Practice problems for the exam

Coding Technology

Problem 1

A binary error correction code has codewords

$$000010, \quad 110011, \quad 101001, \quad 111100.$$

(a) What are the parameters of the code?

(b) Compute the error detection and error correction capabilities of the code.

Problem 1

A binary error correction code has codewords

$$000010, \qquad 110011, \qquad 101001, \qquad 111100.$$

(a) What are the parameters of the code?

(b) Compute the error detection and error correction capabilities of the code.

Solution.

(a) Codewords have length 6, so $n = 6$. The number of codewords is $2^k = 4$, so $k = 2$. This is a C(6,2) code.

Problem 1

A binary error correction code has codewords

$$000010, \qquad 110011, \qquad 101001, \qquad 111100.$$

(a) What are the parameters of the code?

(b) Compute the error detection and error correction capabilities of the code.

Solution.

(a) Codewords have length 6, so $n = 6$. The number of codewords is $2^k = 4$, so $k = 2$. This is a C(6,2) code.

(b) Based on pairwise comparison, the minimal Hamming-distance among codewords is $d_{\min} = 3$:

000010	000010	000010	110001	110011	101001
110011	101001	111100	101011	111100	111100

The code can detect $d_{\min} - 1 = 2$ errors and correct $\lfloor \frac{d_{\min}-1}{2} \rfloor = 1$ error.

Problem 2

A binary linear error-correcting code has parity check matrix

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Is the code systematic?

(b) Determine the generator matrix G.

(c) Is this a Hamming-code?

(d) List all codewords.

(e) How many errors can the code correct?

(f) Compute the syndrome vector and the error group for the error vector $e = (10100)$. What is the group leader?

Problem 2

Solution.

(a) The code is systematic, because the rightmost 3×3 block of H is the identity matrix:

$$H = \left[\begin{array}{cc|ccc|} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right].$$

Problem 2

Solution.

(a) The code is systematic, because the rightmost 3×3 block of H is the identity matrix:

$$H = \left[\begin{array}{cc|ccc} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{array}\right].$$

(b) For systematic linear codes,

$$G = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

where the leftmost 2×2 block of G is the identity matrix, and the rightmost 2×3 block is the transpose of the leftmost 3×2 block of H.

Problem 2

(c) For Hamming codes, $n+1 = 2^{n-k}$ needs to hold, but based on the size of G, $n = 5$ and $k = 2$, for which $n+1 = 2^{n-k}$ does not hold: $n+1 = 6 \neq 8 = 2^{n-k}$. This is not a Hamming code.

Problem 2

(c) For Hamming codes, $n+1 = 2^{n-k}$ needs to hold, but based on the size of G, $n = 5$ and $k = 2$, for which $n+1 = 2^{n-k}$ does not hold: $n+1 = 6 \neq 8 = 2^{n-k}$. This is not a Hamming code.

(d) For linear codes, the codewords are all linear combinations of the rows of G:

$$(00000), \qquad (10101), \qquad (01111), \qquad (11010).$$

Problem 2

(c) For Hamming codes, $n + 1 = 2^{n-k}$ needs to hold, but based on the size of G, $n = 5$ and $k = 2$, for which $n + 1 = 2^{n-k}$ does not hold: $n + 1 = 6 \neq 8 = 2^{n-k}$. This is not a Hamming code.

(d) For linear codes, the codewords are all linear combinations of the rows of G:

$$(00000), \qquad (10101), \qquad (01111), \qquad (11010).$$

(e) Since this is a linear code,

$$d_{\min} = \min_{0 \neq c \text{ codeword}} w(c) = 3,$$

and the code can correct $\lfloor \frac{3-1}{2} \rfloor = 1$ error.

Problem 2

(f) The syndrome vector corresponding to the error vector $e = (10100)$ is

$$s^T = He^T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Problem 2

(f) The syndrome vector corresponding to the error vector $e = (10100)$ is

$$s^T = He^T = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

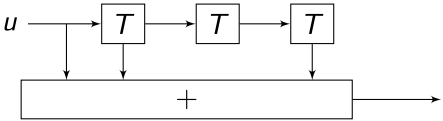
The error group can be obtained by adding each codeword to e:

$$\{(10100), (00001), (11011), (01100)\}.$$

The group leader is the vector with minimal weight: (00001).

Problem 3

Consider GF(8) with the irreducible polynomial $p(y) = y^3 + y + 1$. The following shift register architecture defines a linear code over the GF(8):



(a) What is the generator polynomial $g(x)$ of the code?

(b) What are the parameters of the code?

(c) What are the error correction capabilities of the code?

Problem 3

Solution.

(a) The given shift register architecture is multiplication by a polynomial. It contains no constant factors, so all coefficients of the polynomial are either 0 or 1, depending on whether the corresponding coefficient is included in the sum or not. Altogether, the architecture is multiplication by the polynomial $g(x) = x^3 + x + 1$m, so the generating polynomial of the code is $x^3 + x + 1$.

Problem 3

Solution.

(a) The given shift register architecture is multiplication by a polynomial. It contains no constant factors, so all coefficients of the polynomial are either 0 or 1, depending on whether the corresponding coefficient is included in the sum or not. Altogether, the architecture is multiplication by the polynomial $g(x) = x^3 + x + 1$m, so the generating polynomial of the code is $x^3 + x + 1$.

(b) Over GF(8), generator polynomials with 0-1 coefficients are typical for BCH codes. Let's check whether this polynomial is the generator polynomial of a BCH code.

Problem 3

Solution.

(a) The given shift register architecture is multiplication by a polynomial. It contains no constant factors, so all coefficients of the polynomial are either 0 or 1, depending on whether the corresponding coefficient is included in the sum or not. Altogether, the architecture is multiplication by the polynomial $g(x) = x^3 + x + 1$m, so the generating polynomial of the code is $x^3 + x + 1$.

(b) Over GF(8), generator polynomials with 0-1 coefficients are typical for BCH codes. Let's check whether this polynomial is the generator polynomial of a BCH code.

The conjugate groups and minimal polynomials over GF(8) are

$$\{1\} \to x - 1$$
$$\{y, y^2, y^4\} \to x^3 + x + 1$$
$$\{y^3, y^5, y^6\} \to x^3 + x^2 + 1$$

Problem 3

Solution.

(b) We need to check whether the generator polynomial is a product of some of the minimal polynomials Actually, $g(x)$ is equal to the minimal polynomial of the group $\{y, y^2, y^4\}$, so yes.

Problem 3

Solution.

(b) We need to check whether the generator polynomial is a product of some of the minimal polynomials Actually, $g(x)$ is equal to the minimal polynomial of the group $\{y, y^2, y^4\}$, so yes.

So this is a BCH code; the parameters are $n = 8 - 1 = 7$ and the degree of the generator polynomial is $n - k = 3$, so $k = 4$, and this is a C(7,4) code.

Problem 3

Solution.

(b) We need to check whether the generator polynomial is a product of some of the minimal polynomials Actually, $g(x)$ is equal to the minimal polynomial of the group $\{y, y^2, y^4\}$, so yes.

So this is a BCH code; the parameters are $n = 8 - 1 = 7$ and the degree of the generator polynomial is $n - k = 3$, so $k = 4$, and this is a C(7,4) code.

(c) The roots of $g(x)$ contain y^1 and y^2 (along with their entire conjugate group), but not y^3, so this code can correct $t = 1$ error.

Problem 4

Give the generator matrix of a code over GF(7) that can correct two errors.

Problem 4

Give the generator matrix of a code over GF(7) that can correct two errors.

Solution. We will use Reed–Solomon code. First we compute the parameters n and k. For any RS code over GF(7), $n = 7 - 1 = 6$. A C(n,k) RS code can correct $\lfloor \frac{n-k}{2} \rfloor$ errors;

$$\left\lfloor \frac{6-k}{2} \right\rfloor \qquad \rightarrow \qquad k = 2.$$

Problem 4

Give the generator matrix of a code over GF(7) that can correct two errors.

Solution. We will use Reed–Solomon code. First we compute the parameters n and k. For any RS code over GF(7), $n = 7 - 1 = 6$. A C(n,k) RS code can correct $\lfloor \frac{n-k}{2} \rfloor$ errors;

$$\left\lfloor \frac{6-k}{2} \right\rfloor \qquad \to \qquad k = 2.$$

So we will give a C(6,2) code. We need a primitive element over GF(7). There are two primitive elements over GF(7): 3 and 5; we can choose either one. E.g. picking 3 gives

$$G = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 3 & 2 & 6 & 4 & 5 \end{bmatrix}.$$

Problem 5

A source has the following distribution and coding:

Source symbol	Probability	Codeword
A	0.29	0
B	0.21	10
C	0.20	110
D	0.19	1110
E	0.11	1111

(a) Is the code prefix-free?

(b) Compute the average codelength.

(c) What is the theoretical lower bound for data compression?

(d) Is the code optimal?

Problem 5

Solution.

(a) Yes, it is prefix-free.

Problem 5

Solution.

(a) Yes, it is prefix-free.

(b) The average codelength is

$$0.29 \cdot 1 + 0.21 \cdot 2 + 0.20 \cdot 3 + 0.19 \cdot 4 + 0.11 \cdot 4 = 2.51.$$

Problem 5

Solution.

(a) Yes, it is prefix-free.

(b) The average codelength is

$$0.29 \cdot 1 + 0.21 \cdot 2 + 0.20 \cdot 3 + 0.19 \cdot 4 + 0.11 \cdot 4 = 2.51.$$

(c) The theoretical lower bound for data compression is the entropy:

$$\sum_{i=1}^{5} -p_i \log_2(p_i) = 2.2606.$$

Problem 5

Solution.

(a) Yes, it is prefix-free.

(b) The average codelength is

$$0.29 \cdot 1 + 0.21 \cdot 2 + 0.20 \cdot 3 + 0.19 \cdot 4 + 0.11 \cdot 4 = 2.51.$$

(c) The theoretical lower bound for data compression is the entropy:

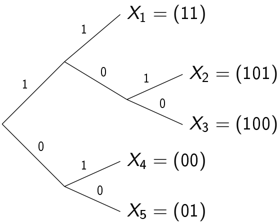
$$\sum_{i=1}^{5} -p_i \log_2(p_i) = 2.2606.$$

(d) Optimality can be checked by comparing it with the Huffman code.

Problem 5

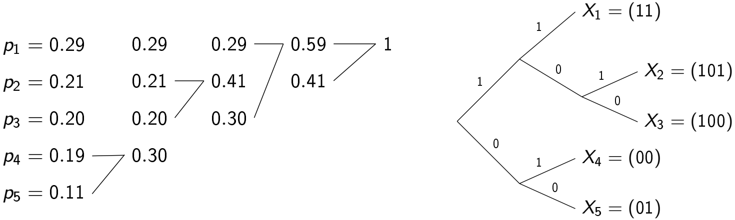
(d)

$p_1 = 0.29$	0.29	0.29	0.59	1
$p_2 = 0.21$	0.21	0.41	0.41	
$p_3 = 0.20$	0.20	0.30		
$p_4 = 0.19$	0.30			
$p_5 = 0.11$				



Problem 5

(d)



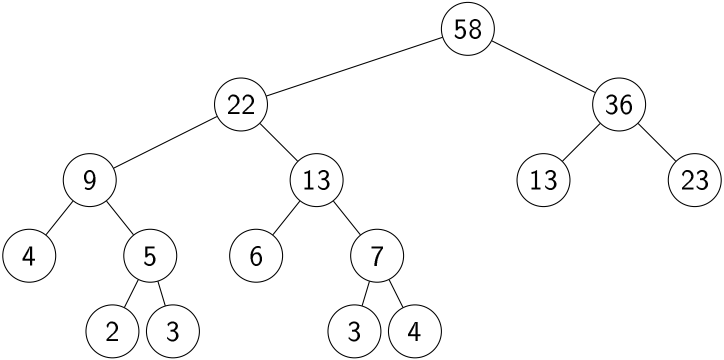
Average codeword length is

$$0.29 \cdot 2 + 0.21 \cdot 2 + 0.20 \cdot 2 + 0.19 \cdot 3 + 0.11 \cdot 3 = 2.30,$$

which is smaller than for the original coding, so the original code is not optimal.

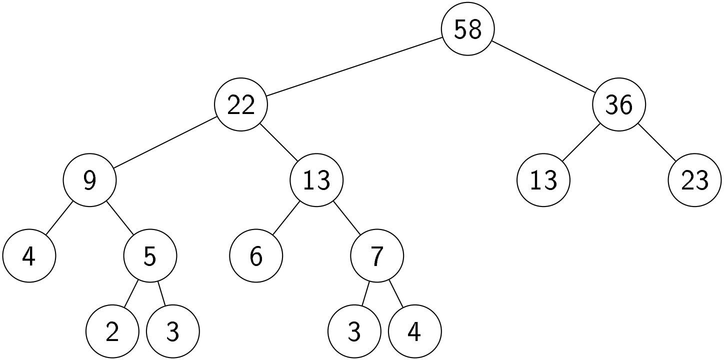
Problem 6

Does the sibling property hold for the following tree?



Problem 6

Does the sibling property hold for the following tree?



Solution. Reading the weights in the tree from bottom row up, left to right: 2, 3, 3, 4, 4, 5, 6, 7, 9, 13, 13, 23, 22, 36, 58. There is a 23 followed by 22, so the sequence is not increasing, the sibling property does not hold.

Problem 7

We use LZ77 to compress the sequence

a p r o b a b a b a b e z u r j a a b a z u r t

Parameter values are $h_s = 6, h_l = 6$. The cursor is initially between characters 6 and 7. Compute the output of the first two steps of the LZ77 algorithm.

Problem 7

We use LZ77 to compress the sequence

$$\text{a p r o b a b a b a b e z u r j a a b a z u r t}$$

Parameter values are $h_s = 6$, $h_l = 6$. The cursor is initially between characters 6 and 7. Compute the output of the first two steps of the LZ77 algorithm.

Solution.

- initial state:

  2

  a p r o b a | b a b a b e | z u r j a a b a z u r t   output: $(2, 5, e)$

  5

Problem 7

We use LZ77 to compress the sequence

a p r o b a b a b a b e z u r j a a b a z u r t

Parameter values are $h_s = 6$, $h_l = 6$. The cursor is initially between characters 6 and 7. Compute the output of the first two steps of the LZ77 algorithm.

Solution.

- initial state:

  $\overset{}{\boxed{\text{a p r o }\overset{2}{\text{b a}}}}\boxed{\underbrace{\text{b a b a b}}_{5}\text{ e}}$ z u r j a a b a z u r t   output: $(2, 5, e)$

- next step:

  a p r o b a $\boxed{\text{b a b a b e}}\boxed{\text{z u r j a a}}$ b a z u r t   output: $(0, 0, z)$

Problem 8

We use RSA encryption with $p = 5, q = 7$.

(a) What is the public key (n, e) if we make the smallest possible choice for e?

(b) Compute the private key d corresponding to e.

(c) Encrypt the text $x = 3$.

Problem 8

Solution.

(a) $n = pq = 35, \quad \phi(n) = (p-1)(q-1) = 24.$

$$\gcd(\phi(n), e) = 1 \qquad \text{és} \qquad 1 < e < \phi(n) = 24$$

needs to hold for e, the smallest possible choice is $e = 5$. The public key is (35,5).

Problem 8

Solution.

(a) $n = pq = 35, \quad \phi(n) = (p-1)(q-1) = 24.$

$$\gcd(\phi(n), e) = 1 \qquad \text{és} \qquad 1 < e < \phi(n) = 24$$

needs to hold for e, the smallest possible choice is $e = 5$. The public key is (35,5).

(b) For $e = 5$, the private key d is the solution of $de = 1 \mod \phi(n)$, that is,

$$5d = 1 \mod 24$$

whose solution is

Problem 8

Solution.

(a) $n = pq = 35, \quad \phi(n) = (p-1)(q-1) = 24.$

$$\gcd(\phi(n), e) = 1 \qquad \text{és} \qquad 1 < e < \phi(n) = 24$$

needs to hold for e, the smallest possible choice is $e = 5$. The public key is (35,5).

(b) For $e = 5$, the private key d is the solution of $de = 1 \mod \phi(n)$, that is,

$$5d = 1 \mod 24$$

whose solution is $d = 5$.

Problem 8

Solution.

(a) $n = pq = 35, \quad \phi(n) = (p-1)(q-1) = 24.$

$$\gcd(\phi(n), e) = 1 \qquad \text{és} \qquad 1 < e < \phi(n) = 24$$

needs to hold for e, the smallest possible choice is $e = 5$. The public key is (35,5).

(b) For $e = 5$, the private key d is the solution of $de = 1 \mod \phi(n)$, that is,

$$5d = 1 \mod 24$$

whose solution is $d = 5$.

(c) The text $x = 3$ is encrypted as

$$c = x^e = 3^5 = 243 \mod 35 = 33.$$