E4.07
CS7.23
SO11
ISE4.15

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2009

MSc and EEE/ISE PART IV: MEng and ACGI

Corrected Copy

CODING THEORY

Wednesday, 6 May 10:00 am

Time allowed: 3:00 hours

There are SIX questions on this paper.

Answer FOUR questions.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible	First Marker(s) :	A.A. Ivanov
	Second Marker(s) :	C. Ling

A table of the field of order 16

log	0	1	12	2	9	13	7	3	4	10	5	14	11	8	6
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	3	4	6	8	10	12	14	9	11	13	15	1	3	5	7
3	2	1	5	12	15	10	9	1	2	7	4	13	14	11	8
4	5	6	7	9	13	1	5	11	15	3	7	2	6	10	14
5	4	7	6	1	8	7	2	3	6	9	12	14	11	4	1
6	7	4	5	2	3	13	11	2	4	14	8	3	5	15	9
7	6	5	4	3	2	1	12	10	13	4	3	15	8	1	6
8	9	10	11	12	13	14	15	15	7	6	14	4	12	13	5
9	8	11	10	13	12	15	14	1	14	12	5	8	1	3	10
10	11	8	9	14	15	12	13	2	3	11	1	5	15	8	2
11	10	9	8	15	14	13	12	3	2	1	10	9	2	6	13
12	13	14	15	8	9	10	11	4	5	6	7	6	10	7	11
13	12	15	14	9	8	11	10	5	4	7	6	1	7	9	4
14	15	12	13	10	11	8	9	6	7	4	5	2	3	2	12
15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	3

Below diagonal $a + b$, on or above $a \times b$,
$0 + a = a$, $a + a = 0$, $0 \times a = 0$

1. Let D be the binary $(7,4)$-code given by the check matrix

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

and let C be the *shortened* code obtained from D by removing the final zero from all the code words that end in zero. Thus $u \in C$ if and only if $(u, 0) \in D$.

Construct (with justification) generator and check matrices for C. *You should state but need not prove the relation between standard form generator and check matrices of a binary linear code.*

[10]

Using your check matrix construct (with justification) a syndrome/coset leader decoding table for C.

[10]

Where there is a choice for the coset leader, indicate all possible values.

[5]

2. Let C be a (not necessarily binary) linear (n, m)-code and let C' be the *punctured* code obtained by deleting a particular position from all the code words of C. Show that if C has minimal distance $d > 1$, then C' is a $(n-1, m)$-code with minimal distance $d' \geq d - 1$.

[7]

Hence or otherwise show that the minimal distance of any linear (n, m)-code C is at most $n - m + 1$. *This is the Singleton bound. If the minimal distance is exactly equal to $n - m + 1$ we shall say the code meets the Singleton bound.*

[7]

Define an r-perfect code.

[1]

Show that the binary Hamming code Ham(4) is 1-perfect, but does not meet the Singleton bound.

[5]

Show that the triple error correcting Reed–Solomon code RS$(4, 3)$ considered as a code over GF(16) meets the Singleton bound but it is not perfect. *You may assume that* RS$(4, 3)$ *corrects* 3 *errors and that there are exactly* $1559476 = 4 \times 389869$ *words at distance* ≤ 3 *from a code word of* RS$(4, 3)$.

[5]

3. Explain how the field table of the field of order 16 on page 1 is constructed, emphasizing the properties of the polynomial x^4+x^3+1 that ensure that the element 2 is a *primitive element*, and explaining how the 'logarithms' at the head of the table can be used to multiply and find inverses in the field.

[10]

Prove that every finite field contains a primitive element.

[15]

4. Describe Euclid's algorithm in its four column version.

[10]

Prove that Euclid's algorithm determines the highest common factor of its input.

[5]

Let $t(x)$ be an irreducible binary polynomial and let $s(x)$ be an non-zero binary polynomial such that $\deg(s(x)) < \deg(t(x))$. Stating clearly any properties of Euclid's algorithm you require, show how that algorithm can be used to determine an inverse polynomial $r(x)$ of $s(x) \bmod t(x)$ that is a polynomial $r(x)$ such that $r(x)s(x) \equiv 1 \bmod t(x)$.

[5]

Illustrate your method by finding an inverse of $x^6 + x^4 + x^3 + x + 1 \bmod x^7 + x + 1$ (*you may assume that* $x^7 + x + 1$ *is irreducible*).

[5]

5. Explain how the binary, t-error correcting BCH code $\mathrm{BCH}(k,t)$ is constructed by extending the check matrix of the Hamming code $\mathrm{Ham}(k)$, describing the two check matrices $H_{k,t}$ and $V_{k,t}$.

[5]

Show that the matrices $H_{k,t}$ and $V_{k,t}$ define the same code.

[8]

Use the matrices $H_{k,t}$ and $V_{k,t}$, and the properties of check matrices in general, and of Vandermonde type matrices in particular (which you should state, but need not prove) to give estimates of the parameters of the codes.

[4]

Suppose now that the check matrix of the Hamming code H_4 is written in the form

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

Construct the check matrix $H_{4,2}$ of the code $\mathrm{BCH}(4,2)$.

[8]

6. The 4-error correcting code RS(4, 4) based on the primitive element 2 is used to transmit a message. One received word is

$$u = (1,\ 10,\ 3,\ 4,\ 4,\ 2,\ 2,\ 13,\ 15,\ 3,\ 6,\ 7,\ 1,\ 1,\ 0).$$

The first 7 syndromes of $S_1, \ldots, S_7$ of u are 10, 8, 12, 4, 7, 0, 1.

Calculate the remaining syndrome S_8.

[5]

Using the syndromes calculate the error locator and the error evaluator for the word.

[10]

Verify that the error locator has roots 6, 11, 12, 15 and calculate a code word at distance ≤ 4 from u.

[10]

E4.07 | I & 4.15 | 2011 | CS7.23

1/9

Coding Theory 2009

SOLUTION 1

We use the fact that the standard form matrix $\binom{I}{A}$ is a generator for a binary code if and only if the standard form matrix $(A, J))$ is a check matrix for that code.

Thus a generator matrix for D is

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}^{\mathsf{T}}.$$

A code word of D is the sum of a subset of the columns of this matrix and it will end in 0 iff that set does not include the last column. Hence a generator matrix G for C, and the corresponding check matrix H are given by

unseen 10

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}^{\mathsf{T}}, \quad H = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The syndromes of single errors are the columns of H and ambiguities occur only if the same column occurs several times Thus the first part of the syndrome/coset leader table has the following form (the syndromes and error patterns are written as rows for convenience):

syndr.	error
000	000000
001	010000 or 000001
010	100000 or 000010
100	000100
011	001000

unseen 10

That exhausts all the columns of H but there are still three syndromes left. These we represent as sums of two columns of H. It is helpful to note, that the only column of H with a 1 in the first row is the 4th. Thus an error producing syndrome starting with 1 must include have a 1 in its 4th position. An error word $abc1ef$ will produce the syndrome $1yz$ iff the error word $abc0ef$ produces the syndrome $0yz$. Since all the outstanding syndromes do start with 1 this makes it straightforward to determine all the corresponding error patterns of weight 2.

syndr.	error
101	010100 or 000101
110	100100 or 000110
111	001100

unseen 5

2/9

SOLUTION 2

Since the minimum distance of C is > 1 two distinct code words of C cannot yield the same word when a single symbol is deleted. Thus C' has exactly the same number of code words as C. It is also obviously still linear. Since the number of code words of a linear code on q symbols of rank k is q^k, it follows that C' has the same rank as C.

Let u and v be two code words of C and u' and v' the corresponding code words of C'. Then if the deleted symbols of u and v are the same we have $d(u', v') = d(u, v)$ and otherwise $d(u', v') = d(u, v) - 1$. Hence the minimum distance of C' is $d - 1$ if there are two code words u, v of C with $d(u, v) = d$ and different symbols in the punctured location. Otherwise the minimum distance of C' is d. Ejercic. 5

Now we puncture the code repeatedly until the minimum distance is 1 that happens at the latest when the block length equals the rank m of C which would occur after $n-m$ steps. At each step the minimum distance has dropped by at most 1 so the original minimum distance is at most $1+n-m$ as claimed.

A code of block length n is r-perfect if to every word v of length n there is exactly one code word u with $d(u, v) \leq r$. Teen 10

The parameters of the code Ham(4) are $15, 11, 3$ since its check matrix H_4 has as its columns all non-zero binary words of length 4. It is perfect single-error correcting because for any word v of length 15, we have either $H_4 v = \underline{0}$ in which case $v \in$ Ham(4) or $H_4 v$ is a unique column of H_4. Correcting the corresponding bit of v will then produce a code word and changing any other or none will not do so. This also proves that the minimum distance of Ham(4) is 3 rather than 5 which would correspond to the Singleton bound. unseen 5

The code RS(4) has as its generator polynomial $\prod_{k=1}^{6}(x - \alpha^k)$, where α is a primitive element of GF(16). Since its code words are represented by polynomials of degree 14 it has block length 15 and rank $14 - 6 + 1 = 9$ Since it can correct three errors it has minimum distance at least 7. As $7 = 15 - 9 + 1$ the minimum distance cannot be greater and the code meets the Singleton bound. If it is perfect then $|\mathrm{RS}(4, 3)| \times 1559476 = 16^{15}$. However $|\mathrm{RS}(4, 3)| = 16^9$ which would imply that 1559476 is a power of 16 which is not the case. so the code is not perfect. unseen 5

3/9

SOLUTION 3

The number n represents its binary 4-tuple a, b, c, d, which in turn represents the polynomial $ax^3 + bx^2 + cx + d$. Thus $13 \sim 1, 1, 0, 1 \sim x^3 + x^2 + 1$. Addition (the lower half of the table) is ordinary addition over $\mathbb{B}$ (i.e. XOR). In particular $\alpha + \alpha = 0$ for any α, so this does not appear in the table. Multiplication (the upper half of the table) is multiplication in $\mathbb{B}[x]$, followed if necessary, by taking the remainder after division by $x^4 + x^3 + 1$.

The fact that 2 which corresponds to x is a primitive root is equivalent to the statement that $x^4 + x^3 + 1$ divides $x^{15} - 1$ but not $x^k - 1$ for any smaller power k. Then the powers $2^0, \ldots, 2^{14}$ must all be distinct and so the cover all the non-zero elements of the field. These powers are the logarithms at the head of the table. To multiply field elements α and β we represent them as powers of 2, say 2^k and 2^ℓ. Then $\alpha\beta = 2^{k+\ell}$. Since $2^{15} = 1$ we can reduce $k + \ell$ mod 15. Then the product is the element with that value as its logarithm. In particular, $\alpha^{-1} = 2^{15-k}$

seen to

THEOREM. *Every finite field has a primitive element.*

We first prove a lemma:

LEMMA. *If $\beta, \gamma \in F$ and $\mathrm{ord}(\beta) = m$ and $\mathrm{ord}(\gamma) = n$, and their highest common factor $(m, n) = 1$, then $\mathrm{ord}(\beta\gamma) = mn$.*

Proof. Certainly $(\beta\gamma)^{mn} = 1$. Suppose $(\beta\gamma)^k = 1$. Then $\beta^{kn} = \beta^{kn}\gamma^{kn} = (\beta\gamma)^{kn} = 1$. Hence m divides kn. So n divides km. Now m and n have no prime factors in common, so m divides kn only if it divides k. Similarly $\gamma^{km} = 1$ and n divides km only if it divides k. So both m and n divide k. Thus their least common multiple, mn divides k. □

Proof. Let $p_1, \ldots, p_k$ be the prime factors of $q - 1$ and let s_i be the highest power of p_i that divides the order of some element γ_i of F. By taking γ_i to a suitable power we may assume that $\mathrm{ord}(\gamma_i) = s_i$. Then $\alpha = \gamma_1 \cdots \gamma_k$ has order $s_1 \cdots s_k = u$ by thelemma. Now by the construction of u any $\beta \in F$ has order dividing u. So the non-zero elements of F are all roots of $x^u - 1$. Hence $q - 1 \leq u$. But each s_i is a factor of $q - 1$ (by Lemma 1a), and they are powers of distinct primes. Therefore u divides $q - 1$. Thus they are equal.

look 15

□

4/9

SOLUTION 4

ALGORITHM Euclid's Algorithm.

Step 1. Set up a table with 4 columns (5 if you count the row number) headed Q, R, U, V. Fill in the first two rows (numbers -1 and 0) as follows:

ROW	Q	R	U	V
-1	-	a	1	0
0	-	b	0	1

Step 2. Calculate the Q and R entries for row 1 by dividing a by b: $a = q_1 b + r_1$. The U entry is 1 and the V entry is $-q_1$.

1	q_1	r_1	$u_1 = 1$	$v_1 = -q_1$

Note that $r_1 = 1a + (-q_1)b$.

Step 3. Suppose we have calculated up to row k and the last two rows are as follows

$k-1$	q_{k-1}	r_{k-1}	u_{k-1}	v_{k-1}
k	q_k	r_k	u_k	v_k

If $r_k = 0$ *Stop.*
Otherwise divide r_{k-1} by r_k: $r_{k-1} = q_{k+1} r_k + r_{k+1}$.
That gives the Q and R entries of row $k+1$.
Using the Q entry just calculated, put $u_{k+1} = u_{k-1} - q_{k+1} u_k$ and $v_{k+1} = v_{k-1} - q_{k+1} v_k$.

book 10

PROPOSITION. *The last non-zero element of the R-column is a highest common factor of a and b.*

Proof. Let the last non-zero element be r_n. Then $r_{n-1} = q_{n+1} r_n + 0$. So $r_n \mid r_{n-1}$. Next, $r_{n-2} = q_n r_{n-1} + r_n$. Since r_n divides both summands on the right hand side it divides r_{n-2}. Now suppose we have shown r_n divides r_{k+1} and r_k. As $r_{k-1} = q_{k+1} r_k + r_{k+1}$, it follows in the same way that r_n divides r_{k-1}. Finally, r_n divides $r_0 = b$ and $r_{-1} = a$. So r_n is a common factor of a and b.

book 5

Conversely, it will follow from (b) that if $c \mid a$ and $c \mid b$, then $c \mid r_n = u_n a + v_n b$. So r_n is indeed (a, b). □

5/9

Since $t(x)$ is irreducible and cannot divide $s(x)$, the highest common factor of the two polynomials must be a non-zero constant, and the only binary non-zero constant is 1. So the last non-zero entry in the R column must be 1. Now for each row of the table the entries satisfy $r_k = au_k + bv_k$. Therefore in the last row we have $1 = t(x)u_k(x) + s(x)v_k(x)$. That implies that $v_k(x)s(x) \equiv 1 \mod t(x)$. So v_k is the required polynomial $r(x)$.

seen

We execute Euclid's algorithm with the two given polynomials below

Q	R	U	V
	1 0 0 0 0 0 1 1	1	0
	1 0 1 1 0 1 1	0	1
1 0	1 1 0 1 0 1	1	1 0
1 1	1 0 0	1 1	1 1 1
1 1 0 1	1	1 0 1 1 0	1 0 0 0 0 1

The final row shows that $(x^5 + 1)(x^6 + x^4 + x^3 + x + 1) \equiv 1 \mod x^7 + x + 1$. So $x^5 + 1$ is the required polynomial.

seen

6/9

SOLUTION 5

We first consider the columns of the check matrix of Ham(k) as the binary representation of the integers $1, \ldots, 2^k - 1$, which in turn we take to represent the non-zero elements of GF(2^k). Using this interpretation, the check matrix $V_{k,t}$ extends the column α to

$$\begin{pmatrix} \alpha \\ \alpha^2 \\ \vdots \\ \alpha^{2t} \end{pmatrix}.$$

the check matrix $H_{k,t}$ is obtained by deleting the rows corresponding to even powers of α.

seen 5

PROPOSITION. *$V_{k,t}$ is a check matrix for the same code as $H_{k,t}$.*

Proof. We need only show that a binary word u of length $2^k - 1$ satisfying $H_{k,t}u = \underline{0}$ also satisfies $V_{k,t}u = \underline{0}$. Write the Hamming check matrix as $\alpha_1, \ldots, \alpha_{2^k-1}$. Then the condition $H_{k,t}u = \underline{0}$ is equivalent to

$$\sum_{r=1}^{2^k-1} \alpha_r^s u_r = 0 \qquad \text{for all odd } s \leq 2t.$$

Since the field we are calculating in has characteristic 2 we have

$$\sum_{r=1}^{2^k-1} \alpha_r^{2s} u_r^2 = \left(\sum_{r=1}^{2^k-1} \alpha_r^s u_r \right)^2 = 0.$$

But since $u_r = 0, 1$ we have $u_r^2 = u_r$ so

$$\sum_{r=1}^{2^k-1} \alpha_r^{2s} u_r = 0,$$

unseen 10

proving that $V_{k,t}u = \underline{0}$.

The block length of BCH(k, t is the number of columns in either of its check matrices, $2^k - 1$. Since each binary row of a check matrix can only reduce the rank of a code by at most one using $H_{k,t}$ we see that the rank of BCH(k, t) is at least $2^k - 1 - kt$.

Noting that each $(2t \times 2t)$–submatrix of $V_{k,t}$ considered over GF(2^k) is a Vandermonde matrix with non-zero determinant, we see that no $2t$ columns of $V_{k,t}$ are linearly dependent, even when the are considered as binary vectors. It follows that the code has minimum distance at least $2t + 1$.

Writing the matrix over GF(16) it takes the form

$$(15\ 14\ 13\ 12\ 11\ 10\ 9\ 8\ 7\ 6\ 5\ 4\ 3\ 2\ 1)$$

Adding the cubes beneath it gives the matrix

$$\begin{pmatrix} 15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ 8 & 5 & 8 & 3 & 1 & 1 & 3 & 5 & 15 & 5 & 3 & 15 & 15 & 8 & 1 \end{pmatrix}.$$

In binary this is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

unseen 10

8/9

SOLUTION 6

Received Word:

1 10 3 4 4 2 2 13 15 3 6 7 1 1 0

Syndrome Calculation using Horner's Scheme (*the candidate need only calculate the last row*):

	1	10	3	4	4	2	2	13	15	3	6	7	1	1	0
2 :	1	8	10	9	15	5	8	4	7	13	5	13	2	5	10
4 :	1	14	9	11	3	14	8	6	14	9	9	8	10	2	8
8 :	1	2	10	2	13	14	15	8	0	3	7	13	13	13	12
9 :	1	3	1	13	5	4	13	12	7	14	5	1	8	6	4
11 :	1	1	8	10	5	14	4	10	14	5	10	6	9	4	7
15 :	1	5	2	3	12	9	8	8	10	1	9	13	5	0	0
7 :	1	13	11	7	8	8	8	7	3	10	2	9	12	14	1
14 :	1	4	9	7	5	6	13	4	5	7	7	6	14	3	11

unseen

The syndromes are the entries in the final column.

Syndrome Polynomial: 11 1 0 7 4 12 8 10

Euclid's Algorithm:

0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	11	1	0	7	4	12	8	10	0	0	0	0	0	0	0	0	0	1
0	0	0	10	0	0	10	0	4	3	5	6	11	0	0	0	0	0	1	0	0	0	10	0
0	0	0	0	11	0	0	11	4	0	2	15	5	1	0	0	0	0	1	0	0	0	10	11
0	0	0	1	0	0	0	5	0	5	11	9	9	10	0	0	0	1	0	0	0	10	11	1
0	0	0	0	9	0	0	0	15	5	0	3	15	3	0	0	0	1	9	0	0	10	7	4
0	0	0	12	0	0	0	0	10	0	15	4	8	1	0	0	12	8	1	0	5	15	8	11
0	0	0	0	9	0	0	0	0	6	15	6	2	3	0	0	12	1	15	0	5	3	5	4
0	0	0	14	0	0	0	0	0	9	15	6	4	3	0	7	14	13	9	4	11	14	13	4
0	0	0	0	15	0	0	0	0	0	12	15	3	11	0	7	5	2	10	4	10	6	12	10

The third column is optional. The last entry in the second column is the error evaluator (it is the first time the entry in this column drops below degree 4). The last entry in the final column is the error locator.

unseen

Verification of the zeros of the error locator and calculation of derivatives at the zeros:

	4	10	6	12	10
6 :	4	11	14	3	0
	4	10	0	3	
11 :	4	13	4	11	0
	4	10	5	7	
12 :	4	8	2	13	0
	4	10	7	2	
15 :	4	4	8	9	0
	4	10	10	11	

The entries in the last column are the values of the error locator and as expected they are all zero. The final entries of the shorter rows are the values of derivative of the error locator (which could also be found by differentiating and calculating directly). Roots: 6 11 12 15. Corresponding error locations: 2 10 1 9

Error Evaluator:

	12	15	3	11
6 :	12	12	0	11
11 :	12	6	11	1
12 :	12	9	11	2
15 :	12	4	13	15

Error Values (error evaluator/derivative of error locator): 14 14 1 2

Corrected Word:

1 10 3 4 10 0 2 13 15 3 6 7 15 0 0.