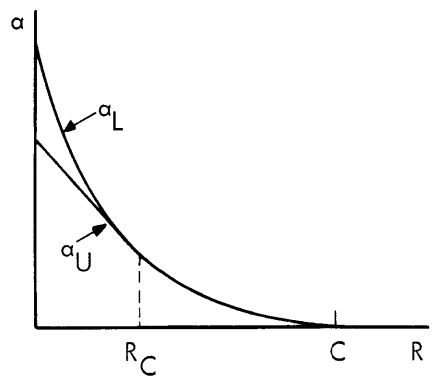

Fig. XX-2. Error probability exponent — upper and lower bounds.

In Fig. XX-1b and 1c the noise is not Markovian, since the conditional probability of a noise symbol 0 or 1 at time t + 1 depends upon the entire past sequence of noise symbols, not just that at time t. However, these state diagrams do possess a useful property: To every possible noise sequence there corresponds only one sequence of Markov states. It follows that the probability of a particular channel-noise sequence can be expressed as a single product of the Markov state transition probabilities. A state diagram that does not possess this property is shown in Fig. XX-1d.

Bounds on the probability of a decoding error for block codes of length N have been found for channels that possess this property. The results are identical to those just discussed (Eq. 1 and Fig. XX-2) except that

1) In Eq. 1, N is replaced by $N^{\frac{m}{2}(m-1)}$, where m is the number of Markov states, and

2) α_L may be infinite for R less than some R.

It should be noted that the bounds of Eq. 1 and Fig. XX-2 are identical in form to those obtained by Fano[1] for general discrete memoryless channels. However, this does not imply that the probability of error given by Eq. 1 and Fig. XX-2 can be achieved without exploiting the channel memory – in general, it cannot.

The channel described by Fig. XX-1d typifies the most general channel that can occur in the class investigated, and the one that is most difficult to analyze. The essential difficulty is that a simple expression for the probability of a noise sequence cannot be obtained. Thus although the sphere-packing bound for the channel can be formulated, the solution cannot be completed without further approximation and a consequent deterioration of the lower bound. The bound thus obtained is, however, of the form of Eq. 1 and Fig. XX-2.

A further complication is that the random-coding argument for the channel cannot be completed for a maximum likelihood decoding procedure. This, again, is due to the lack of a simple expression for the probabilities of the noise sequence. However, this difficulty can be circumvented by considering other than maximum likelihood decoding procedures. In particular, we consider decoding procedures that would be maximum likelihood for channels typified by Fig. XX-1a, 1b, or 1c. A random-coding argument is then used to upper-bound the probability of error when such decoders are used in conjunction with channels of the form illustrated by Fig. XX-1d. Since the probability of error could only be reduced by using maximum likelihood decoding, it follows that we obtain an upper bound to the attainable probability of error.

This problem has been solved, but is still not in a form that permits comparison with the (approximated) sphere-packing lower bound.

3. Non-Maximum Likelihood Decoders

The preceding discussion has dealt primarily with maximum likelihood channel decoders. However, the determination of bounds on Pe for non-maximum likelihood decoding is also of interest, as will be seen from the following considerations.

First, for the class of channels considered here, the implementation of maximum likelihood decoding requires a knowledge of the state transition probabilities. On the other hand, in many physical communication systems one can hardly hope to know these probabilities exactly – or perhaps even the state connections or number of states. Hence, the channel decoder used will not, in general, perform a maximum likelihood decoding procedure, and it is natural to ask: For a given channel how sensitive is the Pe to design errors in the assumed channel parameters? The question is also applicable to the general discrete memoryless channel. Both are in marked contrast to the memoryless binary symmetric channel for which the decoder design is independent of the cross over probability, p, for all $p \leq \frac{1}{2}$.

Furthermore, even if the channel were known exactly, it might be advantageous to use other than maximum likelihood decoding in order to reduce equipment requirements. Again, the question arises concerning the deterioration in system performance as a function of decoder complexity.

During the next quarter these questions will be examined in detail. In particular, convolutional coding and sequential decoding techniques will be investigated.

R. S. Kennedy

References

1. R. M. Fano, Transmission of Information (The M.I.T. Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1961).

D. MAJORITY DECODING OF CONVOLUTIONAL CODES

We report here the development of a simple algebraic algorithm for decoding certain binary convolutional codes.

1. Convolutional Encoding

The concept of convolutional encoding can best be shown by a specific example. (This type of group code was first discovered and named by Elias.[1]) Suppose that it is desired to encode a sequence of information bits that are spaced two units of time apart, which we designate by the symbols $i_o, i_2, i_4, i_6, i_8, \ldots$. The encoding rule is to begin transmitting at time t = 0 the generator sequence

$$\vec{g} = 1, g_1, 0, g_3, 0, g_5, \tag{1}$$

if $i_o = 1$, and the all-zero sequence otherwise. The same rule is used at time t = 2 for bit i_2, and so on for the other information symbols. The actual transmitted signal is the sum (modulo 2) of these signals. It is clear that the first six-bit output from the encoder can be found from the matrix product

$$t = [t_o, t_1, t_2, t_3, t_4, t_5] = [i_o, i_2, i_4] \begin{bmatrix} 1 & g_1 & 0 & g_3 & 0 & g_5 \\ 0 & 0 & 1 & g_1 & 0 & g_3 \\ 0 & 0 & 0 & 0 & 1 & g_1 \end{bmatrix} \tag{2}$$

where t_j is the transmitted digit at time j, and that the information symbol i_o affects only the first six transmitted digits; moreover, the information bits appear unaltered at the even instants of time. The symbols appearing at odd instants of time are the check bits or parity bits. This code has rate $\frac{1}{2}$, since half of the bits output is information bits. It may also be said to have constraint length 6. We shall designate such a code as a $\left[6, \frac{1}{2}\right]$ code.

2. Checking for Parity Failures

We shall use an algebraic approach to introduce the concept of parity sets for a convolutional code. To correspond to the generator sequence $\vec{g}$ of Eq. 1, we define the polynomial

$$g(x) = 1 + g_1x + g_3x^3 + g_5x^5. \tag{3}$$

Here and hereafter all arithmetic is performed in the field of binary numbers.

We now define

$$x^6 \equiv 0. \tag{4}$$

It can be readily checked that the transmitted sequence of Eq. 2 corresponds to the polynomial t(x), given by

$$\begin{aligned} t(x) &= t_o + t_1x + t_2x^2 + t_3x^3 + t_4x^4 + t_5x^5 \\ &= \left(i_o+i_2x^2+i_4x^4\right) g(x) \end{aligned} \tag{5}$$

Let the first six bits at the receiver be represented by the sequence

$$\vec{r} = r_o, r_1, r_2, r_3, r_4, r_5, \tag{6}$$

where r_j is the received bit at time j. Let $\bar{r}$ be represented by the sum

$$\vec{r} = \vec{t} + \vec{e}, \tag{7}$$

where $\vec{e}$ is the error pattern that occurs, given by

$$\vec{e} = e_o, e_1, e_2, e_3, e_4, e_5, \tag{8}$$

in which e_j is a one if the received bit at time j has been changed during transmission, and is zero otherwise.

We now form the polynomial S(x), given by

$$S(x) = \left(r_o+r_2x^2+r_4x^4\right) g(x) + r(x). \tag{9}$$

Substituting Eqs. 8, 7, 5, and 3 in Eq. 9, we obtain

$$S(x) = \left(e_o+e_2x^2+e_4x^4\right) g(x) + e(x). \tag{10}$$

This equation illustrates that only the error pattern $\vec{e}$ affects S(x). Carrying out the multiplication in Eq. 10, we obtain

$$\begin{aligned} S(x) &= (e_o g_1 + e_1)\, x + (e_o g_3 + e_2 g_1 + e_3)\, x^3 + (e_o g_5 + e_2 g_3 + e_4 g_1 + e_5)\, x^5 \\ &= s_1 x + s_3 x^3 + s_5 x^5. \end{aligned} \tag{11}$$

We call the components s_1, s_3, and s_5 the parity sets. We have

$$\begin{aligned} s_1 &= e_o g_1 + e_1 \\ s_3 &= e_o g_3 + e_2 g_1 + e_3 \\ s_5 &= e_o g_5 + e_2 g_3 + e_4 g_1 + e_5. \end{aligned} \tag{12}$$

From these formulas, the general form of the parity sets for any constraint length is evident. It is clear also that all of the s_j are zero when no errors occur, and the subscript j indicates which parity bit enters into that parity set.

3. Majority Decoding

We now introduce the concept of majority decoding. The decoding problem for convolutional codes reduces simply to determining whether e_o is a zero or a one on the basis of the first set of received digits equal in number to the constraint length. For majority decoding we select a generator g in such a manner that a set of 2T linear combinations of the parity sets of Eq. 12 can be formed with the following property: e_o appears in each and every one of these composite parity checks, but no other e_j appears in more than one composite check. It is then possible to make a correct decision on e_o, provided that T, or fewer, errors occur within the constraint span, by using the simple algorithm: Decode $e_o = 1$ if a majority of the composite checks are "ones," otherwise decode $e_o = 0$.

Let us now consider a specific example of a $\left[12, \frac{1}{2}\right]$ code that is capable of correcting 2, or fewer, errors in this manner. For our generator, we take $g_1 = 1$, $g_3 = 0$, $g_5 = 0$, $g_7 = g_9 = g_{11} = 1$. From Eq. 12 we find the parity sets to be

$$\begin{aligned} s_1 &= e_o + e_1 \\ s_3 &= e_2 + e_3 \\ s_5 &= e_4 + e_5 \\ s_7 &= e_o + e_6 + e_n \\ s_9 &= e_o + e_2 + e_8 + e_9 \\ s_{11} &= e_o + e_2 + e_4 + e_{10} + e_{11}. \end{aligned} \tag{13}$$

We now form the four composite checks given by

$$
\begin{aligned}
s_1 &= e_o + e_1 \\
s_3 + s_9 &= e_o + e_3 + e_8 + e_9 \\
s_7 &= e_o + e_6 + e_7 \\
s_{11} &= e_o + e_2 + e_4 + e_{10} + e_{11}.
\end{aligned}
\tag{14}
$$

These four equations have the property that e_o enters into each one and no other e_j enters into more than one; hence the code corrects 2, or fewer, errors by the majority algorithm.

Table XX-1. Convolutional codes for majority decoding.

Code	Errors	$\vec{g}$	Composite Parity Checks
$\left[4, \frac{1}{2}\right]$	≤ 1	1101	s_1, s_3
$\left[3, \frac{1}{3}\right]$	≤ 1	111	s_1, s_2
$\left[12, \frac{1}{2}\right]$	≤ 2	110000010101	$s_1, s_3 + s_9, s_7, s_{11}$
$\left[9, \frac{1}{3}\right]$	≤ 2	111011000	$s_1, s_2, s_4, s_5 + s_7$
$\left[24, \frac{1}{2}\right]$	≤ 3	110000000000010100010101	$s_1, s_3 + s_{15}, s_5 + s_{19}, s_7 + s_{21}, s_9 + s_{17} + s_{23}, s_{17}$
$\left[15, \frac{1}{3}\right]$	≤ 3	111001001011010	$s_1, s_2, s_4, s_7 + s_8, s_{10}, s_{11} + s_{14}$

In Table XX-1 we give the generator sequence and the rules for forming the composite parity checks for the best known single, double, and triple error-correcting codes for rates $\frac{1}{2}$ and $\frac{1}{3}$.

4. Instrumentation

The convolutional codes can easily be encoded by using the canonic-form circuit described by Wozencraft and Reiffen.[2] The circuit for the double-error-correcting

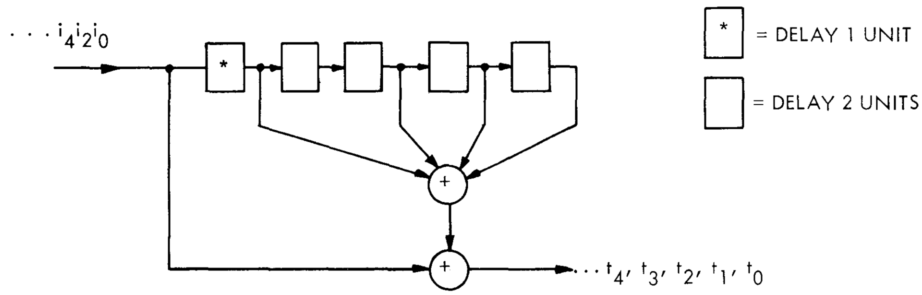

Fig. XX-3. Encoder for $\left[12, \frac{1}{2}\right]$ code.

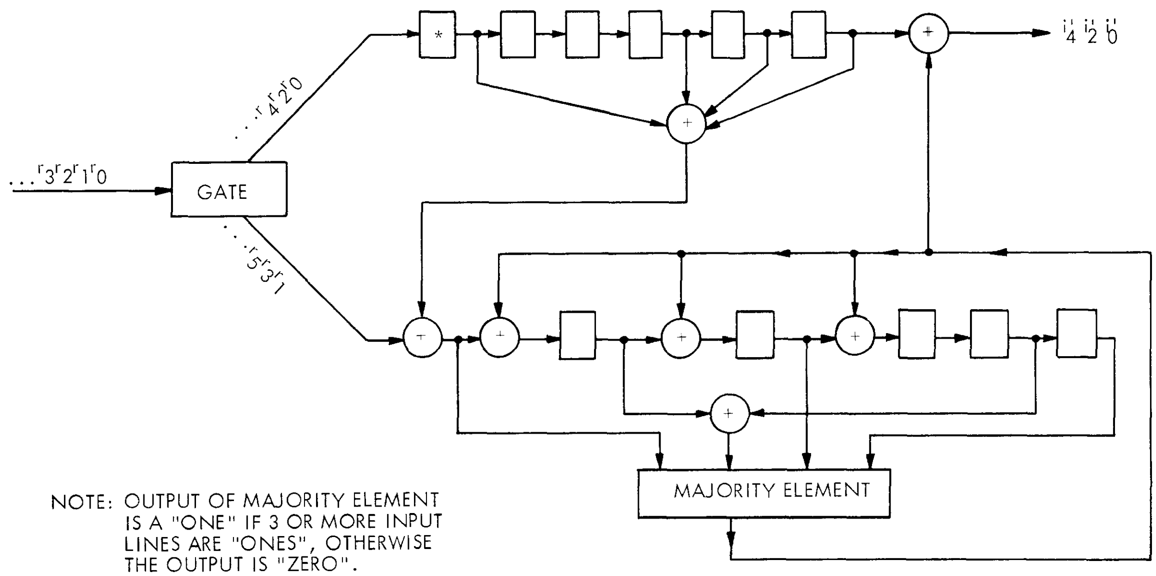

Fig. XX-4. Decoder for $\left[12, \frac{1}{2}\right]$ code.

$\left[12, \frac{1}{2}\right]$ code is shown in Fig. XX-3. A plain box indicates a shift register stage with two units of delay; a starred box indicates one unit of delay.

Majority decoding can be instrumented very simply by carrying out the operations indicated by Eqs. 9 and 14. The decoder for the previous code is shown in Fig. XX-4. The received bits enter in a stream and the decoded information bits flow out in a stream at the same rate, but delayed by 11 units of time. The circuit is self-explanatory except, perhaps, for the feedback of the majority output to 3 adders in the lower section. This is done to remove the effect of e_o on the parity checks in order to prepare for the determination of e_2 at the next shift.

In general, the decoder requires 2c − 1 shift registers, where c is the number of check symbols in the constraint length n. The decoder also requires a simple majority logic circuit, a simple gating element, and approximately 2c adders.

5. Extensions

The results above are easily generalized for the case in which the rate is any rational fraction. It is still an open question whether or not any convolutional code can be decoded up to its guaranteed error-correcting ability by using a majority-rule decoding algorithm. We suspect that this question can be answered in the affirmative, and we hope that either a proof of this or a counterexample will soon be forthcoming.

J. L. Massey

References

1. P. Elias, Coding for noisy channels, IRE Convention Record, Part IV, 1955, pp. 37-44.

2. J. M. Wozencraft and B. Reiffen, Sequential Decoding (The M. I. T. Press, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1961), pp. 54-55.