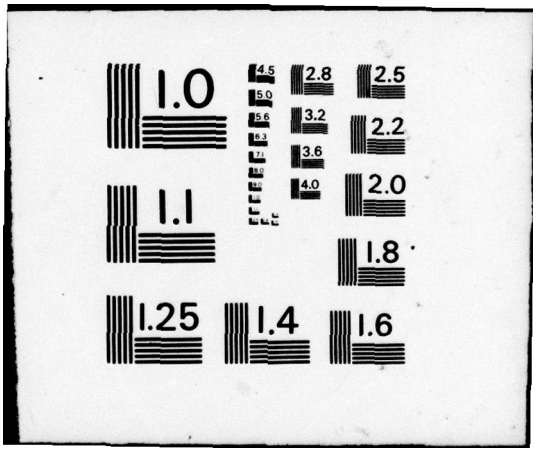
AD-A035 356 NORTHWESTERN UNIV EVANSTON ILL DEPT OF ELECTRICAL E--ETC F/G 9/2
ON UNIQUELY DECIPHERABLE CODES WITH GIVEN COMPOSITIONS,(U)
1976 S C NTAFOS, S L HAKIMI AF-AFOSR-3017-76

UNCLASSIFIED AFOSR-TR-77-0026 NL

1 OF 1
ADA035356

END
DATE FILMED
3-77


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6

ON UNIQUELY DECIPHERABLE CODES WITH GIVEN COMPOSITIONS[1]

S. C. NTAFOS
Department of Electrical Engineering
Northwestern University
Evanston, Illinois

S. L. HAKIMI
Departments of Electrical Engineering & Computer Sciences
Northwestern University
Evanston, Illinois

From the Proceedings of
1976 Allerton Conference
on Circuit and System
Theory
Sponsored by Univ. of Ill
Urbana, Ill.

ABSTRACT

It has been conjectured that a uniquely decipherable (U.D.) code can be replaced by a prefix code with the same codeword compositions. The conjecture is proved for the two length case, that is for U.D. codes with codeword lengths from the set $\{\ell_1,\ell_2\}$. This result is then extended to a more general class of U.D. codes and an equivalent conjecture is proposed for the general case.

1. Introduction

The basic elements of a discrete communication system are its source, encoder, channel, decoder and destination. The source may be represented as a random variable, X, taking on values from a source alphabet $\{x_1,x_2,\ldots,x_M\}$ with propabilities $p_1,p_2,\ldots,p_M$ respectively. A message is a sequence of source characters. To facilitate transmission, the encoder associates with every source character, x_i, a finite sequence of code characters from the code alphabet $\{a_1,a_2,a_3,\ldots,a_D\}$. Such a sequence of code characters is called a codeword. A code is the collection of all codewords. The encoded message is then transmitted over the channel which we assume to be noiseless. At the receiving end, the decoder attempts to reproduce the original message by assigning a set of source characters to the coded message.

To avoid ambiguity, every finite sequence of code characters must correspond to no more than one message. A code that conforms with this requirement is said to be a uniquely decipherable (U.D.) code. Furthermore, if no codeword is a prefix to some other codeword, the code is said to be a prefix (or instantaneous) code. A codeword w_j, is a prefix of codeword w_i, if $w_i=w_jb_j$, where w_j, b_j are in juxtaposition and b_j is some sequence of code characters.

The composition of a codeword w_i, is written as $(d_1(i),d_2(i),\ldots,d_D(i))$ where $d_j(i)$ is the number of times code character a_j appears in codeword w_i. Given a set of costs $(c_1,c_2,\ldots,c_D)$ associated with the respective code characters $\{a_1,a_2,\ldots,a_D\}$ the code cost is given by:

[1] This work is supported by the U.S. Air Force Office of Scientific Research, Systems Command, Grant AF-AFOSR-76-3017.

$$\text{Code cost} = \sum_{i=1}^{M} \sum_{j=1}^{D} p_i d_j(i) c_j$$

The problem of minimizing the code cost has been treated for prefix codes. Karp [2] gave an integer programming formulation and, if all the costs are equal, the problem is solved by the well-known Huffman algorithm [3]. However, no results exist for uniquely decipherable codes. It is known that a U.D. code can be replaced by a prefix code with the same codeword lengths.

Theorem 1: (Kraft [4]). A prefix code with codeword lengths $n_1, n_2, \ldots, n_M$ exists if, and only if, the following condition holds:

$$\sum_{i=1}^{M} D^{-n_i} \leq 1$$

where D is the number of characters in the code alphabet.

McMillan [5] showed that the same inequality must hold for uniquely decipherable codes. Thus, every U.D. code can be replaced by a prefix code with the same codeword lengths.

If the costs $(c_1, c_2, \ldots, c_D)$ are not all equal, the composition of the codewords becomes important. The following theorem, independently discovered by Block [6] and Carter and Gill [7], establishes necessary and sufficient conditions for the existence of a prefix code with a given set of compositions.

Theorem 2: Given a set of compositions of the form $C = \{(m_1, m_2, \ldots, m_D)\}$, there exists a prefix code with these exact codeword compositions if, and only if, the following inequality is satisfied for each $(m_1, m_2, \ldots, m_D) \in C$.

$$\prod_{k=1}^{D-1} \binom{\sum_{i=k}^{D} m_i}{m_k} \geq \sum_{1=1}^{\sum_{i=1}^{D} m_i} \sum_{\substack{n_1+n_2+\ldots=1 \\ n_i \leq m_i}} \left\{ \prod_{k=1}^{D-1} \binom{\sum_{i=k}^{D} (m_i - n_i)}{m_k - n_k} \right\} [n_1, n_2, \ldots, n_D]$$

where $[n_1, n_2, \ldots, n_D]$ is the number of codewords with composition $(n_1, n_2, \ldots, n_D)$.

A summary of the proof is described as follows:

The necessity of the condition is a direct result of the prefix property. Let a word be a sequence of code characters. Then the product term on the right side represents the number of words of composition $(m_1, m_2, \ldots, m_D)$ that have as prefix a word with composition $(n_1, n_2, \ldots, n_D)$. The product term on the left side is the number of all words of composition $(m_1, m_2, \ldots, m_D)$. The inequality states that the number of words of composition $(m_1, m_2, \ldots, m_D)$ must be greater than or equal to the number of words of composition $(m_1, m_2, \ldots, m_D)$ that are eliminated by the prefix property, plus the number $[m_1, m_2, \ldots m_D]$.

The sufficiency of the condition can be shown by constructing a prefix code with the given compositions. The construction proceeds from the shorter to the longer codewords. At each step, if we require $[m_1,m_2,\ldots,m_D]$ words of composition $(m_1,m_2,\ldots,m_D)$, it follows from the composition inequalities that there are at least $[m_1,m_2,\ldots,m_D]$ words of composition $(m_1,m_2,\ldots,m_D)$ and such that they do not have a prefix in the set of words chosen so far. The code, thus constructed, will be a prefix code with the required codeword compositions.

Example 1. Given the compositions (1,1),(1,1),(2,0),(1,2), is there a binary prefix code whose codewords have the given compositions?

The composition inequalities are:

$$\text{for } (1,1); \quad \binom{1+1}{1} \geq \binom{(1-1)+(1-1)}{1-1}[1,1] = \binom{0}{0}2$$

$$\text{or} \quad 2 \geq 2$$

$$\text{for } (2,0); \quad \binom{2+0}{2} \geq \binom{(2-2)+(0-0)}{2-2}[2,0] = \binom{0}{0}1$$

$$\text{or} \quad 1 \geq 1$$

$$\text{for } (1,2); \quad \binom{1+2}{1} \geq \binom{(1-1)+(2-1)}{1-1}[1,1] + \binom{(1-1)+(2-2)}{1-1}[1,2]$$

$$\text{or} \quad 3 \geq 3$$

All the composition inequalities are satisfied. The only prefix code with the given codeword compositions is: $W = \{01,10,00,110\}$. However, if one more composition (1,2) is added to the above set, the last inequality becomes:

$$\binom{1+2}{1} \geq \binom{(1-1)+(2-1)}{1-1}[1,1] + \binom{(1-1)+(2-2)}{1-1}[1,2]$$

$$\text{or} \quad 3 \geq 4$$

Thus, there is no prefix code with codeword compositions $\{(1,1),(1,1),(2,0),(1,2),(1,2)\}$. This is easily confirmed by constructing a binary tree.

It has been conjectured [6,7] that theorem 2 holds for U.D. codes. If this is true, the problem of constructing minimum cost uniquely decipherable codes reduces to that of constructing minimum cost prefix codes.

2. The Two Length Case

In this section we show that every uniquely decipherable code with codeword lengths from the set $\{\ell_1, \ell_2\}$ can be replaced by a prefix code with the same codeword compositions.

Let $W^2 = \{w_1, w_2, \ldots, w_M\}$ be a non-prefix uniquely decipherable code such that $W^2 = L_1 \cup L_2$ where:

$$L_1 = \{w_i \mid \lambda(w_i) = \ell_1\}$$

$$L_2 = \{w_j \mid \lambda(w_j) = \ell_2 > \ell_1\}$$

and $\lambda(w)$ is the length of codeword w.

Lemma 1. Let W be a non-prefix U.D. code, $w_i, w_j \in W$ and $w_j = w_i b_i$. Then $b_i w_i \notin W$.

Proof. Suppose that $b_i w_i = w_k \in W$. Consider the message $w_j w_i$. This message may be written as:

$$w_j w_i = w_i b_i w_i = w_i w_k$$

Hence, there would be two interpretations for the same message (i.e. $w_j . w_i$ or $w_i . w_k$) and the code is not a U.D. code. This contradicts our assumption that W is a U.D. code. Thus, $b_i w_i \notin W$.

Consider now the code W^2. For each codeword $w_j \in W^2$ that has a prefix in W^2, we construct a codeword chain, A_{w_j}. The first word in the chain is $w_j = w_i b_i$; the next word is $b_i w_i$. If $b_i w_i$ does not have a prefix in W^2 the chain terminates with $b_i w_i$. Otherwise, we can write $b_i w_i = w_k b_k$ ($w_k \in W^2$) and the next word in the chain is $b_k w_k$. This process is repeated until a word that has no prefix in W^2 is reached. It will be shown later that such a chain must terminate. We represent the chain associated with $w_j = w_i b_i$ by:

$$A_{w_j} = w_j = w_i b_i \rightarrow b_i w_i = w_k b_k \rightarrow b_k w_k = \ldots = w_x b_x \rightarrow b_x w_x = \ldots$$

Example 2. Let $W^2 = \{000, 001, 010, 100, 00101\}$. The codeword 00101 has the prefix 001 and the chain A_{00101} is given by:

$$A_{00101} = 00101 \rightarrow 01001 \rightarrow 01010 \rightarrow 10010 \rightarrow 10100$$

Clearly, all words in a chain A_{w_j}, have the same composition as they are *cyclic permutations of* w_j.

Theorem 3. Let $w_j \in W^2$, $w_j = w_i b_i$ ($w_i \in W^2$) and let A_{w_j} be the chain associated with w_j. Then, the first word of A_{w_j} is the only member of the chain that is a codeword of W^2.

Proof. Let A_{w_j} be:

$$A_{w_j} = w_j = w_i b_i \rightarrow b_i w_i = w_k b_k \rightarrow b_k w_k = w_l b_l \rightarrow b_l w_l = \ldots = w_x b_x \rightarrow b_x w_x = \ldots$$

Suppose that $b_x w_x$ is a member of the chain and also a codeword of W^2. Consider the message $B = w_j w_i w_k w_l \ldots . w_{x-1} w_x$. Since $w_j = w_i b_i$, we can write:

$$B = w_i b_i w_i w_k w_l \ldots . w_{x-1} w_x$$

but $b_i w_i = w_k b_k \Rightarrow B = w_i w_k b_k w_k w_l \ldots . w_{x-1} w_x$

but $b_k w_k = w_l b_l \Rightarrow B = w_i w_k w_l b_l w_l \ldots . w_{x-1} w_x$

$$\Rightarrow B = w_i w_k w_l \ldots \ldots b_{x-1} w_{x-1} w_x$$

but $b_{x-1} w_{x-1} = w_x b_x \Rightarrow B = w_i w_k w_l \ldots \ldots w_{x-1} w_x b_x w_x$

Now, if $b_x w_x \in W^2$, there would be two interpretations for message B. That is $B = w_j \cdot w_i \cdot w_k \ldots . w_x = w_i \cdot w_k \cdot w_l \ldots . w_x \cdot b_x w_x$. This contradicts our assumption that W^2 is a U.D. code. Therefore, $b_x w_x \notin W^2$ for any x.

We noted before that a chain A_{w_j}. does not terminate, because of the finite number of words of composition equal to that of w_j, the chain must cycle. In other words, at some point a word will be obtained that has already appeared in the chain. Then part of, or the whole chain, will be repeated indefinitely.

Theorem 4. Let $w_i, w_j \in W^2$ and $w_j = w_i b_i$. Let A_{w_j} be the chain associated with w_j. Then A_{w_j} contains no cycles.

Proof. First suppose that the chain cycles to a word other than w_j. That is:

$$A_{w_j} = w_j = w_i b_i \rightarrow b_i w_i = \ldots . = w_n b_n \rightarrow b_n w_n = \ldots . = w_p b_p \rightarrow b_p w_p = b_n w_n$$

Since the codewords w_n and w_p are of equal length (ℓ_1), it follows that $w_n = w_p$ and $b_n = b_p$. But then from the chain we have that:

$$b_{n-1} w_{n-1} = w_n b_n \text{ and } b_{p-1} w_{p-1} = w_p b_p$$

Since $w_n b_n = w_p b_p$ it follows that $b_{n-1} w_{n-1} = b_{p-1} w_{p-1}$. Then, by the same argument, $w_{n-1} = w_{p-1}$ and $b_{n-1} = b_{p-1}$. This process is repeated until we reach w_j. Then we have the following situation:

$$A_{w_j} = w_j = w_i b_i \rightarrow b_i w_i = \ldots \ldots . = w_1 b_1 \rightarrow b_1 w_1 = w_j$$

Theorem 3 states that the first word of A_{w_j} (i.e. w_j) is the only word that belongs to W^2. But if $b_1 w_1 \notin W^2 \Rightarrow w_j \notin W^2$, a contradiction. The

ambiguous message for this case is:

$$B = w_j \cdot w_i \cdot \cdots\cdots w_1 = w_i \cdot w_k \cdot \cdots\cdots w_j .$$

We now describe a method for replacing a uniquely decipherable code W^2 with a prefix code with the same codeword compositions as W^2. Given the U.D. code W^2, define the set S as follows:

$$S = \{w_j \in W^2 \mid w_j = w_i b_i \text{ for some } w_i \in W^2\}$$

Then, for each member of S, we construct a chain and form the set S' where:

$$S' = \{b_x w_x \mid w_x \in W^2 \text{ and } b_x w_x \text{ is the last word in a chain } A_{w_j},\ w_j \in S\}$$

Lemma 2. If $|C|$ is the cardinality of set C, we have $|S| = |S'|$.

Proof. This follows from the fact that no two chains can have any words in common. Suppose that the chains A_{w_k}, A_{w_j} have common words. Then,

$$A_{w_j} = w_j = w_1 b_1 \rightarrow b_1 w_1 = \cdots\cdots = w_i b_i \qquad b'_f w'_f = b_i w_i = w_{i+1} b_{i+1} \rightarrow \cdots$$

$$A_{w_k} = w_k = w'_1 b'_1 \rightarrow b'_1 w'_1 = \cdots\cdots = w'_f b'_f$$

By the argument used in the proof of Theorem 4, we have: $w_i = w'_f$, $b_i = b'_f$ which implies that $w_{i-1} = w'_{f-1}$, $b_{i-1} = b'_{i-1}$ and so on. The two chains cannot have equal lengths because that would require that $w_i = w_k$. Thus, the shorter of the two chains, say A_{w_k}, is a part of the longer one, A_{w_j}. But then the codeword $w_k \in W^2$ appears in A_{w_j}, while, by Theorem 3, w_j is the only member of A_{w_j} that belongs to W^2. Thus $w_k \notin W^2$, a contradiction.

The code $W = W^2 - S + S'$ is then a prefix code with the exact same codeword compositions as W^2.

Example 3. Consider the binary U.D. code W given by:

$$W^2 = \{000, 001, 010, 100, 00110, 10011, 01011, 01111, 11011, 00101\}$$

The set S is: $S = \{00101, 00110, 10011, 01011\}$

We now construct the chains associated with each member of S.

$$A_{00101} = 00101 \rightarrow 01001 \rightarrow 01010 \rightarrow 10010 \rightarrow 10100$$

$$A_{00110} = 00110 \rightarrow 10001 \rightarrow 01100$$

$$A_{10011} = 10011 \rightarrow 11100$$

$$A_{01011} = 01011 \rightarrow 11010$$

Then the set S' is: $S' = \{10100, 01100, 11100, 11010\}$ and the prefix code W' is:

$$W' = \{000, 001, 010, 100, 10100, 01100, 11100, 11010, 01111, 11011\}$$

3. An Extension of the Two Length Case

The method used for the two length case can be applied to a more general class of U.D. codes. The basic requirements are that the concept of a codeword chain is applicable and that Lemma 2 holds. Let W_p be a non-prefix U.D. code such that:

$$W_p = \{w_1, w_2, w_3, \ldots, w_M\} = P_1 \cup P_2$$

$$P_1 = \{w_i \mid \lambda(w_i) = \ell, \quad \ell > \lambda(w_j) \; \forall \; w_j \in P_2\}$$

and P_2 is both a prefix and a suffix code. A code P is said to be a suffix code if no codeword $w_i \in P$ is a suffix to some other codeword $w_j \in P$, i.e. $w_j \neq b_i w_i$ for any $w_i, w_j \in P$.

Lemma 3. All suffix codes are uniquely decipherable.

Proof. Suppose that suffix code $P = \{w_1, w_2, \ldots, w_M\}$ is not uniquely decipherable. Then there is at least one ambiguous message. Let that message be: $B = w_1 w_2 w_3 \ldots w_f = w'_1 w'_2 w'_3 \ldots w'_g$. Then we can either write $w_f = b'_g w'_g$ or $w'_g = b_g w_f$ for some sequences of code characters b_f, b'_g. Either case contradicts our assumption that P is a suffix code.

Prefix and suffix codes are very similar in structure. However, suffix codes are impractical because decoding may have to wait until the whole message is received. As in the two length case, we can construct codeword chains A_{w_j} for all $w_j \in P_1$ that have a prefix in P_2. Also, Theorem 3 still holds.

Lemma 4. Let $A_{w_j}, A_{w_k}, \ldots A_{w_n}$ be the chains associated with the codewords $w_j, w_k, \ldots, w_n \in P_1$. Then

a) no two (or more) chains can have common words
b) all chains must terminate

Proof. a) From Theorem 3 it follows that no chain can be part of another chain. Also, since $w_j \neq w_k \neq \ldots \neq w_n$ and P_2 is a prefix code, at any step in the construction of the chains there is no more than one way for the chain to grow (i.e. a chain can not split up into two or more subchains). Suppose that chains A_{w_j}, A_{w_k} merge:

$$A_{w_j} = w_j = w_1 b_1 \rightarrow b_1 w_1 = \ldots = w_i b_i$$

$$b_i w_i = b'_f w'_f = w_g b_g \rightarrow b_g w_g = \ldots.$$

$$A_{w_k} = w_k = w'_1 b'_1 \rightarrow b'_1 w'_1 = \ldots = w'_f b'_f$$

From Theorem 3, it follows that $b_i w_i = b'_f w'_f \in P_1$. Then, $w_i, w'_f \in P_2$ and

either w_i is a suffix of w'_f, or w'_f is a suffix of w_i. This contradicts our assumption that P_2 is a suffix code.

b)If a chain does not terminate, it must cycle. Let A_{w_j} be:

$$A_{w_j} = w_j = w_1 b_1 \rightarrow b_1 w_1 = w_2 b_2 \rightarrow b_2 w_2 = \ldots = w_f b_f \rightarrow b_f w_f = \ldots = w^j_{g-1} b_{g-1} \rightarrow$$
$$\rightarrow b_{g-1} w_{g-1} = w_g b_g = w_f b_f$$

Since $w_g, w_f \in P_2$ and P_2 is a prefix code, we must have $w_g = w_f$. Also, since P_2 is a suffix code it follows that $w_{g-1} = w_{f-1}$ and so on until we reach w_j. Then, by Theorem 3, $w_j \in W_p$ a contradiction. Q.E.D.

Example 4. Consider the code $W = \{001,010,100,1010,00101\}$. We have:

$$A_{00101} = 00101 \rightarrow 01001 \rightarrow 01010 \rightarrow 10100 \rightarrow 01010 \rightarrow \ldots$$

The chain cycles. By Theorem 4b the code is not uniquely decipherable. One ambiguous message is:

$$B = 001 \cdot 010 \cdot 010 \cdot 100 \cdot 1010 = 00101 \cdot 00101 \cdot 001 \cdot 010$$

We have then shown that the U.D. code W_p can be replaced by a prefix code with the same codeword compositions.

Example 5. Consider the U.D. code

$$W_p = \{00,010,101,0110,1110,00111,10111,01001,11111\}$$

We have: $P_1 = \{00111,10111,01001,11111\}$

$P_2 = \{00,010,101,0110,1110\}$

and $A_{00111} = 00111 \rightarrow 11100 \rightarrow 01110$

$A_{10111} = 10111 \rightarrow 11101 \rightarrow 11110$

$A_{01001} = 01001 \rightarrow 01010 \rightarrow 10010$

The required prefix code is $W' = W_p - S + S'$ or

$$W' = \{00,010,101,0110,1110,01110,11110,10010,11111\}$$

4. The General Case

We start with the following conjecture.

Conjecture A. Every uniquely decipherable code must satisfy the composition inequalities established for prefix codes.

Suppose that conjecture A is false. Then there is a uniquely decipherable code such that at least one of the composition inequalities is not satisfied. Let $W^* = \{w_1, w_2, \ldots, w_M\}$ be such a code, and let $(m_1, m_2, \ldots, m_D)$ be a composition of minimum length,

$(\sum_{i=1}^{D} m_i)$, for which the corresponding inequality is not satisfied. Consider the code W defined by:

$$W = \{w_i \mid m_j(i) \leq m_j,\ \exists\, j \ni m_j(i) < m_j,\ 1 \leq j \leq D\}$$

Code W is a U.D. code because it is a subcode of W^*. Furthermore, all the composition inequalities are satisfied for W. Therefore, there exists a prefix code, W', with the same codeword compositions as W. Let R be the set:

$$R = \{w_j \in W \mid k(w_j) = (m_1, m_2, \ldots, m_D)\}$$

where $k(w_j)$ is the composition of w_j. Then, the code WUR is a U.D. code but there is no prefix code with the same codeword compositions as WUR. This means that $|R|$ is greater than the number of words of composition $(m_1, m_2, \ldots, m_D)$ that can be added to W' without violating the prefix property. Equivalently, the U.D. code W eliminates less words of composition $(m_1, m_2, \ldots, m_D)$ than the prefix code W'. Thus, conjecture A is equivalent to:

Conjecture B. Given any non-prefix U.D. code $W = \{w_1, w_2, \ldots, w_M\}$ and a prefix code $W' = \{w_1', w_2', \ldots, w_M'\}$ such that $k(w_i) = k(w_i') = (m_1(i), m_2(i), \ldots, m_D(i))$ for $1 \leq i \leq M$, let the composition $(m_1, m_2, \ldots, m_D)$ satisfy: $m_j \geq m_j(i)$ for all i,j and $\exists\, j \ni m_j > m_j(i)$, $i \leq i \leq M$, $1 \leq j \leq D$. Then code W eliminates at least as many words of composition $(m_1, m_2, \ldots, m_D)$ as code W'.

Code W' eliminates those words of composition $(m_1, m_2, \ldots m_D)$ that can be written as $w_i' b_j(i)$, $w_i' \in W'$. Consider the word pairs $(w_i b_j(i), b_j(i) w_i)$. From Lemma 1, it follows that at least one word from each pair is eliminated by W (if borh words are added to W, the augmented code is not a U.D. code). Since the number of pairs is equal to the number of words of composition $(m_1, m_2, \ldots, m_D)$ that are eliminated by W', it follows that conjecture B is true if the words $w_i b_j(i)$, $b_j(i) w_i$ are distinct for all i,j. Unfortunately, this is not the case. In certain cases, the matching that may occur, can be resolved by considering non-cyclic permutations. In general however, the question remains open.

5. Conclusion

It is proved that any uniquely decipherable code with codeword lengths from the set $\{\ell_1, \ell_2\}$ can be replaced by a prefix code with the same codeword compositions. The proof is based on the concept of a codeword chain. This result is then extended to the more general class of uniquely decipherable codes described by: $W_p = P_1 U P_2$, where $P_1 = \{w_j \mid \lambda(w_j) = \ell,\ \ell > \lambda(w_i) \forall\, w_i \in P_2\}$ and P_2 is both a prefix and a suffix code.

The general case is then discussed and, using a theorem discovered by Block [6] and Carter and Gill [7], a conjecture on the replacement of uniquely decipherable codes by prefix codes with the same codeword compositions is proposed. Some implications of this conjecture are then discussed. No device analogous to a codeword chain could be found for the general case. Attempts to prove or disprove the conjecture have been unsuccesful so far.

BIBLIOGRAPHY

1. R. Ash, Information Theory, Interscience Publishers, New York, 1965, pp. 1-44.

2. R. M. Karp, "Minimum Redundancy Coding for Discrete Memoryless Channel," IRE Trans. Information Theory, Vol. IT-7, 1961, pp. 27-38.

3. D. A. Huffman, "A Method for the Construction of Minimum Redundancy Codes," Proc. IRE, Vol. 40, 1952, pp. 1098-1101.

4. L. G. Kraft, "A Device for Quantizing, Grouping and Coding Amplitude Modulated Pulses," M.S. Thesis, 1949, Electrical Engineering Dept., MIT, Cambridge, Mass.

5. B. McMillan, "Two Inequalities Implied by Unique Decipherability," IRE Trans. Information Theory, Vol. IT-w, 1956, pp. 115-116.

6. J. R. Block, "On Variable Length Codes with Prescribed Composition," M.S. Thesis, 1975, Electrical Engineering Dept., Northwestern University, Evanston, Ill.

7. L. Carter and J. Gill, "Conjectures on Uniquely Decipherable Codes," IEEE Trans. Information Theory, Vol. IT-20, 1974, pp. 394-396.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR-77-0026
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): ON UNIQUELY DECIPHERABLE CODES WITH GIVEN COMPOSITIONS
5. TYPE OF REPORT & PERIOD COVERED: Interim rept.
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): S. C. Ntafos, S. L. Hakimi
8. CONTRACT OR GRANT NUMBER(s): AF-AFOSR-3017-76
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Northwestern University, Dept of Electrical Engineering ~~& Computer Science~~, Evanston, Illinois 60201
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F, 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS: Air Force Office of Scientific Research/NM, Bolling AFB, Washington, DC 20332
12. REPORT DATE: 1976
13. NUMBER OF PAGES: 10
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 11p.
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

variable length codes, codeword composition, uniquely decipherable codes, prefix codes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It has been conjectured that a uniquely decipherable (U.D.) code can be replaced by a prefix code with the same codeword compositions. The conjecture is proved for the two length case, that is for U.D. codes with codeword lengths from the set $\{l_1, l_2\}$. This result is then extended to a more general class of U.D. codes and an equivalent conjecture is proposed for the general case.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 400 351

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)