J. KSIAM **Vol.6**, No.1, 47-52, 2002

DECODING OF LEXICODES $\mathbf{S_{10,4}}$

D. G. Kim

Abstract. In this paper we propose a simple decoding algorithm for the 4-ary lexicographic codes (or lexicodes) of length 10 with minimum distance 4, write $S_{10,4}$. It is based on the syndrome decoding method. That is, using a syndrome vector we detect an error and it will be corrected an error from the four parity check equations.

1. Introduction

In this paper, we shall introduce the surprising arithemetical operations which are used in the Game of Nim. Under these operations, the lexicodes are linear over some finite field. Their definition is derived from a greedy algorithm, that is, each codeword is chosen as the first word not prohibitively near to previous codewords.

The main aim of this paper is to find an decoding algorithm of the 4-ary [10, 6, 4] lexicodes, write $\mathbf{S_{10,4}}$. Using a syndrome vector and the four parity check equations, we correct one error in received vector.

This paper is arranged as follows. The nim operation is introduced in section 2, the lexicodes with base 2^{2^a} are discussed in section 3. In particular we obtain the six basis of the 4-ary lexicodes $\mathbf{S_{10,4}}$. Section 4 gives a decoding algorithm and decoding examples for this code.

2. Nim operation

First, we define the two operations which are called the nim-addition $\oplus$ and nim-multiplication $\otimes$ in that game.

Research partially supported by Chungwoon University Grant.

1991 *Mathematics Subject Classification.* 94B35.

Key words and phrases. nim-operations, minimum distance, lexicographic codes, parity check matrix, syndrome.

Definition 1. *Let x' be a variable that ranges over all elements strictly less than x and mex the least non-negative integer not of the form. Then we define the two operations:*
(1) $a \oplus b = mex\{a' \oplus b,\ a \oplus b'\}$
(2) $a \otimes b = mex\{(a' \otimes b) \oplus\ (a \otimes b') \oplus (a' \otimes b')\}$

Two operations, $\oplus$ and $\otimes$, convert the numbers $0, 1, 2, \cdots$ into a field of characteristic 2. Also, for $a \geq 0$, the numbers less than 2^{2^a} form a subfield and isomorphic to the Galois field GF(2^{2^a}).

Theorem 2 ([2]). *The nim-operations turn the set of non-negative integers into a field of characteristic* 2*.*

Using the field laws, we shall fill out the first 4 by 4 corner of the addition and multiplication tables in nim. Consider the nim-addition of any two numbers from $0, 1, 2, 3$.

Theorem 3 ([1]). *We have $x \oplus 0 = 0 \oplus x = x$, for every number x.*

Since $\{0, 1, 2, 3\}$ is a field of characteristic 2, we have $x \oplus x = 0$ for all $x \in \{0, 1, 2, 3\}$. By Theorem 3, $1 \oplus 2$ can not be one of $0, 1, 2$ and so must be 3. Since $1 \oplus 3 \neq 0, 1, 3$, it must be 2. In the same way, we have $2 \oplus 3 = 1$. Therefore the sum of any two distinct numbers from $1, 2, 3$ is the third.

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	0	3	2
2	2	3	0	1
3	3	2	1	0

There is a nim-multiplication $\otimes$ which together with nim-addition $\oplus$ converts the integers into a field [2]. With nim-multiplication, we know that $0 \otimes x$ must be 0 which is the zero of the field. Also $1 \otimes x$ must be x. Since the elements other than 0, 1 satisfy $x^2 = x \oplus 1$ (here x^2 means $x \otimes x$) in the field GF(4), we have $2 \otimes 2 = 2 \oplus 1 = 3$ and $3 \otimes 3 = 3 \oplus 1 = 2$. Next $2 \otimes 3$ can not be one of $0, 2, 3$ and so must be 1.

$\otimes$	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	3	1
3	0	3	1	2

The following is a rule enabling us to perform nim-additions. In its statement, the term 2-power means a power of 2, such as $1, 2, 4, 8, \cdots$, in the ordinary sense:

(i) If x is a 2-powers and $y < x$, then $x \oplus y = x + y$.
(ii) $x \oplus x = 0$ for any x.

For example, $15 \oplus 5 = (8 \oplus 4 \oplus 2 \oplus 1) \oplus (4 \oplus 1) = 8 \oplus 2 = 10$, since both 4's and 1's are cancelled.

For finite numbers, the nim-multiplication follows from the following rules, similar to those for nim-addition. In the following statement, the term *Fermat 2-power* means the number 2^{2^n}, such as $2, 4, 16, 256 \cdots$, in the ordinary sense:

(i) If x is a Fermat 2-powers and $y < x$, then $x \otimes y = x \times y$.
(ii) $x \otimes x = \frac{3}{2} \times x$ for any Fermat 2-power x.

For example $16 \otimes 2 = 32$, since $16 = 2^{2^2}$. By an equation (ii), we have $2^2 = 2 \times \frac{3}{2} = 3$, $4^2 = 4 \times \frac{3}{2} = 6$, $16^2 = 16 \times \frac{3}{2} = 24, \cdots$.
Using the associative and distributive laws, $19 \otimes 11 = (16 \oplus 2 \oplus 1) \otimes (8 \oplus 2 \oplus 1) = (16 \otimes 8) \oplus (16 \otimes 2) \oplus (16 \otimes 1) \oplus (2 \otimes 8) \oplus (2 \otimes 2) \oplus (2 \otimes 1) \oplus (8 \oplus 2 \oplus 1) = 128 \oplus 32 \oplus 16 \oplus (2 \otimes 8) \oplus 2 \oplus 8 = 128 \oplus 32 \oplus 16 \oplus 4 \oplus 2 = 182$, since $2 \otimes 8 = 2 \otimes (4 \otimes 2) = 4 \otimes 2^2 = 4 \otimes 3 = 8 \oplus 4$.
Next, we compute the inverse value 15^{-1} satisfying $15 \otimes 15^{-1} = 1$. $15 \otimes 4 = (8 \oplus 4 \oplus 2 \oplus 1) \otimes 4 = (8 \otimes 4) \oplus (4 \otimes 4) \oplus (2 \otimes 4) \oplus (1 \otimes 4) = (2 \otimes 4 \otimes 4) \oplus 6 \oplus 8 \oplus 4 = (2 \otimes 6) \oplus (4 \oplus 2) \oplus 8 \oplus 4 = (2 \otimes (4 \oplus 2)) \oplus 2 \oplus 8 = 8 \oplus 3 \oplus 2 \oplus 8 = 3 \oplus 2 = 1$. Hence $15^{-1} = 4$.

3. Lexicodes

Consider the lexicodes with base $B = 2^{2^a}$. A word of this codes is a sequence $\mathbf{x} = \cdots x_3 x_2 x_1$, $x_i \in \{0, 1, \cdots, 2^{2^a} - 1\}$. For a convenience, we omit leading zeros (i.e., $012 = 12$). The set of words is ordered lexicographically, i.e., the word $\mathbf{x} = \cdots x_3 x_2 x_1$ is smaller than the word $\mathbf{y} = \cdots y_3 y_2\, y_1$, written $\mathbf{x} < \mathbf{y}$, if for some n we have $x_n < y_n$, but $x_N = y_N$ for all $N > n$. For example, $123 < 132$, $312 < 1032$.
Lexicodes are defined by saying a word is in the code if it does not conflict with any earlier codewords. That is, the lexicode with minimum distance d is defined by saying that two words do not conflict if the Hamming distance between them is not less than d. We write $S_{n,d}$ for the lexicode consisting of the codewords with base 4, length n or less and minimum distance d.

Example 1. Applying the greedy algorithm, then the lexicode $S_{4,3}$ contains the codewords, 0, 111, 222, 333, 1012, 1103, 1230, 1321, 2023, 2132, 2201, 2310, 3031, 3120, 3213, 3302.

In [3], Conway and Sloane show that the lexicode with base $B = 2^a$ is closed under coordinatewise nim-addition, and if $B = 2^{2^a}$, the lexicode is closed under coordinatewise nim-multiplication by scalars k, $k \in \{0, 1, \cdots, 2^{2^a} - 1\}$. As a result we provide the following Lexicode Theorem.

Theorem 4 ([3]). *If B is of the form 2^{2^a}, then the lexicode is a linear code over the Galois field GF(B).*

Now we consider the lexicodes $S_{10,4}$. Let $\mathbf{e}_i$ be the basis of lexicode $S_{10,4}$. It is easily checked that we have the first 3 bases $\mathbf{e}_1 = 1111$, $\mathbf{e}_2 = 10123$ and $\mathbf{e}_3 = 100132$. Since $S_{10,4}$ is a 6-dimensional vector space, this code has 6 bases. So we need to find the basis $\mathbf{e}_4$, $\mathbf{e}_5$ and $\mathbf{e}_6$ of this code.

Theorem 5. *For each i $(3 \leq i \leq 5)$, if $\mathbf{e}_{i+1}$ is the smallest codeword with more digits than $\mathbf{e}_i$, then $\mathbf{e}_4 = 11000011$. Moreover we have $\mathbf{e}_5 = 101000023$ and $\mathbf{e}_6 = 1001000032$.*

Proof. In [3, Table IV], 7 digit codewords are not possible. So we find the smallest eight digit codeword. For k, $a \in \{0,1,2,3\}$, $10aaaaaa$ is impossible for the same reasons that $01aaaaaa$ is impossible. Also 11000000, $1100000a$ and $110000a0$ conflict with 00000000. So $\mathbf{e}_4$ may be $110000aa$. Assume $\mathbf{e}_4 = 11000011$. If $\mathbf{c}$ is a linear sum of any two bases of $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$, then $d(\mathbf{e}_4,\ k \otimes \mathbf{e}_i) = d(\mathbf{e}_4,\ \mathbf{c}) \geq 4$, $i = 1,2,3$, from the last 2 places of $\mathbf{e}_4$, and at least 2 places of $k \otimes \mathbf{e}_i$ and $\mathbf{c}$.

If $\mathbf{c}$ is a linear sum of $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$, then $d(\mathbf{e}_4,\ \mathbf{c}) \geq 5$ from the last 2 places of $\mathbf{e}_4$ and at least 3 places of $\mathbf{c}$.

By the similar way, we can obtain the bases $\mathbf{e}_5$ and $\mathbf{e}_6$. □

4. Decoding Method

In this section, we shall obtain a 4 by 10 parity check matrix H using the 6 bases of $S_{10,4}$. For a given received vector $\mathbf{r}$, this matrix H gives a syndrome vector $\mathbf{s} = \mathbf{r} \otimes H^T$, where H^T is a transpose of H. If the syndrome is nonzero, this implies that an error occurred in the received vector.

Let $\mathbf{r}$ be a received vector, $\mathbf{r} = r_{10}\ r_9\ r_8\ r_7\ r_6\ r_5\ r_4\ r_3\ r_2\ r_1$. If r_i $(i = 1,2,3,7)$ is incorrect, these equations yield three 0s and one nonzero, and respectively three nonzeros and one 0 if r_i $(i = 4,5,6,8,9,10)$ is incorrect. In other cases, we conclude that more than one error has been made. In particular, if the syndrome $\mathbf{s}$ is a multiple of the ith column vector of H, then r_i is not correct. Using the syndrome vector, we can detect an errored coordinate in the received vector.

Now, all the arithmetic operations are in the nim-sense (nim-additon and nim-multiplication). So we write $x + y$ for $x \oplus y$, and xy for $x \otimes y$.

Note : Let $\mathbf{c}$ be a codeword, $\mathbf{c} = c_{10}\ c_9\ c_8\ c_7\ c_6\ c_5\ c_4\ c_3\ c_2\ c_1$, $\mathbf{c} = \sum_{i=1}^{6} x_i \mathbf{e}_i$, $x_i \in \{0,1,2,3\}$ Then we have $c_1 = x_1 + 3x_2 + 2x_3 + x_4 + 3x_5 + 2x_6$, $c_2 = x_1 + 2x_2 + 3x_3 + x_4 + 2x_5 + 3x_6$, $c_3 = x_1 + x_2 + x_3$, $c_{i+3} = x_i$ $(i = 1,2,3)$, $c_7 = x_4 + x_5 + x_6$ and $c_{i+4} = x_i$ $(i = 4,5,6)$. If $\mathbf{r}$ has no error, then the four parity check equations yield $0,0,0,0$ as the following these :

$$r_{10} + r_9 + r_8 + r_7 = 0 \tag{1}$$

$$r_6 + r_5 + r_4 + r_3 = 0 \tag{2}$$

$$3r_{10} + 2r_9 + r_8 + 3r_6 + 2r_5 + r_4 + r_2 = 0 \tag{3}$$

$$2r_{10} + 3r_9 + r_8 + 2r_6 + 3r_5 + r_4 + r_1 = 0 \tag{4}$$

From the four parity check equations and a property of $x \oplus x = 0$, we obtain a coordinate c_i, where $c_1 = r_4{+}3r_5{+}2r_6{+}r_8{+}3r_9{+}2r_{10}$, $c_2 = r_4{+}2r_5{+}3r_6{+}r_8{+}2r_9{+}3r_{10}$, $c_3 = r_4{+}r_5{+}r_6$, $c_4 = r_3{+}r_5{+}r_6$, $c_5 = r_3{+}r_4{+}r_6$, $c_6 = r_3{+}r_4{+}r_5$, $c_7 = r_8{+}r_9{+}r_{10}$, $c_8 = r_7 + r_9 + r_{10}$, $c_9 = r_7 + r_8 + r_{10}$, $c_{10} = r_7 + r_8 + r_9$. Therefore we can obtain a desired codeword.

These the four equations give a parity check matrix H as the following this :

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 3 & 2 & 1 & 0 & 1 & 0 \\ 2 & 3 & 1 & 0 & 2 & 3 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

For example, let $\mathbf{r} = 3012221020$ be a received vector. Then by nim-multiplication of matrix, we have the syndrome $\mathbf{r}H^T = (0,1,2,3)$. Since this vector is the 5th column vector of H, the 5th coordinate of $\mathbf{r}$ is not correct. Therefore we obtan $c_5 = r_3{+}r_4{+}r_6 = 0 + 1 + 2 = 3$ and then have a desired codeword $\mathbf{c} = 3012231020$.

Now, we give a decoding algorithm of $\mathbf{S_{10,4}}$.

Algorithm

Step 1 : First, we compute the syndrome vector $\mathbf{s}$. If $\mathbf{s}$ is a multiple of the ith column of H, we go to step 2.

Step 2 : Since r_i is not correct, r_i is replaced by c_i.

Example 2. Let $\mathbf{r} = 1232012331$. Since $\mathbf{s} = (2,0,0,0)$ is a multiple of the 7th column vector of H, then r_7 is not correct. Hence $c_7 = r_8 + r_9 + r_{10} = 1 + 2 + 3 = 0$, and so we get the desired codeword $\mathbf{c} = 1230012331$.

Example 3. Let $\mathbf{r} = 2131112202$. Since $\mathbf{s} = (1,0,3,2)$ is a multiple of the 10th column vector of H, then r_{10} is not correct. So we have $c_{10} = 3r_1 + 3r_4 + 2r_5 + r_6 + 3r_8 + 2r_9 = 1 + 1 + 2 + 1 + 2 + 2 = 3$. Hence we get $\mathbf{c} = 3131112202$.

Example 4. Let $\mathbf{r} = 3012221020$. Since $\mathbf{s} = (0, 1, 2, 3)$ is a multiple of the 5th column vector of H, then r_5 is not correct, and so $c_5 = 3r_1+2r_2+r_4+r_8+r_9 = 0+3+1+1+0 = 3$. Therefore we get $\mathbf{c} = 3012231020$.

Example 5. Let $\mathbf{r} = 213313011$. Since $\mathbf{s} = (0, 1, 0, 0)$ is a multiple of the 3th column vector of H, then r_3 is not correct. Hence we obtain $c_3 = r_4 + r_5 + r_6 = 3 + 1 + 3 = 1$. Therefore we get $\mathbf{c} = 213313111$.

REFERENCES

[1] J.H. Conway, *Integral Lexicographic Codes*, Discrete Mathematics 83(1990) 219-235.
[2] J.H. Conway, On Numbers and Games, Academic Press, New York, 1976.
[3] J.H.Conway and N.J.A. Sloane, *Lexicographic Codes: Error-Correcting Codes from Game Theory*, IEEE Trans. Inform. Theory IT-32(3) (1986) 337-348.

Liberal Arts and Science
Chungwoon University, Hongsung
Chungnam 350-701, Korea
e-mail codekim@www.chungwoon.ac.kr