BINARY CODING USING STANDARD RUN LENGTHS

Thesis by

John I. Molinder

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1969

(Submitted May 12, 1969)

ACKNOWLEDGMENT

I greatly appreciate the guidance of my research advisor Dr. T. L. Grettenberg who suggested the problem and provided advice and direction throughout the course of the research. Several useful discussions with Dr. Jerry Butman are appreciated.

I also appreciate the Financial Support provided by a Ford Foundation Fellowship, a Teaching Assistantship, and the GI Bill during my graduate studies at Caltech.

Last but by no means least I appreciate the excellent typing of Mrs. Doris Schlicht, without whom the deadline for submission of this thesis for June graduation would have never been met.

ABSTRACT

Run length coding using standard run lengths has been proposed by Cherry et al [7]. Their analysis has been mostly experimental for specific types of data.

In this thesis the globally optimum single standard run length has been derived for the binary independent source and globally optimum single standard run lengths of zeros and ones have been derived for the binary first order Markov source. It is assumed that the output symbols are subsequently block coded in each case. A recursion relationship between standard run lengths is derived for two specific coding algorithms. A simple single standard run length scheme using a non-block code on the output symbols has also been derived for the binary independent source.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

INTRODUCTION

A field of interest to communications engineers has been the minimization of the amount of data required to be transmitted to describe the behavior of a random source. This field is known by various names including noiseless coding, redundancy reduction, and data compression. Various schemes have been described in the literature [2],[3],[4],[5],[6],[7],[9],[10]. The theoretical performance limit of any such scheme is of course that derived by Shannon [8]. A large portion of the analysis of various data compression schemes has been experimental. Davisson [3], Ehrman [4], and Tunstall [9] have only recently theoretically analyzed some of the schemes by assuming a specific source model. This is the approach followed in this thesis.

Efficient coding for an unsymmetrical binary independent or Markov source may be attained by Huffman coding an extension of the original source rather than the source itself. As the lack of symmetry increases a higher extension must be coded to maintain a given efficiency. This requires an increasing number of code symbols.

Another scheme is to use run length coding. Here the number of successive zeros say, up to some maximum run length, is transmitted rather than the zeros themselves. Again to increase the maximum run length encoded (and thus the efficiency) requires increasing the number of code symbols.

A different approach is to decide to use $n \geq 2$ code symbols where each symbol represents a fixed run length of zeros or ones. To insure all possible sequences can be encoded, two symbols must be used to represent a zero and one respectively. This leaves $n - 2$ symbols to be chosen. The technique is known as run length coding using standard run lengths and the problem now is to choose these standard run lengths optimally. This technique has been studied experimentally by Cherry et al [2] with the best standard run lengths for a specific type of data being determined by exhaustive search.

In this thesis the globally optimum single standard run length has been derived for the binary independent source and globally optimum single standard run lengths of zeros and ones have been derived for the binary first order Markov source. It is assumed that the output symbols are subsequently block coded in each case. Maxima have been found for the binary independent source when Huffman coding is subsequently used to code the output symbols and in some cases these have been shown to be global optimums. A recursion relationship between standard run lengths is derived for two specific coding algorithms. This recursion relationship holds for an arbitrary number of standard run lengths. A simple single standard run length scheme using a non-block code on the output symbols has also been derived for the binary independent source.

CHAPTER I

CODING TECHNIQUE

1.1. Introduction.

In this thesis a binary source is coded into $n \geq 2$ code symbols where each symbol represents a fixed run length of zeros or ones. To insure all possible sequences can be encoded two symbols must be used to represent a zero and a one respectively. This leaves $n - 2$ symbols to be chosen. The problem now is to choose these standard run lengths optimally.

1.2. Optimality Criterion.

The optimality criterion selected for this thesis is the maximization of the compression ratio. The compression ratio is defined as the expected ratio of the number of binary digits in the input sequence to the number of binary digits in the output sequence as the length of the input sequence tends to infinity. The optimal code is then defined by the standard run lengths that maximize the compression ratio. As will be pointed out later, the formulation of the problem is general enough so that cost functions other than the length of the output sequence can be used. This does not change the method of analysis, however.

CHAPTER II.

OPTIMAL RUN LENGTH CODING USING ONE STANDARD RUN LENGTH FOR THE INDEPENDENT BINARY SOURCE

2.1. Introduction.

In this chapter the optimal single run length is determined for the binary independent source. Of course runs of the most likely symbol are encoded which is arbitrarily chosen to be 0. In the next chapter the optimum single run lengths of 0's and 1's for the binary first order Markov source are derived. Since a first order Markov source may be made equivalent to an independent source by assigning appropriate transition probabilities, this chapter is really a special case of the following one. The analysis is much more straightforward for the independent source, however, and it clearly illustrates the method of analysis used in the following chapter. For this reason analysis of the independent source is given separately.

2.2. Definition of Coding Technique.

An independent binary source emitting zeros and ones with probabilities q and $p = 1-q$ respectively where $q \gg p$ is encoded as follows:

$$0 \rightarrow x_1$$

$$1 \rightarrow x_2$$

$$N \text{ 0's in a row} \rightarrow x_3$$

The operation of the coder may be defined by observing that no action is taken until the occurrence of one of the following two events:

A. A one is reached in the input sequence, or

B. N zeros have been accumulated.

Thus the coder operation may be viewed as a mapping of certain input sequences into their corresponding output sequences as shown below.

$$
\begin{aligned}
1 &\rightarrow x_2 \\
01 &\rightarrow x_1 x_2 \\
001 &\rightarrow x_1 x_1 x_2 \\
&\vdots \\
\underbrace{0 \cdots 0}_{N-1} 1 &\rightarrow \underbrace{x_1 \cdots x_1}_{N-1} x_2 \\
\underbrace{0 \cdots 0\,0}_{N} &\rightarrow x_3
\end{aligned}
\tag{2.1}
$$

The mapping of one of the above input sequences into the corresponding output sequence will be denoted as a coder action (CA).

2.3. Definition of Compression Ratio.

The compression ratio (CR) is defined in Chapter I to be the expected ratio of the number of binary digits in the input sequence to the number of binary digits in the output sequence as the length of the input sequence tends to infinity. In this case this reduces to

$$CR = \lim_{n \to \infty} \frac{n}{n(x_1)\ell_1 + n(x_2)\ell_2 + n(x_3)\ell_3} \tag{2.2}$$

where

n = number of input symbols

$n(x_i)$= number of x_i's $(i = 1,2,3)$ in the output sequence

ℓ_i = cost of the code word for $x_i (i = 1,2,3)$ in binary digits.

The optimum code is then defined by the N that maximizes the compression ratio (2.2). Obviously ℓ_1, ℓ_2, and ℓ_3 may be considered as the cost of outputing an x_1, x_2, or x_3 respectively rather than the length of the code words. This does not change the method of analysis, however.

2.4. Derivation of Compression Ratio in Terms of Coder Actions.

From (2.1) it is evident that the probability that a coder action results in an output consisting of a string of J x_1's $(0 \le J \le N-1)$ followed by an x_2 is given by

$$P_{CA}(Jx_1\text{'s},x_2) = P(J0\text{'s},1) = pq^J \quad (0 \le J \le N-1)$$

while the probability that a coder action outputs an x_3 is given by

$$P_{CA}(x_3) = P(N\ 0\text{'s}) = q^N$$

Thus the expected number of x_1's, x_2's, and x_3's emitted per coder action is given by

$$E(x_1) = \sum_{J=1}^{N-1} Jpq^J = \frac{q[1+(N-1)q^N-Nq^{N-1}]}{p}$$

$$E(x_2) = \sum_{J=1}^{N-1} pq^J = 1-q^N \tag{2.3}$$

$$E(x_3) = q^N$$

Now consider Q coder actions and let

$$m_i = \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_i) \qquad (i = 1,2,3)$$

where $n_j(x_i)$ is the number of x_i's occurring on the jth coder action. Since the coder actions are independent, the weak law of large numbers [11] gives

$$P[\,|m_i - E(x_i)| \geq \epsilon] \leq \frac{\sigma_i^2}{Q\epsilon^2} \quad (i = 1,2,3)$$

where

$$\sigma_1 = \sum_{J=1}^{N-1} J^2 pq^J - [E(x_1)]^2 < \infty$$

$$\sigma_2 = \sum_{J=1}^{N-1} pq^J - [E(x_2)]^2 < \infty$$

$$\sigma_3 = q^N - [E(x_3)]^2 < \infty$$

Thus

$$\lim_{Q\to\infty} \frac{1}{Q} \sum_{y=1}^{Q} n_j(x_i) = E(x_i) \qquad (i = 1,2,3) \tag{2.4}$$

with probability one. The compression ratio (2.2) may be written as

$$CR = \lim_{Q\to\infty} \frac{\sum_{j=1}^{Q} n_j(x_1) + \sum_{j=1}^{Q} n_j(x_2) + N \sum_{j=1}^{Q} n_j(x_3)}{\ell_1 \sum_{j=1}^{Q} n_j(x_1) + \ell_2 \sum_{j=1}^{Q} n_j(x_2) + \ell_3 \sum_{j=1}^{Q} n_j(x_3)}$$

Dividing numerator and denominator by Q

$$CR = \lim_{Q\to\infty} \frac{\frac{1}{Q} \sum_{j=1}^{Q} n_j(x_1) + \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_2) + N \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_3)}{\ell_1 \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_1) + \ell_2 \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_2) + \ell_3 \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_3)} \tag{2.5}$$

Substituting (2.4) into (2.5)

$$CR = \frac{E(x_1) + E(x_2) + NE(x_3)}{\ell_1 E(x_1) + \ell_2 E(x_2) + \ell_3 E(x_3)} \tag{2.6}$$

with probability one where $E(x_1)$, $E(x_2)$, and $E(x_3)$ are given in (2.3).

2.5. Optimal Code for Output Symbols of Equal Length.

If $\ell_1 = \ell_2 = \ell_3 = L$ (2.6) may be written

$$CR = \frac{E(x_1) + E(x_2) + NE(x_3)}{L[E(x_1) + E(x_2) + E(x_3)]} = \frac{1}{L}\left[1 + \frac{(N-1)\,E(x_3)}{E(x_1) + E(x_2) + E(x_3)}\right] \tag{2.7}$$

Substituting (2.3) into (2.7) and reducing yields

$$CR = \frac{1}{L}\left[1 + \frac{(N-1)pq^N}{1 + (N-1)q^{N+1} - Nq^N}\right] \tag{2.8}$$

To maximize (2.8) it is necessary only to maximize

$$\frac{(N-1)pq^N}{1 + (N-1)q^{N+1} - Nq^N} \tag{2.9}$$

Differentiating (2.9) with respect to N, combining terms and setting the result equal to zero yields

$$\frac{pq^N}{[1+(N-1)q^{N+1}-Nq^N]^2}\,[1-q^N+(N-1)\ln q] = 0 \tag{2.10}$$

Since

$$q^N > 0 \text{ and}$$

$$\begin{aligned} 1 + (N-1)q^{N+1} - Nq^N &= 1 - q + q[1+(N-1)q^N-Nq^{N-1}] \\ &= 1 - q + qp^2 \sum_{J=1}^{N-1} Jq^J > 0 \end{aligned} \tag{2.11}$$

the left hand side of (2.10) is equal to zero only if

$$(N-1)(-\ln q) = 1 - q^N \tag{2.12}$$

The graphical solution of this implicit equation is shown in Figure 1.

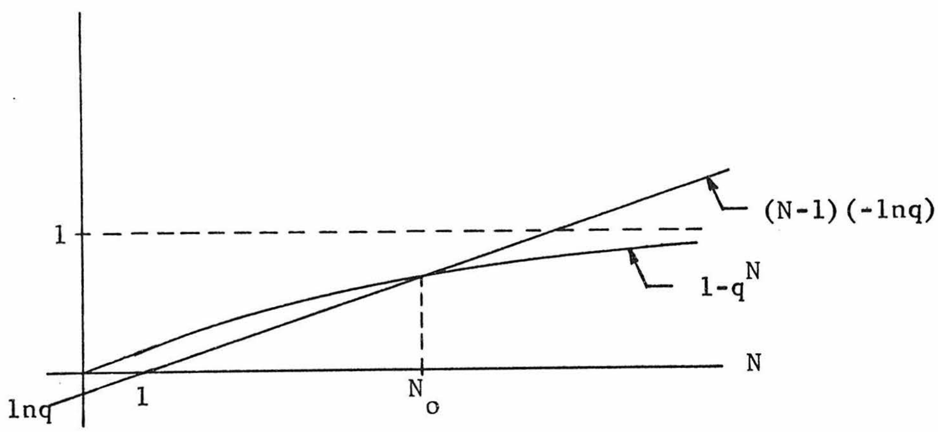


Figure 1

Graphical Solution of $(N-1)(-\ln q) = 1 - q^N$

From Figure 1 and (2.11) it can be seen that if N is decreased from N_o (2.10) is positive while if N is increased from N_o (2.10) is negative. This means that the slope of (2.10) (or equivalently the second derivative of (2.8)) is negative at N_o assuring that N_o determined a maximum. It is geometrically evident from Figure 1 that there is only one solution to (2.12). Thus the integer $N = N_o$ most nearly satisfying (2.12) defines the globally optimum run length within ± 1.

Encoding three output symbols requires a block code length $L = 2$. The solution of (2.12) and the resulting compression ratios for various values of p are given in Table 1. Plots of the optimum N and compression ratio vs. p are given in Figure 2 and Figure 3 respectively at the end of the chapter.

Table 1

Optimum N and CR when output symbols are block coded (L=2)

p	N	CR
0.5000	1	1.000
0.2000	1	1.000
0.1000	5	1.181
0.0500	7	1.636
0.0300	8	2.093
0.0200	10	2.551
0.0100	14	3.583
0.0050	20	5.046
0.0030	26	6.500
0.0020	32	7.951
0.0015	37	9.173
0.0010	45	11.224

2.6. Optimal Code When Huffman Coding is Used to Code Output Symbols.

Block coding is not the optimum way to encode the output symbols. The best way to encode symbols with given probabilities is with the Huffman coding algorithm. To use this algorithm, however, the probabilities of the symbols must be known. The probabilities of x_1, x_2, and x_3 may be defined as the limit of their frequency ratio as the length of the input sequence tends to infinity. Thus

$$P(x_i) = \lim_{n\to\infty} \frac{n(x_i)}{n(x_1) + n(x_2) + n(x_3)} \quad (i = 1,2,3) \tag{2.13}$$

where $n(x_i)$ $(i = 1,2,3)$ is the number of x_i(s) in the output sequence and n is the number of binary digits in the input sequence. In terms of register actions (2.13) may be written

$$P(x_i) = \lim_{Q\to\infty} \frac{\sum_{j=1}^{Q} n(x_i)}{\sum_{j=1}^{Q} [n_j(x_1)+n_j(x_2)+n_j(x_3)]} \tag{2.14}$$

Dividing numerator and denominator by Q and using (2.4) yields

$$P(x_i) = \frac{E(x_i)}{E(x_1) + E(x_2) + E(x_3)} \qquad (i = 1,2,3)$$

with probability one. The optimum N and resulting compression ratio may now be determined by computer search. The values of $P(x_i)$ $(i = 1,2,3)$ are calculated for $N = 2,3, \cdots$, the Huffman algorithm is applied at each step to determine ℓ_1, ℓ_2, and ℓ_3, the compression ratio is determined according to (2.6), and the N yielding the maximum value of the compression ratio (2.6) is selected. Note that this is a fundamentally different process than applying Huffman coding to the optimum N selected for block coding by the method discussed in the previous section. It should also be pointed out that only a finite search is required to determine the globally optimum N for the

Huffman case. This may be shown as follows. Rewriting (2.6) yields

$$CR = \frac{E(x_1) + E(x_2) + NE(x_3)}{[\ell_1 P(x_1)+\ell_2 P(x_2)+\ell_3 P(x_3)][E(x_1)+E(x_2)+E(x_3)]} \qquad (2.15)$$

But this is just the compression ratio for the block coding case with the average code length replacing L. Now clearly

$$\bar{\ell} = \ell_1 P(x_1) + \ell_2 P(x_2) + \ell_3 P(x_3) \geq 1$$

and from the previous section, (2.7) and (2.8) the quantity

$$\frac{E(x_1) + E(x_2) + NE(x_3)}{E(x_1) + E(x_2) + E(x_3)} \qquad (2.16)$$

is a monotonically decreasing function of N approaching 1 for $N > N_o$ (since it has only one maximum). Thus the search need only be carried out until (2.16) is less than or equal to the maximum of (2.15) up to that point. The results of the computer search are given in Table 2 and plotted in Figures 2 and 3. The points for which the search has been carried out far enough to guarantee a global maximum are marked with an asterisk. A comparison of the efficiency of this scheme with various other coding schemes is given in Figure 9 at the end of Chapter IV.

Table 2

Optimum N and CR when output symbols are Huffman coded

p	N	CR
0.5000	1	1.000
0.2000	5	1.102
0.1000	7	1.559*
0.0500	10	2.207*
0.0300	12	2.856*
0.0200	15	3.503*
0.0100	21	4.964*
0.0050	29	7.034*
0.0030	37	9.091*
0.0020	45	11.141*
0.0015	52	12.871*
0.0010	64	15.772

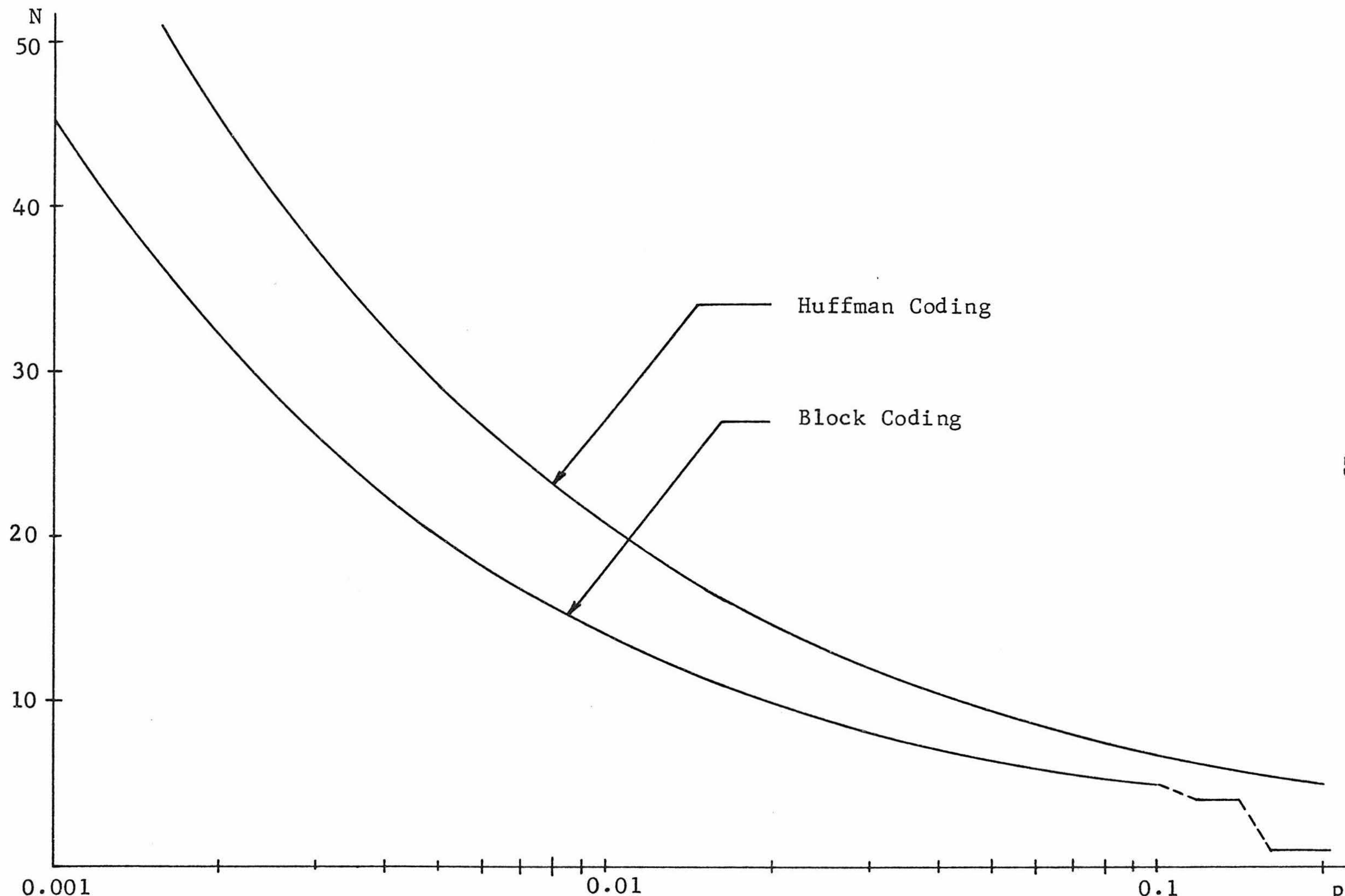


Figure 2

Optimum N vs p for single run length

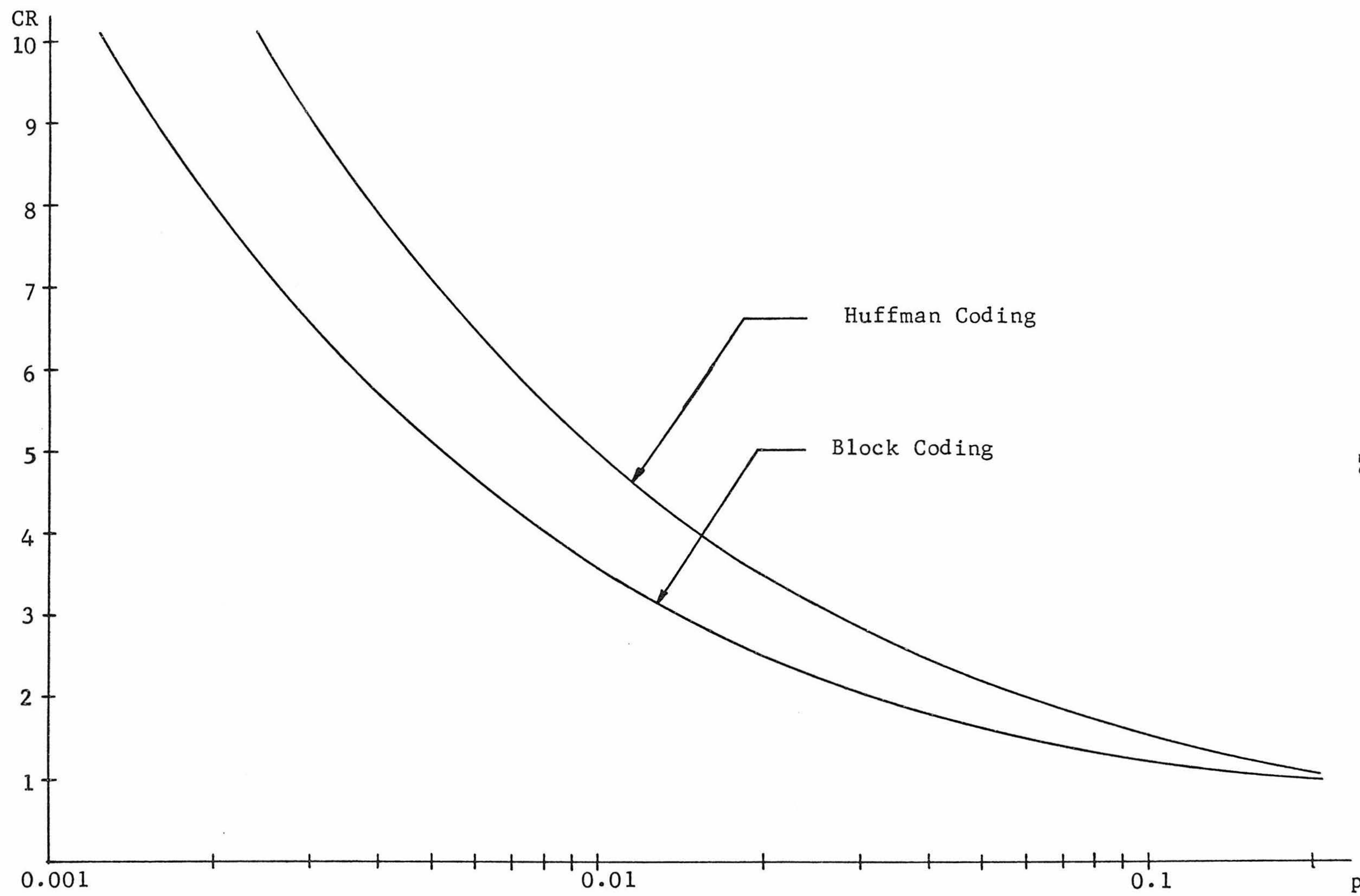


Figure 3

Compression ratio vs p for single run length

CHAPTER III

OPTIMUM SINGLE RUN LENGTHS OF 0'S AND 1'S FOR THE BINARY FIRST ORDER MARKOV SOURCE

3.1. Introduction.

As was pointed out in Section 2.1, Chapter II is really a special case of Chapter III. The independent source is considerably easier to analyze, however, and it clearly illustrates the basic method used in both chapters. For this reason the analysis of the independent source was given separately in Chapter II.

3.2. Definition of Coding Technique.

A binary first order Markov source is defined by the following transition probabilities

$$P(0|0) = q_o \qquad P(0|1) = p_1$$
$$P(1|0) = p_o \qquad P(1|1) = q_1$$

where $p_o = 1-q_o$ and $p_1 = 1-q_1$. This corresponds to the state diagram shown in Figure 4.

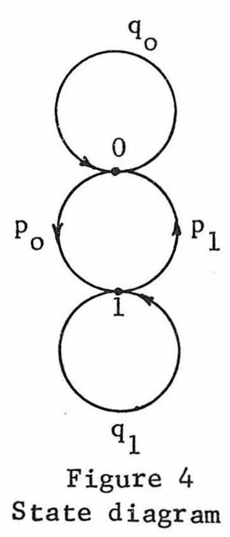


Figure 4
State diagram

This source is then encoded as follows:

$$\begin{aligned} 0 &\rightarrow x_1 \\ 1 &\rightarrow x_2 \\ K\ 0\text{'s in a row} &\rightarrow x_3 \\ N\ 1\text{'s in a row} &\rightarrow x_4 \end{aligned} \qquad (3.1)$$

The operation of the coder may be defined by observing that no action is taken until the occurrence of one of the following events:

A. the source changes from state 0 to state 1

B. the source changes from state 1 to state 0

C. K 0's have been accumulated

D. N 1's have been accumulated.

If the source changes from state 0 to 1 (event A) the J 0's ($1 \leq J \leq K-1$) which have been accumulated thus far are coded as J x_1's and the 1 produced by the state change is stored until it is determined whether or not N-1 additional 1's in a row will occur (thus allowing coding into an x_4). The source is in state 1 at the end of the coding operation. If K 0's have been accumulated (event C) they are coded as an x_3, no input symbol is stored, and the source is in state 0 at the end of the coder operation. Similar arguments apply to events B and D. Thus the probability of a certain coder operation is dependent on whether the preceding coder operation was triggered by event A, B, C, or D. As in Chapter II the coder operation may be defined as a mapping of certain input sequences into their corresponding output sequences as shown in Figure 5. This mapping is again denoted as a coder action (CA). Event C is equivalent to a coder output of x_3 and event D is equivalent to a coder output of x_4. Thus to simplify notation, events C and D are denoted x_3 and x_4 respectively for the remainder of the chapter.

Triggering Event	Coder Action	Remarks
A	$01 \to x_1$ $001 \to x_1 x_1$ $\vdots$ $\underbrace{0\cdots01}_{K-1} \to \underbrace{x_1\cdots x_1}_{K-1}$	A 1 remains to be coded. The source is left in state 1.
B	$10 \to x_2$ $110 \to x_2 x_2$ $\vdots$ $\underbrace{1\cdots10}_{N-1} \to \underbrace{x_2\cdots x_2}_{N-1}$	A 0 remains to be coded. The source is left in state 0.
C	$\underbrace{0\cdots00}_{K} \to x_3$	Nothing remains to be coded. The source is left in state 0.
D	$\underbrace{1\cdots11}_{N} \to x_4$	Nothing remains to be coded. The source is left in state 1.

Figure 5

Coder actions for binary first order Markov source

3.3. Definition of Compression Ratio.

The compression ratio (CR) is defined in Chapter I to be the expected ratio of the number of binary digits in the input sequence to the number of binary digits in the output sequence as the length of the input sequence tends to infinity. In this case this reduces to

$$CR = \lim_{n\to\infty} \frac{n}{n(x_1)\ell_1 + n(x_2)\ell_2 + n(x_3)\ell_3 + n(x_4)\ell_4} \tag{3.2}$$

where

n = number of input symbols

$n(x_i)$ = number of x_i's $(i = 1,2,3,4)$ in the output sequence

ℓ_i = cost of the code word for x_i $(i = 1,2,3,4)$ in binary digits.

The optimum code is defined by the K and N that maximize the compression ratio (3.2).

3.4. Derivation of the Compression Ratio in Terms of Coder Actions.

Referring to Figures 4 and 5 and using the reasoning of Section 3.2 the probabilities of the possible coder actions conditioned on the previous coder action may be determined as follows.

$$P_{CA}(Jx_1\text{'s}\,|\,A) = 0 \qquad P_{CA}(Jx_2\text{'s}\,|\,A) = p_1 q_1^{J-1}$$

$$P_{CA}(Jx_1\text{'s}\,|\,B) = p_o q_o^{J-1} \qquad P_{CA}(Jx_2\text{'s}\,|\,B) = 0$$

$$P_{CA}(Jx_1\text{'s}\,|x_3) = p_o q_o^{\,J} \qquad P_{CA}(Jx_2\text{'s}\,|x_3) = p_o p_1 q_1^{\,J-1}$$

$$P_{CA}(Jx_1\text{'s}\,|x_4) = p_o p_1 q_o^{\,J-1} \qquad P_{CA}(Jx_2\text{'s}\,|x_4) = p_1 q_1^{\,J} \qquad (3.3)$$

$$(J = 1, \cdots, K-1) \qquad (J = 1, \cdots, N-1)$$

$$P_{CA}(x_3|A) = 0 \qquad P_{CA}(x_4|A) = q_1^{\,N-1}$$

$$P_{CA}(x_3|B) = q_o^{\,K-1} \qquad P_{CA}(x_4|B) = 0$$

$$P_{CA}(x_3|x_3) = q_o^{\,K} \qquad P_{CA}(x_4|x_3) = p_o q_1^{\,N-1}$$

$$P_{CA}(x_3|x_4) = p_1 q_o^{\,K-1} \qquad P_{CA}(x_4|x_4) = q_1^{\,N}$$

Thus the conditional expectations of the number of x_1's, x_2's, x_3's, and x_4's emitted per coder action are given by

$$E(x_1|A) = 0 \qquad E(x_2|A) = \sum_{J=1}^{N-1} J p_1 q_1^{\,J-1}$$

$$E(x_1|B) = \sum_{J=1}^{K-1} J p_o q_o^{\,J-1} \qquad E(x_2|B) = 0$$

$$E(x_1|x_3) = \sum_{J=1}^{K-1} J p_o q_o^{\,J} \qquad E(x_2|x_3) = \sum_{J=1}^{N-1} J p_o p_1 q_1^{\,J-1}$$

$$E(x_1|x_4) = \sum_{J=1}^{K-1} p_o p_1 q_o^{\,J-1} \qquad E(x_2|x_4) = \sum_{J=1}^{N-1} J p_1 q_1^{\,J} \qquad (3.4)$$

$$E(x_3|A) = 0 \qquad E(x_4|A) = q_1^{\,N-1}$$

$$E(x_3|B) = q_o^{\,K-1} \qquad E(x_4|B) = 0$$

$$E(x_3|x_3) = q_o^K \qquad E(x_4|x_3) = p_o q_1^{N-1}$$

$$E(x_3|x_4) = p_1 q_o^{K-1} \qquad E(x_4|x_4) = q_1^N$$

Since A, B, C, and D are disjoint events whose union covers the probability space of coder actions

$$E(x_i) = E(x_i|A)P_{CA}(A) + E(x_i|B)P_{CA}(B) + E(x_i|x_3)P_{CA}(x_3) + E(x_i|x_4)P_{CA}(x_4)$$

$$(i = 1,2,3,4) \qquad (3.5)$$

where $P_{CA}(A)$ is the stationary probability of event A, etc. Now consider Q coder actions and let

$$m(x_i|z) = \frac{1}{Q}\sum_{j=1}^{Q} n_j(x_i|z) \qquad (i = 1,2,3,4) \quad (z = A,B,x_3,x_4)$$

where $n_j(x_i|z)$ is the number of x_i's occurring on the jth coder action given that the previous coder action belonged to event z. Since the conditional coder actions are independent, the weak law of large numbers [11] gives

$$[P|m(x_i|z) - E(x_i|z)| \geq \epsilon] \leq \frac{\sigma(x_i|z)^2}{Q\epsilon^2} \qquad (i = 1,2,3,4) \quad (z = A,B,x_3,x_4)$$

where

$$\sigma(x_1|B) = \sum_{J=1}^{K-1} J^2 p_o q_o^J - [E(x_1|B)]^2 < \infty \quad \text{etc.}$$

Thus

$$\lim_{Q\to\infty} \frac{1}{Q} \sum_{j=1}^{Q} n_j(x_i|z) = E(x_i|z) \qquad \begin{matrix}(i = 1,2,3,4) \\ (z = A,B,x_3,x_4)\end{matrix} \tag{3.6}$$

with probability one.

The source may equivalently be thought of as having states A, B, x_3, x_4 with transitional probabilities $P_{CA}(A|A)$, $P_{CA}(A|B)$, etc. It has been shown [1] that

$$\lim_{Q\to\infty} \frac{n(z)}{Q} = P_{CA}(z) \qquad (z = A,B,x_3,x_4) \tag{3.7}$$

where $P_{CA}(z)$ are the unconditional state probabilities.

The compression ratio (3.2) may be written

$$CR = \lim_{Q\to\infty} \frac{\sum\limits_{z=A,B,x_3,x_4} \left\{ n(z) \sum\limits_{j=1}^{Q} [n_j(x_1|z)+n_j(x_2|z)+Kn_j(x_3|z)+Nn_j(x_4|z)] \right\}}{\sum\limits_{z=A,B,x_3,x_4} \left\{ n(z) \sum\limits_{j=1}^{Q} [\ell_1 n_j(x_1|z)+\ell_2 n_j(x_2|z)+\ell_3 n_j(x_3|z)+\ell_4 n_j(x_4|z)] \right\}}$$

Dividing numerator and denominator by Q^2 and substituting (3.6) and (3.7) yields

$$CR = \frac{E(x_1) + E(x_2) + KE(x_3) + NE(x_4)}{\ell_1 E(x_1) + \ell_2 E(x_2) + \ell_3 E(x_3) + \ell_4 E(x_4)} \tag{3.8}$$

where $E(x_i)$ $(i = 1,2,3,4)$ is given in (3.5).

$P_{CA}(A)$, $P_{CA}(B)$, $P_{CA}(x_3)$, and $P_{CA}(x_4)$ must now be determined to specify $E(x_i)$ $(i = 1,2,3,4)$. This may be done by observing that the stationary probabilities of these events must satisfy the following equations. Since all probabilities refer to coder actions, the subscript CA will be dropped throughout the derivation for notational convenience.

$$P(A)+P(B)+P(x_3)+P(x_4) = 1$$

$$P(x_4|A)P(A)+P(x_4|B)P(B)+P(x_4|x_3)P(x_3)+P(x_4|x_4)P(x_4) = P(x_4)$$

$$P(x_3|A)P(A)+P(x_3|B)P(B)+P(x_3|x_3)P(x_3)+P(x_3|x_4)P(x_4) = P(x_3) \qquad (3.9)$$

$$P(A|A)P(A)+P(A|B)P(B)+P(A|x_3)P(x_3)+P(A|x_4)P(x_4) = P(A)$$

$$P(B|A)P(A)+P(B|B)P(B)+P(B|x_3)P(x_3)+P(B|x_4)P(x_4) = P(B)$$

Of course these five equations are dependent since there are only four unknowns. The first four equations will be used.

$P(x_3|z)$ and $P(x_4|z)$ $(z = A,B,x_3,x_4)$ are given in (3.3). Also from (3.3)

$$P(A|A) = 0$$

$$P(A|B) = \sum_{J=1}^{K-1} p_o q_o^{J-1} = 1-q_o^{K-1}$$

$$P(A|x_3) = \sum_{J=1}^{K-1} p_o q_o^{\,J} = q_o(1-q_o^{\,K-1})$$

$$P(A|x_4) = \sum_{J=1}^{K-1} p_o p_1 q_o^{\,J-1} = p_1(1-q_o^{\,K-1})$$

Thus the first four equations of (3.9) become

$$P(A)+P(B)+P(x_3)+P(x_4) = 1 \tag{3.10a}$$

$$q_1^{\,N-1}P(A)+p_o q_1^{\,N-1}P(x_3)+q_1^{\,N}P(x_4) = P(x_4) \tag{3.10b}$$

$$q_o^{\,K-1}P(B)+q_o^{\,K}P(x_3)+p_1 q_o^{\,K-1}P(x_4) = P(x_3) \tag{3.10c}$$

$$(1-q_o^{\,K-1})P(B)+q_o(1-q_o^{\,K-1})P(x_3)+p_1(1-q_o^{\,K-1})P(x_4) = P(A) \tag{3.10d}$$

Solving (3.10b) for $P(A)$

$$P(A) = -\,p_o P(x_3) + \frac{(1-q_1^{\,N})}{q_1^{\,N-1}}\,P(x_4) \tag{3.11}$$

Solving (3.10c) for $P(B)$

$$P(B) = \frac{(1-q_o^{\,K})}{q_o^{\,K-1}}\,P(x_3) - p_1 P(x_4) \tag{3.12}$$

Dividing (3.10d) by $(1-q_o^{\,K-1})$ and rewriting

$$- \frac{1}{1-q_o^{K-1}} P(A) + P(B) + q_o P(x_3) + p_1 P(x_4) = 0 \tag{3.13}$$

Substituting (3.11) and (3.12) into (3.13) and reducing

$$\frac{(1-q_o^K)}{q_o^{K-1}} P(x_3) - \frac{(1-q_1^N)}{q_1^{N-1}} P(x_4) = 0 \tag{3.14}$$

Substituting (3.11) and (3.12) into (3.10a) and reducing

$$\frac{1}{q_o^{K-1}} P(x_3) + \frac{1}{q_1^{N-1}} P(x_4) = 1 \tag{3.15}$$

Solving (3.14) and (3.15) for $P(x_3)$ and $P(x_4)$

$$P(x_3) = \frac{q_o^{K-1}(1-q_1^N)}{(1-q_o^K) + (1-q_1^N)} \tag{3.16}$$

$$P(x_4) = \frac{q_1^{N-1}(1-q_o^K)}{(1-q_o^K) + (1-q_1^N)} \tag{3.17}$$

Substituting (3.16) and (3.17) into (3.11) and (3.12) and reducing

$$P(A) = \frac{(1-q_o^{K-1})(1-q_1^N)}{(1-q_o^K) + (1-q_1^N)}$$

$$P(B) = \frac{(1-q_o^K)(1-q_1^{N-1})}{(1-q_o^K) + (1-q_1^N)}$$

Reinserting the CA notation and summarizing the results

$$P_{CA}(A) = \frac{(1-q_o^{K-1})(1-q_1^N)}{(1-q_o^K) + (1-q_1^N)}$$

$$P_{CA}(B) = \frac{(1-q_o^K)(1-q_1^{K-1})}{(1-q_o^K) + (1-q_1^N)} \qquad (3.18)$$

$$P_{CA}(x_3) = \frac{q_o^{K-1}(1-q_1^N)}{(1-q_o^K) + (1-q_1^N)}$$

$$P_{CA}(x_4) = \frac{q_1^{N-1}(1-q_o^K)}{(1-q_o^K) + (1+q_1^N)}$$

Note that for $K = 1$, $N = 1$, the state probabilities reduce to

$$P_{CA}(A) = 0 \qquad P_{CA}(x_3) = \frac{p_1}{p_o + p_1}$$

$$P_{CA}(B) = 0 \qquad P_{CA}(x_4) = \frac{p_o}{p_o + p_1}$$

where $P_{CA}(x_3)$ and $P_{CA}(x_4)$ are just the state probabilities of 0 and 1 respectively.

Using (3.3), (3.5), and (3.18), the expected number of x_i's ($i = 1,2,3,4$) emitted per coder action is given by

$$E(x_1) = P_{CA}(B) \sum_{J=1}^{K-1} Jp_o q_o^{J-1} + P_{CA}(x_3) \sum_{J=1}^{K-1} Jp_o q_o^{J} + P_{CA}(x_4) \sum_{J=1}^{K-1} Jp_o p_1 q_o^{J-1}$$

$$= p_o \sum_{J=1}^{K-1} Jq_o^{J-1} \left[P_{CA}(B) + q_o P_{CA}(x_3) + p_1 P_{CA}(x_4)\right]$$

$$= \frac{[1-Kq_o^{K-1} + (K-1)q_o^{K}](1-q_1^{N})}{p_o[(1-q_o^{K}) + (1-q_1^{N})]} \qquad (3.19a)$$

Similarly

$$E(x_2) = \frac{[1-Nq_1^{N-1} + (N-1)q_1^{N}](1-q_o^{K})}{p_1[(1-q_o^{K}) + (1-q_1^{N})]} \qquad (3.19b)$$

$$E(x_3) = P_{CA}(x_3) = \frac{q_o^{K-1}(1-q_1^{N})}{(1-q_o^{K}) + (1-q_1^{N})} \qquad (3.19c)$$

$$E(x_4) = P_{CA}(x_4) = \frac{q_1^{N-1}(1-q_o^{K})}{(1-q_o^{K}) + (1-q_1^{N})} \qquad (3.19d)$$

In summary, the compression ratio is given by (3.8)

$$CR = \frac{E(x_1) + E(x_2) + KE(x_3) + NE(x_4)}{\ell_1 E(x_1) + \ell_2 E(x_2) + \ell_3 E(x_3) + \ell_4 E(x_4)} \qquad (3.8)$$

where $E(x_i)$ $(i = 1,2,3,4)$ is given by (3.19).

3.5. Optimal Code for Output Symbols of Equal Length.

If $\ell_1 = \ell_2 = \ell_3 = \ell_4 = L$ (3.8) may be written

$$CR = \frac{E(x_1)+E(x_2)+KE(x_3)+NE(x_4)}{L[E(x_1)+E(x_2)+E(x_3)+E(x_4)]} = \frac{1}{L}\left[1 + \frac{(K-1)E(x_3)+(N-1)E(x_4)}{E(x_1)+E(x_2)+E(x_3)+E(x_4)}\right] \tag{3.20}$$

Substituting (3.19) into (3.20) and reducing yields

$$CR = \frac{1}{L}\left[1 + \frac{p_o p_1 (K-1) q_o^{K-1}(1-q_1^N) + p_o p_1 (N-1) q_1^{N-1}(1-q_o^K)}{p_1[1-(K-1)q_o^{K-1}+(K-2)q_o^K](1-q_1^N)+p_o[1-(N-1)q_1^{N-1}+(N-2)q_1^N](1-q_o^K)}\right] \tag{3.21}$$

To maximize (3.21) it is necessary only to maximize

$$\frac{p_o p_1 (K-1) q_o^{K-1}(1-q_1^N) + p_o p_1 (N-1) q_1^{N-1}(1-q_o^K)}{p_1[1-(K-1)q_o^{K-1}+(K-2)q_o^K](1-q_1^N)+p_o[1-(N-1)q_1^{N-1}+(N-2)q_1^N](1-q_o^K)} \tag{3.22}$$

Differentiating (3.22) wrt K and setting the result equal to zero yields

$$\begin{aligned} 0= \frac{p_o p_1}{Y^2}&\left\{p_1[1-(K-1)q_o^{K-1}+(K-2)q_o^K](1-q_1^N)+ p_o[1-(N-1)q_1^{N-1}+(N-2)q_1^N](1-q_o^K)\right\} \\ &\times\left\{(1-q_1^N)[(K-1)q_o^{K-1}\ln q_o + q_o^{K-1}] - (N-1)q_1^{N-1}q_o^K \ln q_o\right\} \\ &- \frac{p_o p_1}{Y^2}\left\{(K-1)q_o^{K-1}(1-q_1^N)+ (N-1)q_1^{N-1}(1-q_o^K)\right\}\left\{p_1[-q_o^{K-1} - (K-1)q_o^{K-1}\ln q_o \right. \\ &\left. + q_o^K+(K-2)q_o^K \ln q_o](1-q_1^N) - p_o[1-(N-1)q_1^{N-1}+(N-2)q_1^N]q_o^K \ln q_o\right\} \end{aligned} \tag{3.23}$$

where Y is the denominator of (3.22). Expanding the numerator of (3.23) gives (neglecting the constant $p_o p_1$)

①

$$\left\{p_1(1-q_1^N)^2[1-(K-1)q_o^{K-1}+(K-2)q_o^K][(K-1)q_o^{K-1}\ln q_o+q_o^{K-1}]\right\}$$

②

$$-\left\{p_1(1-q_1^N)[1-(K-1)q_o^{K-1}+(K-2)q_o^K](N-1)q_1^{N-1}q_o^K\ln q_o\right\}$$

③

$$+\left\{p_o(1-q_o^K)(1-q_1^N)[1-(N-1)q_1^{N-1}+(N-2)q_1^N][(K-1)q_o^{K-1}\ln q_o+q_o^{K-1}]\right\}$$

④

$$-\left\{p_o(1-q_o^K)[1-(N-1)q_1^{N-1}+(N-2)q_1^N](N-1)q_1^{N-1}q_o^K\ln q_o\right\}$$

⑤

$$-\left\{p_1(K-1)q_o^{K-1}(1-q_1^N)^2[-q_o^{K-1}-(K-1)q_o^{K-1}\ln q_o+q_o^K+(K-2)q_o^K\ln q_o]\right\}$$

⑥

$$+\left\{p_o(K-1)q_o^{K-1}(1-q_1^N)[1-(N-1)q_1^{N-1}+(N-2)q_1^N]q_o^K\ln q_o\right\}$$

⑦

$$-\left\{p_1(N-1)q_1^{N-1}(1-q_o^K)(1-q_1^N)[-q_o^{K-1}-(K-1)q_o^{K-1}\ln q_o+q_o^K+(K-2)q_o^K\ln q_o]\right\}$$

⑧

$$+\left\{p_o(N-1)q_1^{N-1}(1-q_o^K)[1-(N-1)q_1^{N-1}+(N-2)q_1^N]q_o^K\ln q_o\right\}$$

Terms 4 and 8 cancel. Regrouping the remaining terms

(1) and (5)

$$p_1(1-q_1^N)^2\Big\{[1-(K-1)q_o^{K-1}+(K-2)q_o^K][(K-1)q_o^{K-1}\ln q_o+q_o^{K-1}]$$

$$-[(K-1)q_o^{K-1}][-q_o^{K-1}-(K-1)q_o^{K-1}\ln q_o+q_o^K+(K-2)q_o^K\ln q_o]\Big\}$$

(3) and (6)

$$+\,p_o(1-q_1^N)[1-(N-1)q_1^{N-1}+(N-2)q_1^N]\Big\{[(K-1)q_o^{K-1}\ln q_o+q_o^{K-1}](1-q_o^K)$$

$$+(K-1)q_o^{K-1}q_o^K\ln q_o\Big\}$$

(2) and (7)

$$-\,p_1(1-q_1^N)(N-1)q_1^{N-1}\Big\{[1-(K-1)q_o^{K-1}+(K-2)q_o^K]q_o^K\ln q_o$$

$$+(1-q_o^K)[-q_o^{K-1}-(K-1)q_o^{K-1}\ln q_o+q_o^K+(K-2)q_o^K\ln q_o]\Big\}$$

Multiplying out the terms in { } and reducing

$$q_o^{K-1}(1-q_1^N)^2(p_1+p_o)[(K-1)\ln q_o+1-q_o^K]$$

Thus setting the derivative of (3.22) wrt K equal to zero yields

$$p_o p_1\,\frac{q_o^{K-1}(1-q_1^N)^2(p_1+p_o)}{Y^2}\,[(K-1)\ln q_o+1-q_o^K]=0 \qquad (3.24)$$

where Y is the denominator of (3.22).

For $K \geq 1$, $N \geq 1$

$$p_o p_1 q_o^{K-1}(1-q_1^N)^2 \; (p_1+p_o) > 0$$

$$1-(K-1)q_o^{K-1}+(K-2)q_o^K = p_o^2 \sum_{J=1}^{K-1} Jq_o^{J-1}+p_o q_o^{K-1} > 0$$

and

$$1-(N-1)q_1^{N-1}+(N-2)q_1^N = p_1^2 \sum_{J=1}^{N-1} Jq_1^{J-1}+p_1 q_1^{N-1} > 0$$

implying Y and $Y^2 > 0$. Thus the left hand side of (3.24) is equal to zero only if

$$(K-1)(-\ln q_o) = 1-q_o^K \tag{3.25}$$

This is the same implicit equation as that of Section 2.5 and its graphical solution is shown in Figure 1. Also by the same argument as given in Section 2.5, the integer K most nearly satisfying (3.25) defines the global maximum of (3.22) with respect to K. Since (3.22) is symmetrical in K and N it is clear that (3.22) is maximized with respect to N by choosing N to be the integer most nearly satisfying (within ± 1)

$$(N-1)(-\ln q_1) = 1-q_1^N \tag{3.26}$$

Thus the globally optimum code is defined by the integers K and N most nearly satisfying (3.25) and (3.26) respectively. The solutions of (3.25) and (3.26) and the resulting compression ratios for various values of p_o and p_1 are given in Table 2.

Note that in Table 2 the compression ratio for $p_o = 0.001$ and $p_1 = 0.500$ is greater than that for $p_o = 0.001$ and $p_1 = 0.005$ but lower than that for $p_o = p_1 = 0.001$. This seems strange since in the second case more strings of 1's should occur than in the first case and thus, perhaps, a greater overall compression ratio should be expected. This behavior can be intuitively explained by the fact that for $p_o \ll p_1$ the state probability of a zero is nearly one as shown below.

$$p(0) = \frac{p_1}{p_1 + p_o} \approx \frac{p_1}{p_1} \approx 1$$

Thus the source is almost always in the state 0 where high compression ratios are obtained. As p_1 approaches p_o the source is less likely to be in state zero and the overall compression ratio decreases even though the compression ratio obtained in state 1 is increasing. Finally, as the compression ratio in state 1 increases further the overall compression ratio increases again.

3.6. Optimal Code When Huffman Coding is Used to Code Output Symbols.

The probabilities of $x_i (i = 1,2,3,4)$ may be defined as the limit of their frequency ratio as the length of the input sequence tends to infinity. Thus

$$P(x_i) = \lim_{n\to\infty} \frac{n(x_i)}{n(x_1)+n(x_2)+n(x_3)+n(x_4)} \tag{3.27}$$

Table 3

Optimum K,N and CR when output symbols are block coded (L=2)

p_1 \ p_o	0.500	0.100	0.050	0.010	0.005	0.001
0.500	(1,1) 1.000	(5,2) 1.218	(7,2) 1.639	(14,2) 3.537	(20,2) 4.989	(45,2) 11.152
0.100	(2,5) 1.218	(5,5) 1.392	(7,5) 1.672	(14,5) 3.298	(20,5) 4.666	(45,5) 10.703
0.050	(2,7) 1.639	(5,7) 1.672	(7,7) 1.859	(14,7) 3.249	(20,7) 4.528	(45,7) 10.415
0.010	(2,14) 3.537	(5,14) 3.298	(7,14) 3.249	(14,14) 3.821	(20,14) 4.688	(45,14) 9.704
0.005	(2,20) 4.989	(5,20) 4.666	(7,20) 4.528	(14,20) 4.688	(20,20) 5.288	(45,20) 9.600
0.001	(2,45) 11.152	(5,45) 10.703	(7,45) 10.415	(14,45) 9.704	(20,45) 9.600	(45,45) 11.471

Key: (K,N) / CR

where $n(x_i)$ $(i = 1,2,3,4)$ is the number of x_i's in the output sequence and n is the number of binary digits in the input sequence. In terms of register actions (3.27) may be written

$$P(x_i) = \lim_{Q\to\infty} \frac{\sum_{z=A,B,x_3,x_4} \left\{ n(z) \sum_{j=1}^{Q} n_j(x_i|z) \right\}}{\sum_{z=A,B,x_3,x_4} \left\{ n(z) \sum_{j=1}^{Q} \left[n_j(x_1|z)+n_j(x_2|z)+n_j(x_3|z)+n_j(x_4|z) \right] \right\}} \quad (3.28)$$

$(i = 1,2,3,4)$

where the notation is the same as that of Section 3.4. Dividing numerator and denominator of (3.28) by Q^2 and using (3.6) and (3.7) yields

$$P(x_i) = \frac{E(x_i)}{E(x_1) + E(x_2) + E(x_3) + E(x_4)} \qquad (i = 1,2,3,4)$$

A finite computer search may now be performed to determine the optimum K and N for the Huffman coded output symbols using the same arguments as those given in Section 2.6. In this case the search would fix N, search K from 1 to N, increment N, search K from 1 to N, etc.

3.7. Reduction to the Independent Source.

If p_o and p_1 of the binary first order Markov source are chosen to be p and q respectively, the Markov source is equivalent to an independent binary source with probabilities p and q for a 1 and 0 respectively. Thus Chapter II is really a special case of

Chapter III. If p_o and q_o are chosen as above and $K = 1$ the results of Chapter III reduce to those of Chapter II.

CHAPTER IV

RUN LENGTH CODING USING TWO STANDARD RUN LENGTHS FOR THE INDEPENDENT BINARY SOURCE

4.1. Introduction.

In this chapter a closed form expression is derived for the compression ratio when a binary independent source is encoded using two standard run lengths. The coder is assumed to have a memory of N binary digits where N is the length of the longest standard run length. A computer search is then performed to select the best run lengths. It is strongly suspected that the results of the computer search are global optimums although this has not been proved.

The above must be considered a coding algorithm constrained by the fact that the coder has a memory of only N binary digits. If memory is unconstrained the problem is much more difficult and a simple coding algorithm is not possible. This may be illustrated with a simple example. Suppose it is desired to code a string of 19 0's using the following equal cost symbols.

$$0 \rightarrow x_1$$

$$1 \rightarrow x_2$$

$$6 \text{ 0's in a row} \rightarrow x_3$$

$$7 \text{ 0's in a row} \rightarrow x_4$$

Using the algorithm discussed above this string would be coded as

$$2\ x_4\text{'s and } 5\ x_1\text{'s} = 7 \text{ code symbols}$$

whereas the optimum coder would code the sequence as

$$1\ x_4 \text{ and } 2\ x_3\text{'s} = 3 \text{ code symbols.}$$

Thus the technique described in this chapter always codes a string of zeros by using the maximum number of x_4's, then the maximum number of x_3's followed by x_1's.

4.2. Definition of Coding Technique.

An independent binary source emitting zeros and ones with probabilities q and $p = 1-q$ respectively is encoded as follows.

$$0 \rightarrow x_1$$

$$1 \rightarrow x_2$$

$$K\ 0\text{'s in a row} \rightarrow x_3$$

$$N\ 0\text{'s in a row} \rightarrow x_4$$

The remaining 0's are then coded as x_1's. Note that K and N are distinct from those in Chapter III.

The operation of the coder is defined as follows. No action is taken until the occurrence of one of the following two events:

A. a 1 is reached in the input sequence, or

B. N 0's have been accumulated.

If event A occurs the source encodes the J0's $(0 \leq J \leq N-1)$ and 1 accumulated as $[\frac{J}{K}]$ x_3's, $(J-K[\frac{J}{K}])$ x_1's, and an x_2 where [ ] is defined as the integer part of the expression enclosed. If event B occurs the coder simply outputs an x_4. Thus as in preceding chapters the coder operation may be viewed as a mapping of certain input sequences into their corresponding output sequences as shown below. The mapping of one of these input sequences into the corresponding output sequence is again denoted as a coder action (CA).

$$\begin{array}{ll}
1 \rightarrow x_2 & \\
01 \rightarrow x_1 x_2 & \\
\vdots & \\
\underbrace{0\cdots0}_{K-1}1 \rightarrow \underbrace{x_1\cdots x_1}_{K-1}\, x_2 & \\
\underbrace{0\cdots00}_{K}\,1 \rightarrow x_3 x_2 & \underbrace{0\ldots00}_{N} \rightarrow x_4 \\
\vdots & \\
\underbrace{0\cdots0}_{N-1}\,1 \rightarrow \underbrace{x_3\cdots x_3}_{[\frac{N-1}{K}]}\, \underbrace{x_1\cdots x_1}_{N-1-K[\frac{N-1}{K}]}\, x_2 &
\end{array} \tag{4.1}$$

4.3. Definition of Compression Ratio.

The compression ratio is defined to be the expected ratio of the number of binary digits in the input sequence to the number of binary digits in the output sequence as the length of the input sequence tends to infinity.

$$CR = \lim_{n\to\infty} \frac{n}{n(x_1)\ell_1+n(x_2)\ell_2+n(x_3)\ell_3+n(x_4)\ell_4} \tag{4.2}$$

where

n = number of input symbols

$n(x_i)$ = number of x_i's $(i = 1,2,3,4)$ in the output sequence

ℓ_i = cost of the code word for x_i $(i = 1,2,3,4)$ in binary digits.

The optimum code is again defined by the K and N that maximize the compression ratio (4.2).

4.4. Derivation of Compression Ratio in Terms of Coder Actions.

Let $[\frac{N-1}{K}]=M$. From (4.1) it is evident that the probability that a coder action results in a string of J x_1's $(0 \le J \le K-1)$ followed by an x_2 is given by

$$P_{CA}(Jx_1\text{'s},x_2) = P(J0\text{'s},1)+P(K+J0\text{'s},1)+ \dots + P((M-1)K+J0\text{'s},1)$$
$$+ P(MK+J0\text{'s},1) = pq^J+pq^{K+J}+\dots+pq^{(M-1)K+J}+pq^{MK+J}$$

if $J \le N-1-MK$. If $N-1-MK < J \le K-1$

$$P_{CA}(Jx_1\text{'s},x_2) = P(J0\text{'s},1)+P(K+J0\text{'s},1)+ \dots + P((M-1)K+J0\text{'s},1)$$
$$= pq^J+pq^{K+J}+ \dots +pq^{(M-1)K+J}$$

Similarly

$$P_{CA}(x_2) = P(J0\text{'s},1) \qquad 0 \le J \le N-1$$

$$P_{CA}(Jx_3\text{'s}) = P(JK+L0\text{'s}) \qquad 0 \le L \le K-1 \;,\; 0 \le J \le M-1$$

$$P_{CA}(x_4) = q^N$$

Thus the expected number of x_i's $(i = 1,2,3,4)$ emitted per coder action are given by

$$E(x_1) = p \sum_{J-1}^{K-1} Jq^J + \left[p \sum_{J=1}^{K-1} Jq^J \right] \left[\sum_{J=1}^{M-1} q^{JK} \right] + pq^{MK} \sum_{J=1}^{N-MK-1} Jq^J$$

$$E(x_2) = p \sum_{J=0}^{N-1} q^J$$

$$E(x_3) = \left[\sum_{J=1}^{M-1} Jq^{JK} \right] \left[p \sum_{J=0}^{K-1} q^J \right] + Mpq^{MK} \sum_{J=0}^{N-MK-1} q^J$$

$$E(x_4) = q^N$$

Performing the indicated summations and reducing yields

$$E(x_1) = \frac{q[1-Kq^{K-1}+(K-1)q^K](1-q^{MK})}{p(1-q^K)} + \frac{q^{MK+1}-(N-MK)q^N+(N-MK-1)q^{N+1}}{p} \qquad (4.3)$$

$$E(x_2) = 1-q^N$$

$$E(x_3) = \frac{q^K[1-Mq^{(M-1)K}+(M-1)q^{MK}]}{(1-q^K)} + M(q^{MK}-q^N)$$

$$E(x_4) = q^N$$

By the same arguments as given in Section 2.4 it can be shown that the compression ratio (4.2) converges to

$$\frac{E(x_1) + E(x_2) + KE(x_3) + NE(x_4)}{\ell_1 E(x_1) + \ell_2 E(x_2) + \ell_3 E(x_3) + \ell_4 E(x_4)}$$

with probability one. Thus in summary

$$CR = \frac{E(x_1) + E(x_2) + KE(x_3) + NE(x_4)}{\ell_1 E(x_1) + \ell_2 E(x_2) + \ell_3 E(x_3) + \ell_4 E(x_4)} \tag{4.4}$$

where $E(x_i)$ $(i = 1,2,3,4)$ is given in (4.3).

Using the same arguments as given in Section 2.6 it can also be shown that the output symbol probabilities converge to

$$P(x_i) = \frac{E(x_i)}{E(x_1) + E(x_2) + E(x_3) + E(x_4)} \qquad (i = 1,2,3,4)$$

with probability one.

4.5. Optimal Coding.

The integers K and N maximizing (4.4) may now be found by computer search. This has been done for both the case of equal length output symbols $(\ell_1=\ell_2=\ell_3=\ell_4=2)$ and when the output symbols were Huffman coded. The search was carried out well beyond the point where (4.4) appeared to be maximized. It is strongly suspected that the results of the computer search are global optimums although this has not been proved. Results of the computer search are given in Table 4 and Figure 6, 7 and 8. A comparison of the efficiencies of the coding techniques presented in Chapters II and IV with various other coding schemes is given in Figure 9. The results of Figure 9 are for Huffman coding of the output symbols in each case.

TABLE 4

Compression ratio and run lengths vs p

p	CR_B	CR_H	K_B	N_B	K_H	N_H
0.200	1.014	1.102	2	5	1	5
0.100	1.512	1.574	3	8	6	14
0.050	2.286	2.356	4	14	8	20
0.030	3.130	3.183	4	18	8	29
0.020	4.033	4.155	5	23	9	41
0.010	6.235	6.536	6	39	11	61
0.005	9.719	10.287	8	60	14	92
0.003	13.512	14.402	9	77	16	136

Key

p = probability of a 1

CR_B = compression ratio when block coding is used on output symbols

CR_H = compression ratio when Huffman coding is used on output symbols

K_B, N_B = standard run lengths associated with CR_B

K_H, N_H = standard run lengths associated with CR_H

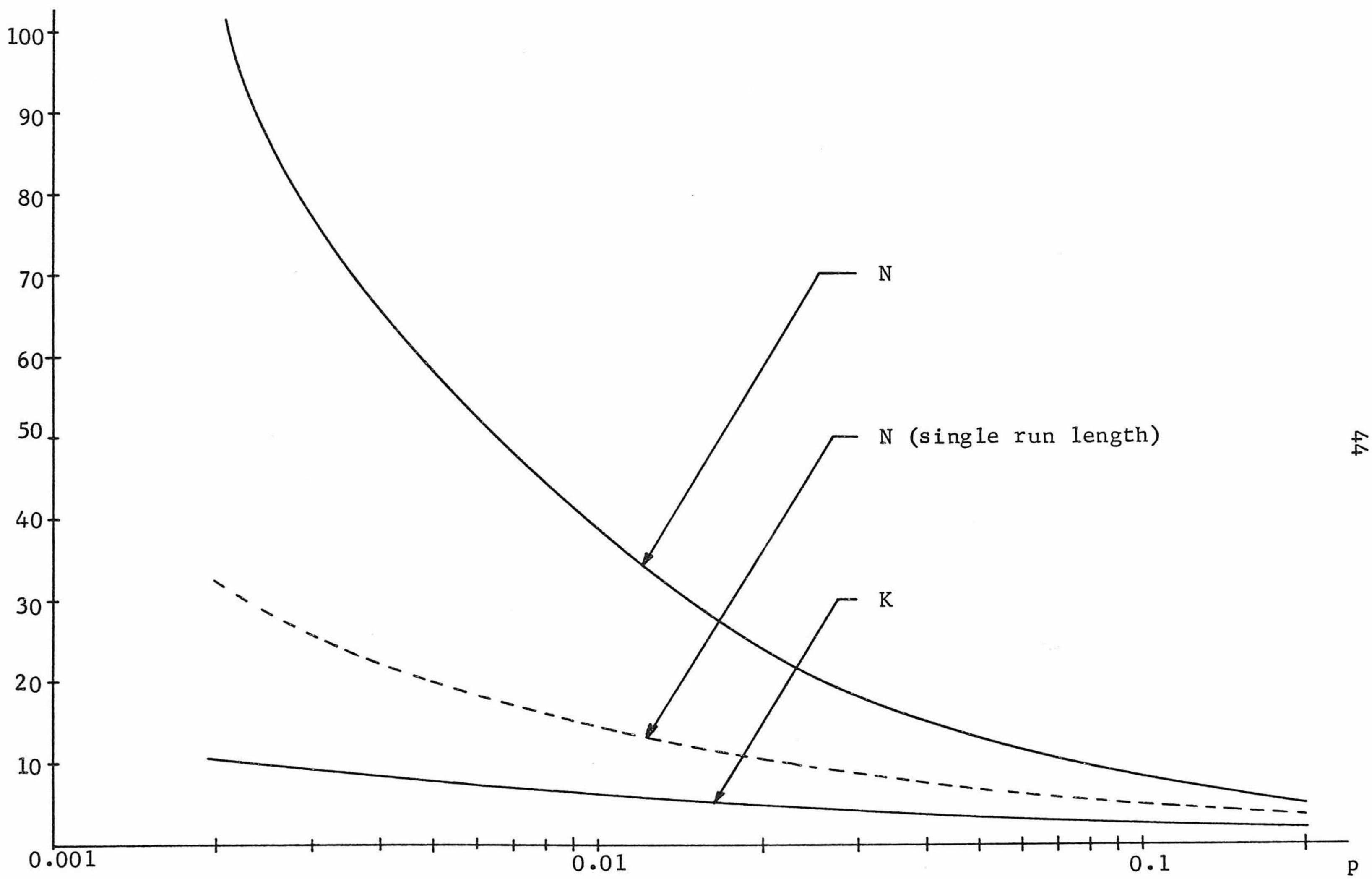


Figure 6

K and N vs p for two standard run lengths using block coding

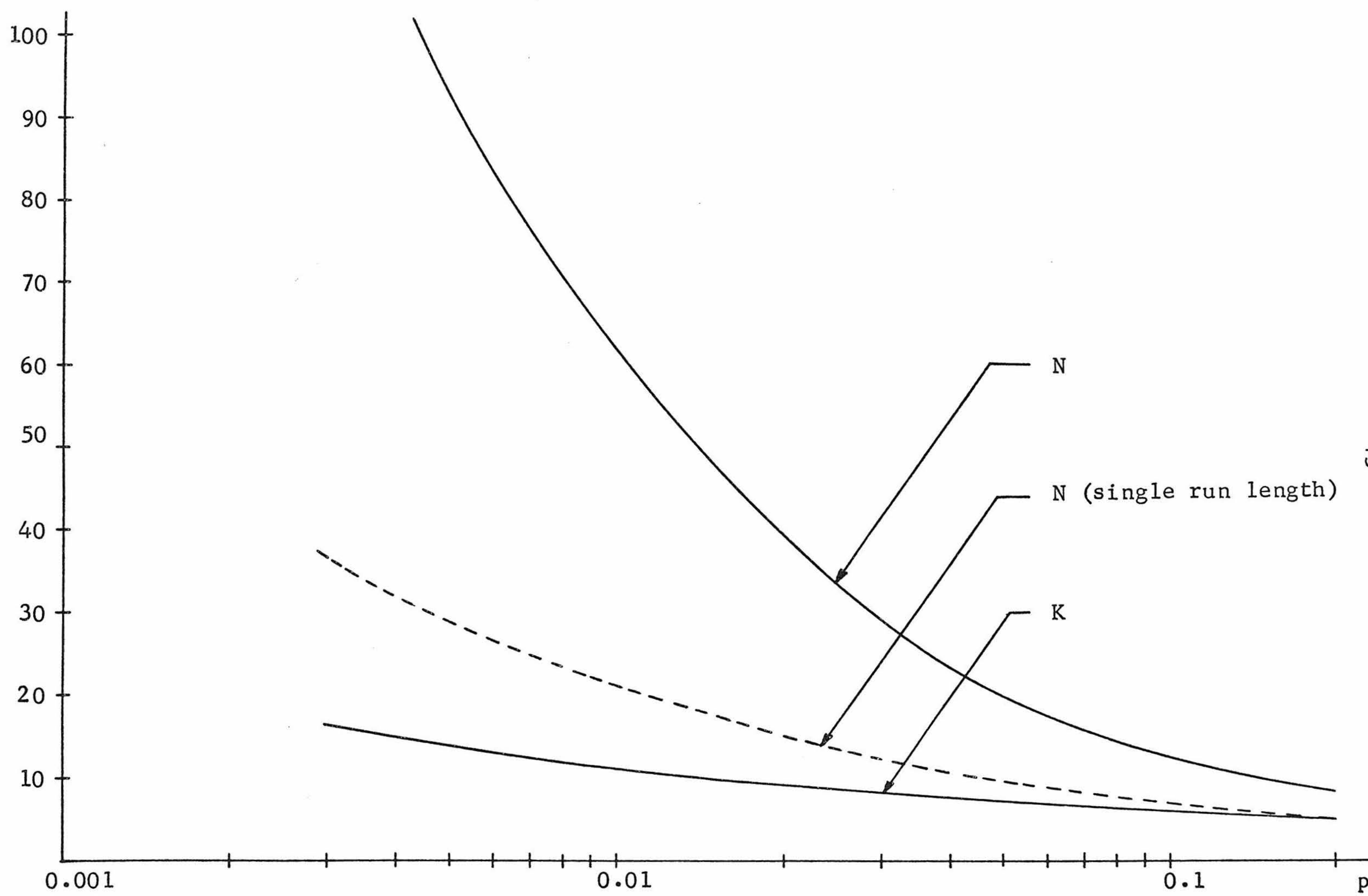


Figure 7

K and N vs p for two standard run lengths using Huffman coding

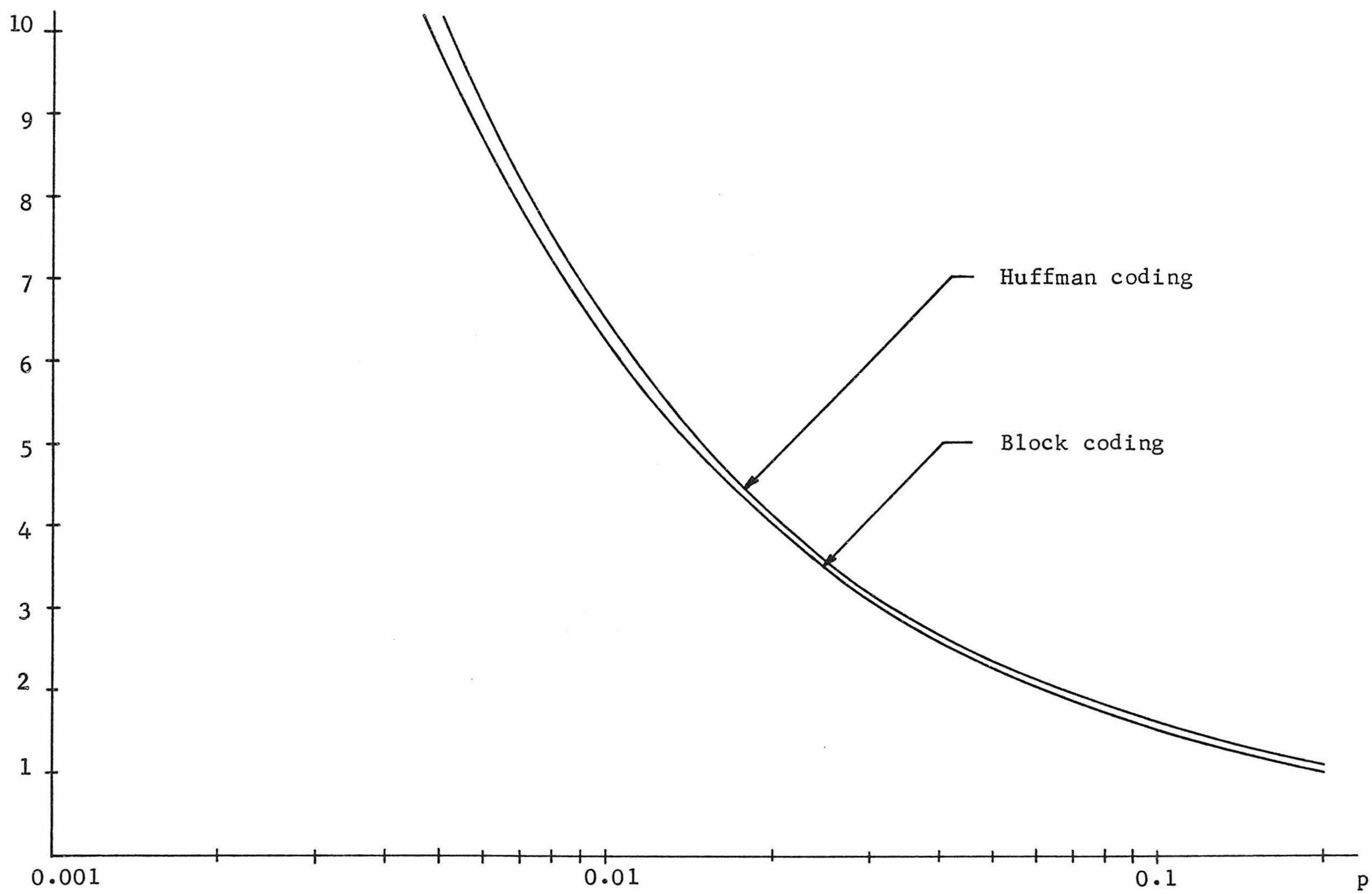


Figure 8

Compression ratio vs p for two standard run lengths

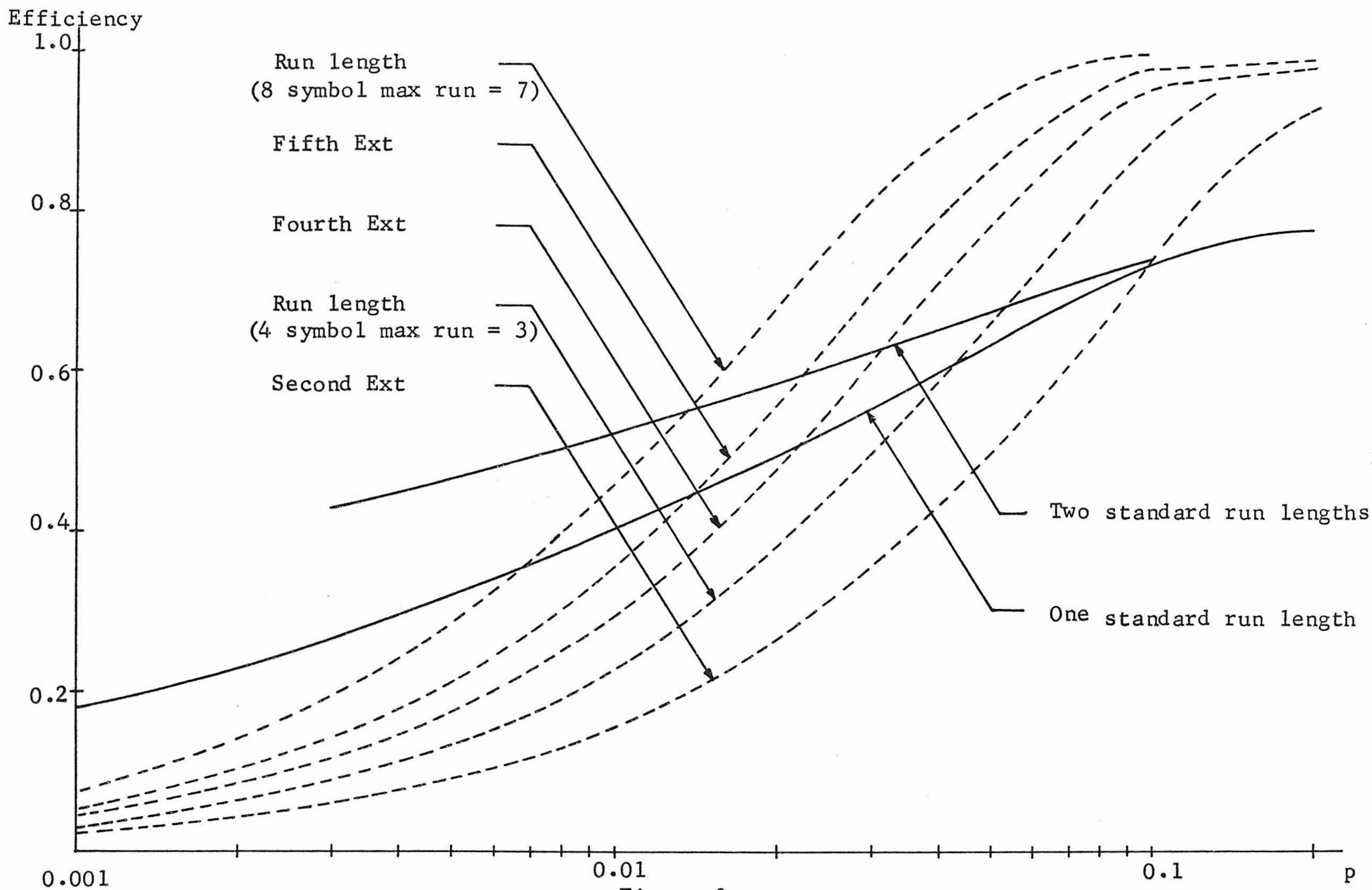


Figure 9

Comparison of coding schemes

CHAPTER V

CODING USING AN ARBITRARY NUMBER OF STANDARD RUN LENGTHS

5.1. Introduction.

The optimum coding scheme using single run lengths of 0's and 1's was derived in Chapter III. In Chapter IV closed form expressions for the output symbol probabilities and compression ratio of a coding algorithm using two standard run lengths of 0's with an independent binary source were derived. The coder was constrained to have a memory of N binary digits where N is the length of the longest standard run length. A computer search was then used to determine the optimum run lengths over the region searched. It would be desirable to generalize the results of Chapter III to an arbitrary number of run lengths. This is a difficult problem since the compression ratio must be simultaneously maximized over all the standard run lengths. Even if it is assumed that a run is encoded using the maximum number of the longest standard run lengths followed by the maximum number of the next longest run lengths, etc. (so that the compression ratio can at least be written in closed form), the expressions for the compression ratios involve integer parts of the ratios of the various run lengths which cannot be easily handled analytically. In this chapter a recursive coding technique is developed which generalizes to any number of run lengths and applies to both the binary independent and first order Markov sources. This technique assumes that the output symbols are block coded and that the ratios of standard run lengths are integers.

5.2. Coding Technique.

The coding algorithm is defined to code an input run by using the maximum number of the longest standard run lengths followed by the maximum number of the next longest standard run length, etc. This algorithm may be performed in two stages as shown in Figure 10. Note that the coder actions for both coders are the same. That is, coder No. 2 can act immediately on any coder action from coder No. 1. The derivation will be carried out for the binary independent source. That the results also apply to the binary first order Markov source is shown in Section 5.3.

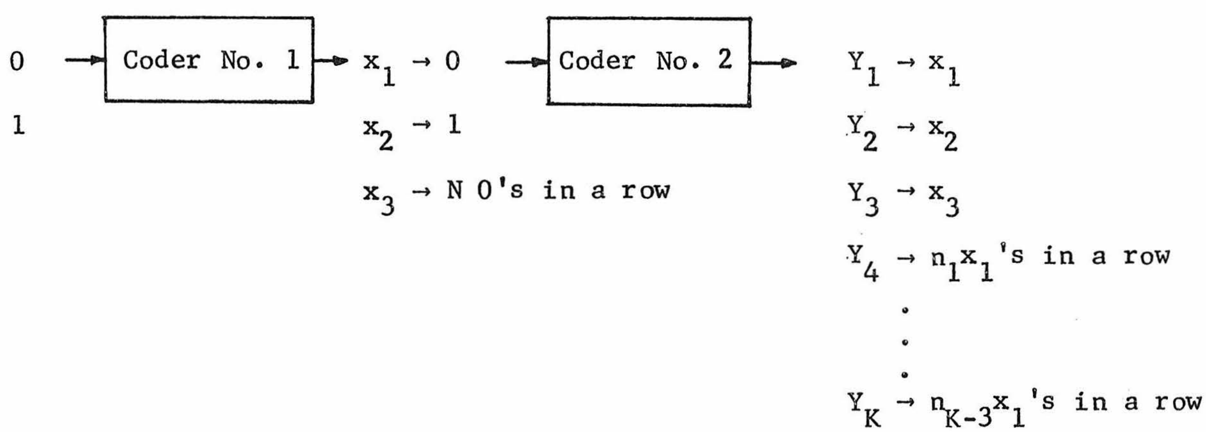


Figure 10
Coding Technique

The overall compression ratio may be written as

$$CR = \frac{1}{L}\left[\frac{E(x_1)+E(x_2)+NE(x_3)}{(E(Y_1)+E(Y_2)+E(Y_3)+E(Y_4)+ \ldots + E(Y_K))}\right] \tag{5.1}$$

where L is the length of the output block code using the same arguments as presented in Chapter II and IV. But

$$E(x_1) = E(Y_1) + n_1E(Y_4) + \dots + n_{K-3}E(Y_K)$$

$$E(x_2) = E(Y_2)$$

$$E(x_3) = E(Y_3)$$

Thus (5.1) may be written as

$$CR = \frac{1}{L}\left[\frac{E(x_1)+E(x_2)+NE(x_3)}{E(x_1)+E(x_2)+E(x_3)-(n_1-1)E(Y_4)- \dots - (n_{K-3}-1)E(Y_K)}\right] \tag{5.2}$$

Now $E(x_i)$ $(i = 1,2,3)$ are functions only of N and the probabilities of a 0 and 1. To determine the optimum code the bracketed quantity of (5.2) must be maximized over $N, n_1, n_2, \dots, n_K$. Assuming $\frac{N}{n_1}, \frac{n_1}{n_2}, \dots, \frac{n_{K-1}}{n_K}$ are integers $E(Y_4), \dots, E(Y_K)$ may be determined as follows. The probability of I Y_4's per coder action $(I = 1, \dots, \frac{N}{n_1} -1)$ is given by

$$P_{CA}(IY_4\text{'s}) = \sum_{S=0}^{n_1-1} P_{CA}(In_1+S \text{ 0's}, 1) \tag{5.3}$$

The probability of I Y_5's per coder action $(I = 1, \dots, \frac{n_1}{n_2} - 1)$ is given by

$$P_{CA}(IY_5\text{'s}) = \sum_{S=0}^{n_2-1} [P_{CA}(In_2+S \text{ 0's},1) + P_{CA}(n_1+In_2+S \text{ 0's},1)$$
$$+ \dots + P_{CA}((M-1)n_1+In_2+S \text{ 0's},1)]$$

where $M = \frac{N}{n_1}$.

Using the results of Chapter II this may be written

$$P_{CA}(IY_5's) = \sum_{S=0}^{n_2-1}\left\{\sum_{F=0}^{M-1} pq^{Fn_1+In_2+S}\right\}$$

from which $E(Y_5)$ may be written as

$$E(Y_5) = \left\{\sum_{I=1}^{\frac{n_1}{n_2}-1} Iq^{In_2}\right\}\left\{\sum_{S=0}^{n_2-1} pq^S\right\}\left\{\sum_{F=0}^{M-1} q^{Fn_1}\right\}$$

Continuing, it can be seen that

$$E(Y_J) = \left\{\sum_{I=1}^{\frac{n_{J-4}}{n_{J-3}}-1} Iq^{In_{J-3}}\right\}\left\{\sum_{S=0}^{n_{J-3}-1} pq^S\right\} f(N,n_1,\cdots,n_{J-4}) \tag{5.4}$$

where $f(N,n_1,\cdots,n_{J-4})$ is a positive summation of q to the various allowable combinations of standard run lengths. Thus differentiating the bracketed quantity of (5.2) with respect to n_{K-3} yields

$$\frac{[E(x_1)+E(x_2)+NE(x_3)]}{D^2} f(N,n_1,\cdots,n_{K-4}) \frac{\partial}{\partial n_{K-3}}\left\{(n_{K-3}-1)\sum_{I=1}^{\frac{n_{K-4}}{n_{K-3}}-1} Iq^{In_{K-3}}\right\}$$

$$\times\left\{\sum_{S=0}^{n_{K-3}-1} pq^S\right\} = 0 \tag{5.5}$$

where D is the denominator of the bracketed quantity of (5.2). But

$$\frac{[E(x_1)+E(x_2)+NE(x_3)]}{D^2} f(N,n_1,\cdots,n_{K-4}) > 0$$

since $E(x_1) + E(x_2) + NE(x_3)$ is the expected number of input symbols per coder action $f(N, n_1, \cdots, n_{K-4})$ is a positive summation as pointed out above and D is equal to the expected number of output symbols per coder action. Thus (5.5) reduces to

$$\frac{\partial}{\partial n_{K-3}} \left\{ (n_{K-3}-1) \sum_{I=1}^{\frac{n_{K-4}}{n_{K-3}} - 1} Iq^{In_{K-3}} \right\} \left\{ \sum_{S=0}^{n_{K-3}-1} pq^S \right\} = 0 \tag{5.6}$$

Now assuming that n_{K-3} is known (5.6) gives a relationship from which n_{K-4} can be determined. Now the same procedure can be applied to n_{K-4} yielding (5.6) with n_{K-4} replacing n_{K-3} and n_{K-5} replacing n_{K-4}. Since n_{K-4} is known this yields a relationship from which n_{K-5} can be determined. Thus the solution of (5.6) gives a recursive relationship between each run length and the next longer run length. This may be determined as follows. Letting $n_{K-4} = N$ and $n_{K-3} = K$ in (5.6) for notational convenience yields

$$\frac{\partial}{\partial K} \left\{ (K-1) \left[\sum_{I=1}^{\frac{N}{K}-1} Iq^{IK} \right] \left[\sum_{S=0}^{K-1} pq^S \right] \right\} = 0$$

Performing the indicated summations yields

$$\frac{\partial}{\partial K} \left\{ \frac{(K-1)\ [q^K - \frac{N}{K} q^N + (\frac{N}{K} - 1) q^{N+K}]}{(1-q^K)} \right\} = 0$$

Performing the differentiation

$$\frac{1}{(1-q^K)^2}\left\{(1-q^K)\left[(K-1)[q^K \ln q + \frac{N}{K^2} q^N + (\frac{N}{K} - 1)q^{N+K} \ln q - \frac{N}{K^2} q^{N+K}] + [q^K - \frac{N}{K} q^N + (\frac{N}{K} - 1) q^{N+K}]\right] + (K-1)[q^K - \frac{N}{K} q^N + (\frac{N}{K} - 1)q^{N+K}] q^K \ln q\right\} = 0$$

Multiplying out expressions and reducing yields

$$\frac{1}{(1-q^K)^2}\left\{(K-1)q^K \ln q(1-q^N) + (1-q^K)[q^K(1-q^N) - \frac{N}{K^2} q^N(1-q^K)\right\} = 0 \qquad (5.7)$$

Since $\frac{1}{(1-q^K)^2} > 0$ (5.7) is equal to zero only if

$$(K-1)(1-q^N)(-\ln q) = \frac{(1-q^K)}{q^K}\left[q^K(1-q^N) - \frac{N}{K^2} q^N(1-q^K)\right]$$

or rearranging

$$(K-1)(-\ln q) = (1-q^K)\left[1 - \frac{Nq^N(1-q^K)}{K^2 q^K(1-q^N)}\right] \qquad (5.8)$$

Now if

$$(K-1)(-\ln q) < (1-q^K) \qquad (5.9)$$

(5.8) will have at least one solution for N as a function of K. Comparing Figure 6 and Figure 2 (5.9) is satisfied at least for the case of two standard run lengths over the range where calculations were made. It is suspected that this is the case in general although this has not been proved.

The optimal code for this algorithm may now be searched out as follows. Start with $n_{K-3} = 2$ and use (5.8) to determine the remaining standard run lengths. Calculate the compression ratio. Increment K and repeat. Select the run length set that maximizes the compression ratio. Global optimality of the search results is not guaranteed although Figure 6 indicates that over a wide range a search over low values of n_{K-3} is probably sufficient. The compression ratio vs. number of run lengths for P=0.005 is given in Table 5.

5.3. Generalization to First Order Markov Source.

The coding technique applied to the first order Markov source is shown in Figure 11.

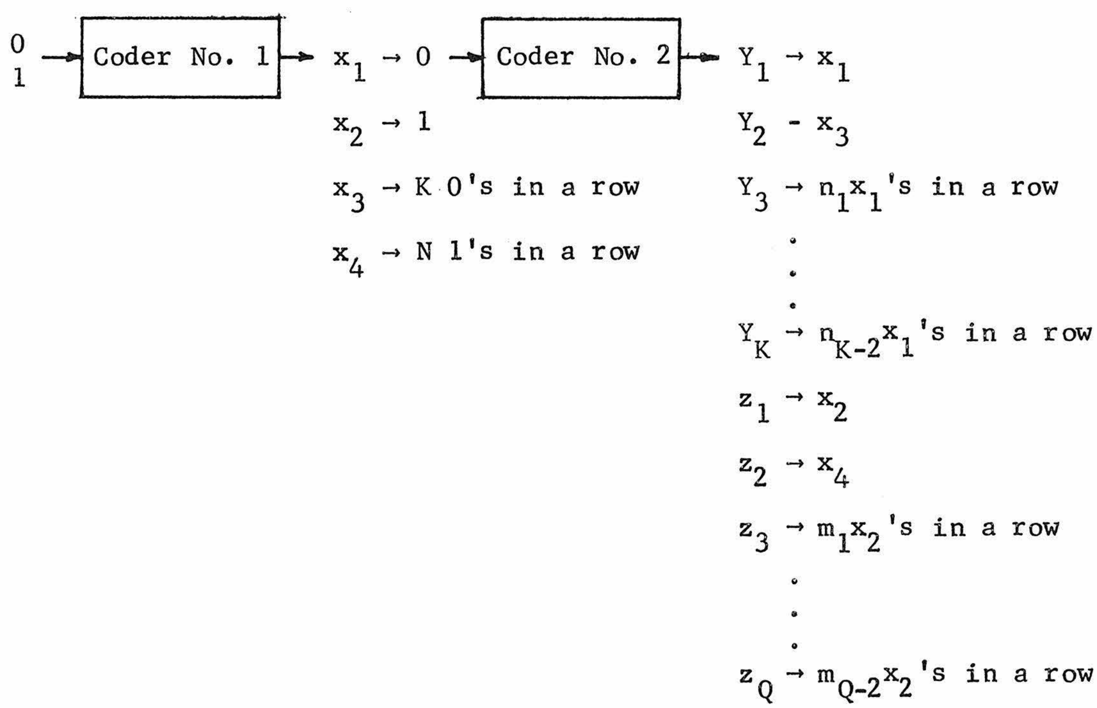


Figure 11

Coding Technique for First Order Markov Source

TABLE 5

Compression ratio vs number of standard run lengths for $p = 0.005$

N_L	N	K	M	CR_S	CR_B
1	20			10.092	5.046
2	64	8		19.217	9.719
3	81	9	3	23.032	7.677*

Key

N_L = number of standard run lengths

N,K,M = lengths of the standard run lengths

CR_S = compression ratio assuming output symbols of unit cost

CR_B = compression ratio when output symbols are block coded

* Note that the compression ratio when the output symbols are block coded is less for three standard runs than for two standard runs. This is because the required length of the output symbol block code increases faster than the compression ratio. The compression ratio assuming unit cost for output symbols ($\ell_1=\ell_2=\ell_3=\ell_4=1$) of course increases. This corresponds to a mapping of the binary source into a five-level source.

From Chapter III, the overall compression ratio may be written as

$$CR = \frac{1}{L}\left[\frac{E(x_1)+E(x_2)+KE(x_3)+NE(x_4)}{E(Y_1)+E(Y_2)+\ldots+E(Y_K)+E(z_1)+E(z_2)+\ldots+E(z_Q)}\right] \tag{5.10}$$

where L is the length of the output block code but

$$E(x_1) = E(Y_1) + n_1E(Y_3)+ \ldots + n_{K-2}E(Y_K)$$

$$E(x_3) = E(Y_2)$$

$$E(x_2) = E(z_1) + m_1E(z_3) + \ldots + m_{Q-2}E(z_Q)$$

$$E(x_4) = E(z_2)$$

Thus (5.10) may be written as

$$CR = \frac{1}{L}\left[\frac{E(x_1)+E(x_2)+KE(x_3)+NE(x_4)}{E(x_1)+E(x_2)+E(x_3)+E(x_4)-(n_1-1)E(Y_3)- \ldots - (n_{K-2}-1)E(Y_K) -(m_1-1)E(z_3)- \ldots - (m_{Q-2}-1)E(z_Q)}\right]$$

From Chapter III

$$P_{CA}(IY_4\text{'s}) = \frac{1}{q_o}\sum_{S=0}^{n_2-1}\left\{\sum_{F=0}^{M-1} p_o q_o^{Fn_1+In_2+S}\frac{(1-q_1N)}{(1-q_o^K)+(1-q_1^N)}\right\}$$

where $M = \frac{N}{n_1}$ and thus

$$E(Y_4) = \frac{1}{q_o}\left\{\sum_{I=1}^{\frac{n_1}{n_2}-1} Iq_o^{In_2}\right\}\left\{\sum_{S=0}^{n_2-1} p_o q_o^S\right\}\left\{\sum_{F=0}^{M-1} q^{Fn_1} \frac{(1-q_1^N)}{(1-q_o^K)+(1-q_1^N)}\right\}$$

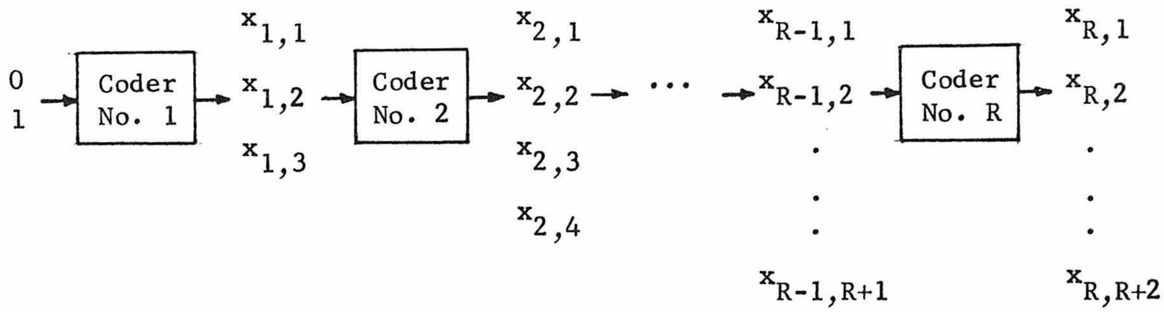
Continuing

$$E(Y_J) = \frac{1}{q_o}\left\{\sum_{I=1}^{\frac{n_{J-3}}{n_{J-2}}} Iq_o^{In_{J-2}}\right\}\left\{\sum_{S=0}^{n_{J-2}-1} p_o q_o^S\right\} f(N,K,n_1,\ldots,n_{J-3})$$

But this is the same as (5.4) with $q = q_o$ and $p = p_o$ except for the constant $\frac{1}{q_o}$ and $f(N,K,n_1,\ldots,n_{J-S})$. Since both of these factors are constants with respect to the differentiations the same recursion formula (5.8) results for $N, n_1, \ldots, n_{K-2}$ with $p = p_o$ and $q = q_o$. An identical argument on $E(z_J)$ shows that the same recursion formula (5.8) holds for $K, m_1, \ldots, m_{Q-z}$ with $p = p_1$ and $q = q_1$.

5.4. Recursive Coding Technique.

Consider the coding technique shown in Figure 12. Again it is assumed that a run is encoded using the maximum number of the longest standard run length followed by the maximum number of the next longest run length, etc. Also it is assumed that the ratios of the standard run lengths are integers.



where the coding sequence is defined as follows.

$$0 \rightarrow x_{1,1} \rightarrow x_{2,1} \rightarrow \cdots \rightarrow x_{R-1,1} \rightarrow x_{R,1}$$

$$1 \rightarrow x_{1,2} \rightarrow x_{2,2} \rightarrow \cdots \rightarrow x_{R-1,2} \rightarrow x_{R,2}$$

$$n_1 \text{ 0's in a row} \rightarrow x_{1,3} \rightarrow x_{2,3} \rightarrow \cdots \rightarrow x_{R-1,3} \rightarrow x_{R,3}$$

$$n_2 \text{ 0's in a row (or } n_2 x_1 \text{'s in a row)} \rightarrow x_{2,4} \rightarrow \cdots \rightarrow x_{R-1,4} \rightarrow x_{R,4}$$

$$\vdots$$

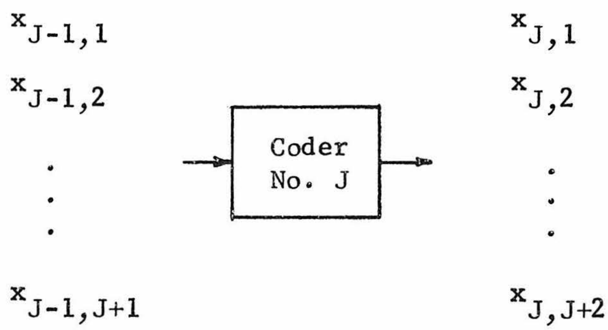
$$n_R \text{ 0's in a row (or } n_R x_{R-1,1} \text{'s in a row)} \rightarrow x_{R,R+2}$$

Figure 12

Recursive coding technique

This time the standard run lengths are selected recursively to maximize the symbol compression ratio of each coder. The symbol compression ratio is defined as the expected ratio of input to output symbols as the length of the input sequence tends to infinity. Thus n_1 is selected to maximize the symbol compression ratio of coder No. 1, n_2 is then selected to maximize the symbol compression ratio of coder No. 2, etc. Note that the coder actions for all the coders are the same. That is, Coder No. 2 can act immediately on any coder action from coder No. 1, etc. The optimal way to select n_1 was derived in Chapters II and III. A recursive technique to optimally select $n_2, n_3, \ldots, n_K$ will now be derived.

Consider coder No. J as shown in Figure 13.



where the coding sequence is defined as follows.

$$x_{J-1,1} \rightarrow x_{J,1}$$
$$x_{J-1,2} \rightarrow x_{J,2}$$
$$\vdots$$
$$x_{J-1,J+1} \rightarrow x_{J,J+1}$$
$$n_J x_{J-1,1}\text{'s in a row} \rightarrow x_{J,J+2}$$

Figure 13

Coder J

The symbol compression ratio for each coder is defined as the expected ratio of the number of symbols in the input sequence to the number of symbols in the output sequence as the length of the input sequence tends to infinity. Using the same reasoning as that given in Sections 2.4 and 3.4 the symbol compression ratio for the Jth coder converges with probability one to

$$CR_J = \frac{E(x_{J,1})+E(x_{J,2})+\ldots+E(x_{J,J+1})+n_J E(x_{J,J+2})}{E(x_{J,1})+E(x_{J,2})+\ldots+E(x_{J,J+1})+E(x_{J,J+2})} \tag{5.11}$$

where $E(x_{ij})$ denotes the expected number of x_{ij}'s emitted per coded action. But for each coder action

$$\begin{aligned} E(x_{J,1}) + n_J E(x_{J,J+2}) &= E(x_{J-1,1}) \\ E(x_{J,2}) &= E(x_{J-1,2}) \\ &\vdots \\ E(x_{J,J+1}) &= E(x_{J-1,J+1}) \end{aligned} \tag{5.12}$$

Thus (5.11) may be written as

$$CR_J = \frac{E(x_{J-1,1})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})}{E(x_{J-1,1})-n_J E(x_{J,J+2})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})+E(x_{J,J+2})}$$

Combining terms

$$CR_J = \frac{[E(x_{J-1,1})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})]}{[E(x_{J-1,1})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})]-(n_J-1)E(x_{J,J+2})} \tag{5.13}$$

Since the quantity

$$E(x_{J-1,1})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})$$

does not depend on n_J, differentiating (5.3) with respect to n_J and setting the result equal to zero yields

$$\frac{E(x_{J-1,1})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})}{Y^2}\left\{\frac{\partial}{\partial n_J}\left[(n_J-1)E(x_{J,J+2})\right]\right\} = 0$$

But

$$\frac{E(x_{J-1,1})+E(x_{J-1,2})+\ldots+E(x_{J-1,J+1})}{Y^2} > 0$$

since the numerator is the expected number of input symbols per coder action and the denominator is equal to the square of the number of output symbols per coder action. Thus to find the maximum of (5.1) it is necessary only to solve

$$\frac{\partial}{\partial n_J}\left[(n_J-1)E(x_{J,J+2})\right] = 0$$

But

$$E(x_{J,J+2}) = E(Y_J)$$

of the previous section except for subscript notation differences. Thus the implicit equation (5.8) results except this time the longer standard run length is fixed and the next shorter one is to be determined. This is just the reverse of the previous section. By the same arguments used previously this can be generalized to the first order Markov case with $p = p_o$ and $q = q_o$ in the case of run lengths of 0's and $p = p_1$ and $q = q_1$ for run lengths of 1's. The compression ratio vs. number of standard run lengths for $p = 0.005$ is given in Table 6. Instead of choosing N by the method of Chapter II and III some improvement may be gained by incrementing N constraining $n_1, \ldots, n_K$ to satisfy (5.8) and searching out the maximum compression ratio.

A test of (5.8) is to try to calculate K of Figure (6) given N and p. This has been done and interestingly enough (5.8) predicted the correct K exactly for every point checked even though in some cases $\frac{N}{K}$ was not an integer as was assumed in the derivation.

TABLE 6

Compression ratio vs number of standard run lengths for p = 0.005 (recursive scheme)

N_L	N	K	M	CR_S	CR_B
1	20			10.092	5.046
2	20	4		14.350	7.175
3	20	4	2	14.882	4.960

Key

p = probability of a 1

CR_B = compression ratio when block coding is used on output symbols

CR_S = compression ratio assuming output symbols of unit cost

N,K,M = lengths of the standard run lengths

CHAPTER VI

A SIMPLE SINGLE STANDARD RUN LENGTH SCHEME USING A NON-BLOCK CODE ON THE OUTPUT SYMBOLS

6.1. Introduction.

In Chapter II the optimum single standard run length for the binary independent source was derived assuming the output symbols were block coded. A non-block output code (Huffman) required computer search to determine the optimum standard run length. In this chapter a simple coding scheme using a single standard run length and a non-block output code is analyzed.

6.2. Coding Technique.

Consider a binary independent source emitting ones and zeros with probabilities p and $q = 1-p$ respectively. This sequence is then encoded as follows. After each M binary digits have been emitted the coder sends

1 if M zeros have been emitted

0 followed by the original sequence otherwise.

Let the average number of output digits used to represent M source symbols be denoted by L. Then

$$L = q^M + (1-q^M)(M+1)$$

or rewriting

$$L = 1 + M(1-q^M)$$

The compression ratio is defined as

$$CR = \frac{M}{L} = \frac{M}{1+M(1-q^M)}$$

Maximizing by differentiating with respect to M and setting the result equal to zero yields

$$\frac{1}{[1+M(1-q^M)]^2}\left[1+M(1-q^M)-M(1-q^M)+M^2 q^M \ln q\right] = 0$$

or

$$q^M(-\ln q) = \frac{1}{M^2} \qquad (6.1)$$

The same type of reasoning as presented in Section 2.5 shows that (6.1) defines a global maximum. The optimum M vs. p and the resulting compression ratio is given in Table 7.

TABLE 7

Optimum M and compression ratio vs p for non-block scheme

p	M	CR
0.2000	3	1.218
0.1000	4	1.684
0.0500	5	2.346
0.0300	6	2.997
0.0200	8	3.646
0.0100	11	5.113
0.0050	15	7.189
0.0030	19	9.249
0.0020	23	11.302
0.0015	26	13.031
0.0010	32	15.934

CHAPTER VII

CONCLUSIONS

The globally optimum single standard run length has been derived for the binary independent source and globally optimum single standard run lengths of zeros and ones have been derived for the binary first order Markov source. It is assumed that the output symbols are subsequently block coded in each case. The optimum standard run lengths depend on whether block or Huffman coding is subsequently used to encode the symbols. If Huffman coding is used on the output symbols the optimum standard run lengths can be determined by a finite computer search. A recursion relationship between standard run lengths is derived for two specific coding algorithms. An area of future study would be to try to remove the restrictions of these coding algorithms. A simple single standard run length scheme using a non-block code on the output symbols has also been derived for the binary independent source.

An advantage of this scheme over the usual run length coding, coding extensions of the source, or picking more general variable length codes [9], is ease of implementation. From a theoretical point of view, for example, Huffman coding a sufficiently large extension of the source will guarantee an efficiency as close to one as desired. Implementing this scheme, however, requires that the coder be able to distinguish between 2^n source sequences of length n where n is the order of the extension. As the source becomes more and more unsymmetrical a high extension must be coded to maintain the same efficiency. In contrast, the schemes proposed here require the coder to recognize only runs of zeros or ones. This can be accomplished with shift registers, counters and simple gating circuitry.

Of course the decision of whether or not to use a particular coding scheme is dependent on the source statistics as well as the complexity of implementation. The schemes presented in this thesis are particularly suited to unsymmetrical binary independent sources or binary first order Markov sources with unsymmetrical transition probabilities. A comparison of the efficiency of various schemes as a function of the source statistics is given in Figure 9.

Finally a coding scheme must be chosen with reference to the type of channel over which the information will be sent. Transmission over any realistic channel produces the possibility of errors. Errors of little concern to one particular coding scheme may be disastrous to another. For example, although the scheme of Chapter VI produces good compression ratios, loss of sync by the decoder essentially requires starting over again. Of course there are other classes of codes which are used because of their immunity to certain types of errors. These usually require more rather than less data be sent.

Thus the choice of a particular coding scheme for data compression is dependent not only upon the compression ratio attainable. Other factors such as ease of implementation, source statistics, and the channel that is to be used for transmission also play a major role.

REFERENCES

[1] R. B. Ash, Information Theory, John Wiley and Sons, Inc., New York, 1965.

[2] C. Cherry, M. H. Kubba, D. E. Pearson, and M. P. Barton, "An Experimental Study of the Possible Bandwidth Compression of Visual Image Signals", Proc. IEEE, Vol. 51, pp. 1507-1517, November 1963.

[3] L. O. Davisson, "Data Compression Using Straight Line Interpolation", IEEE Trans. on Information Theory, Vol. IT-14, pp. 390-394, May 1968.

[4] L. Ehrman, "Analysis of Some Redundancy Removal Bandwidth Compression Techniques", Proc. IEEE, Vol. 55, pp. 278-287, March 1967.

[5] L. W. Gardenhire, "Redundancy Reduction -- the Key to Adaptive Telemetry", 1964 Proc. Nat'l. Telemetering Conf., pp. 1-16.

[6] C. M. Kortman, "A Practical Method of Data Compression", Proc. IEEE, Vol. 55, pp. 253-263, March 1967.

[7] A. H. Robinson and C. Cherry, "Results of a Prototype Television Bandwidth Compression Scheme", Proc. IEEE, Vol. 55, pp. 356-364, March 1967.

[8] C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, The University of Illinois Press, Urbana, 1949.

[9] B. P. Tunstall, "Synthesis of Noiseless Compression Codes", Research Report No. 67-7, School of Electrical Engineering, Georgia Institute of Technology, October 1967.

[10] D. R. Weber, "A Synopsis on Data Compression", 1965 Proc. Nat'l. Telemetering Conf., pp. 9-16.

[11] J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, Inc., New York, 1965.