TUHH
Technische Universität Hamburg

INSTITUT FÜR
NACHRICHTENTECHNIK
PROF. DR.-ING. GERHARD BAUCH

Exam WS 2017/2018

Information Theory and Coding

Name: Sample solution	Student ID:	
	Points	From
Task 1		
Task 2		
Total points		
Grade		

- The following aids are allowed in this exam:
  - 2 DIN A4 sheets, **handwritten** on both sides (4 pages in total)
  - Calculator (non-programmable, not graphical, not capable of communication)
  - Pens
- Other aids are not allowed.
- Please use a separate solution sheet for each task.
- Write your name and matriculation number on each solution sheet.
- An arrow next to a question means that this part of the task can be solved independently of the rest of the task.
- For calculations the approach as well as the steps must be specified.
- Please do not write with pencils and do not use a red pen.
- The duration of the exam is 90 minutes.
- The exam consists of 9 pages (including this cover page).
- Switch off your cell phones!

Task 1: Ternary Huffman Code

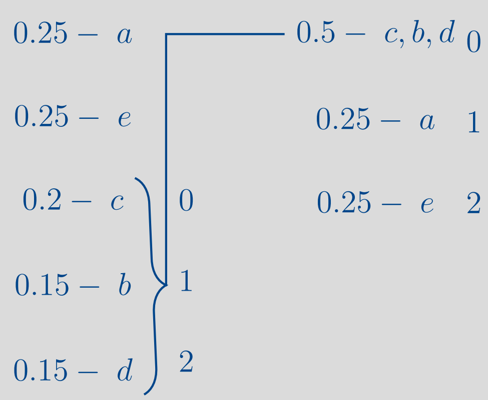
The symbols of a source are described by the random variable X with realizations $x_i \in \{a, b, c, d, e\}$.
The realizations x_i occur according to the following probabilities:

x_i	$p_X(x_i)$
a	0.25
b	0.15
c	0.2
d	0.15
e	0.25

Now, the source symbols shall be encoded by a ternary Huffman code, i.e. the code symbols of the Huffman code take values in the set $c_i \in \{0, 1, 2\}$.

a) Construct a ternary Huffman-Code for encoding of the realizations x_i. Make sure that all steps in the construction of the Huffman code are clearly given. Represent the mapping of realizations x_i to codewords $\mathbf{c}$ of the Huffman code in a table. Answer in complete sentences!

At first all possible realizations are sorted according to its probabilities in descending order. This ordering is not unique since realizations with the same probability can be sorted arbitrarily The realizations with the 3 smallest probabilities are combined and represented by one hyper symbol. We code the 3 most unlikely realizations with a code symbol $c_i \in \{0, 1, 2\}$. In a second step again the remaining symbols and the hyper symbol are sorted according to its probabilities in descending order. No further combining is needed since only 3 distinct symbols exist. Hence, we code these 3 symbols again with a code symbol $c_i \in \{0, 1, 2\}$.

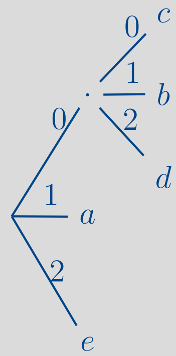


To obtain the codewords $\mathbf{c}$ the code symbol c_i assigned to a realization or a hyper symbol which contains this realizations is noted from right to left.

x_i	$\mathbf{c}$
a	1
b	01
c	00
d	02
e	2

b) Sketch the tree representation of your Huffman code from a).

Based on the codewords from a) the tree can be constructed directly. Starting from the root, a branch for each code symbol is added and labeled. The leafs at the first level are either the realizations itself (if the codeword length is one) or the hyper symbol. For hyper symbols again three branches are created and labeled. The leafs correspond to realizations or hyper symbols and so on.



c) Determine the average codeword length of the Huffman code from a).

The average codeword length is found by summing up all individual codeword lengths weighted by the probability that this codeword occurs, which is equivalent to the probability that the realization x_i is drawn which is mapped onto this particular codeword.

$$E\{L\} = \sum_{\mathbf{c}} P(\mathbf{c})L(\mathbf{c}) = 0.15 \cdot 2 + 0.15 \cdot 2 + 0.2 \cdot 2 + 0.25 \cdot 1 + 0.25 \cdot 1 = 1.5$$

d) Is the code prefix-free? Give reasons and answer in complete sentences!

A code is called a prefix code or a prefix-free code or an instantaneous code, if no codeword is the prefix of another codeword. This is true for the given code. Thus, this code is prefix-free.

Task 2: Polar Codes

Polar Codes are currently discussed for many applications such as 5G cellular systems as they have been shown to be capacity-achieving in binary memoryless channels. The idea of Polar Codes is that the encoder in combination with a successive cancellation decoder transforms the channel into a set of equivalent channels for the information bits, where a fraction of those channels is error-free (capacity $C_i = 1$), whereas the remaining channels are totally noisy (capacity $C_i = 0$). This effect of transforming the physical channel into perfect and noisy (good and bad) channels is called polarization. Information is transmitted only on the perfect channels. We will investigate the principle of Polar Codes in this task.
We consider Polar Codes in a Binary Erasure Channel (BEC). The capacity of a BEC with erasure probability p_e is given by

$$C_{\text{BEC}} = 1 - p_e$$

The basic building block of a Polar Code is depicted in the following figure:

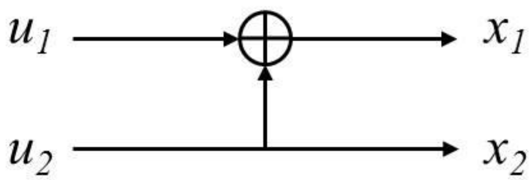
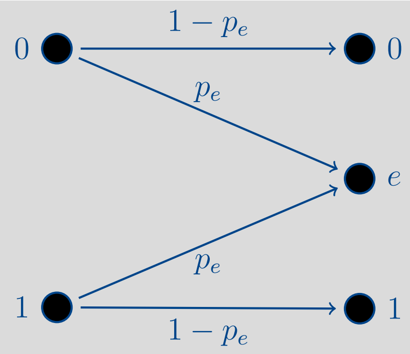


Figure 1: Basic building block of a Polar Code encoder.

Two information bits u_1 and u_2 are mapped to two code bits x_1 and x_2. The code bits x_i are then transmitted through a Binary Erasure Channel (BEC).
The received bits at the output of the BEC are denoted y_i.
The information bits are uniformly distributed, i.e. $P(u_i = 0) = P(u_i = 1) = 1/2$.

a) Sketch the state transition diagram of a Binary Erasure Channel. Label the figure completely.

A binary erasure channel is characterized by two input states, i.e. $+1, -1$ and three output states, i.e. $+1, e, -1$. For a binary erasure channel the transmission can either be correct or an erasure. Hence, no transitions from $-1 \to +1$ or from $+1 \to -1$ are possible. An erasure occurs with probability p_e and the transmission is correct with probability $1 - p_e$. These considerations lead to the state transition diagram depicted below where the arrows are labeled by the probabilities p_e and $1 - p_e$, respectively.



b) State the definition of a memoryless channel and show, that the Binary Erasure Channel is memoryless. Answer in complete sentences!

A channel is called to be memoryless if

$$p(\mathbf{y}|\mathbf{x}) = \prod_i p(y_i|x_i).$$

Thus, the BEC is memoryless.

c) Derive the channel capacity $C_{2\text{BEC}}$ of two parallel Binary Erasure Channels with the same erasure probability p_e.

For two parallel channels the overall capacity is the sum of the capacity of each individual channel. Thus,

$$C_{2\text{BEC}} = 2 \cdot C_{\text{BEC}} = 2(1-p_e)$$

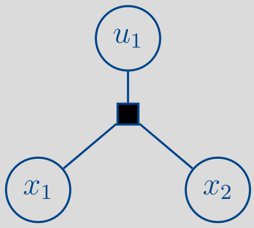
d) Derive the equations for the code bits x_1 and x_2 depending on u_1 and u_2.

As described in the task description and visualized in Figure 1 the code bits x_1 and x_2 are obtained by rather simple algebraic operations. x_1 is just the modulo-2 sum of u_1 and u_2 and x_2 is an unchanged copy of u_2, i.e.

$$x_1 = u_1 \oplus u_2$$
$$x_2 = u_2$$

e) Sketch a Tanner graph element for the information bit u_1, i.e. a Tanner graph which contains u_1,x_1 and x_2 as variable nodes. Which type of code is represented by this Tanner graph element?

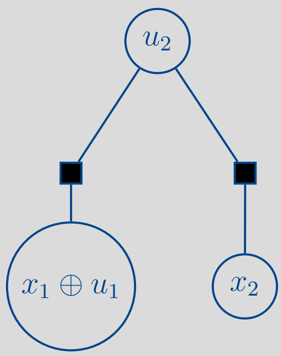
A Tanner graph is a bipartite graph with one set of nodes which holds the so-called variable nodes and a second set holding the so-called check nodes. Variable nodes are visualized as circles with the name of the respective variable inside. Check nodes instead are represented as squares.



Since only one check node exists which is connected to all variable nodes involved, this corresponds to exactly one parity check equation holding x_1, x_2 and u_1, i.e $x_1 \oplus x_2 \oplus u_1 = 0$. Hence, this graph represents a single parity check code.

f) Sketch a Tanner graph element for the information bit u_2. Assume that the bit u_1 is perfectly known due to prior decoding. Which type of code is represented by this Tanner graph element ?

A Tanner graph is a bipartite graph with one set of nodes which holds the so-called variable nodes and a second set holding the so-called check nodes. Variable nodes are visualized as circles with the name of the respective variable inside. Check nodes instead are represented as squares.



Since only x_2 and u_2 are connected to the right check node this indicates that $x_2 = u_2$ and the check node serves rather as an "equality" node than a check node, i.e. $x_2 \oplus u_2 = 0$. The same holds for the left side where due to the perfect knowledge of u_1 no additional variable node is added for u_1. Thus, also the second check node is basically an "equality" node. Hence, this graph represents a repetition code, i.e. $x_1 \oplus u_1$ and x_2 are equivalent to u_2.

The basic decoder of a Polar Code is a successive cancellation decoder: First, the information bit u_1 is decoded. Then, the information bit u_2 is decoded under the assumption, that the

decision on u_1 was correct.
Thus, for each information bit an equivalent channel with a corresponding capacity can be derived.
The channel capacity C_1 for an equivalent channel for the information bit u_1 is given as

$$C_1 = (1 - p_e)^2.$$

The channel capacity C_2 for an equivalent channel for the information bit u_2 is given as

$$C_2 = 1 - p_e^2.$$

g) Show that $C_2 > C_1$. Give an intuitive explanation for this relation. Answer in complete sentences!

Using the hints from the task description in the first step to show that the inequality holds, the difference $C_1 - C_2$ is shown to be smaller than 0.

$$\begin{aligned} C_1 &= (1 - p_e)^2 = 1 - 2p_e + p_e^2 \\ C_2 &= 1 - p_e^2 \\ C_1 - C_2 &= 1 - 2p_e + p_e^2 - (1 - p_e^2) \\ &= \underbrace{2p_e}_{>0} \underbrace{(p_e - 1)}_{<0} < 0 \,\text{for}\, p_e > 0 \end{aligned}$$

Thus, $C_2 > C_1$ for $p_e > 0$.

Intuitively, u_2 is contained in both code bits x_1 and x_2, while u_1 is contained in x_1 only. Hence, u_2 is better protected (under the assumption that u_1 is known).

h) Derive the sum capacity $C_1 + C_2$ based on the given channel capacities for the equivalent channels. Compare the sum capacity to your result from c).

$$\begin{aligned} C_1 + C_2 &= (1 - p_e)^2 + (1 - p_e^2) \\ &= 1 - 2p_e + p_e^2 + (1 - p_e^2) \\ &= 2(p_e - 1) \\ &= 2C_{BEC} \\ &= C_{2BEC} \end{aligned}$$

i) Interprete your results from g) and h) in view of the target of *polarization* as explained in the introduction to the task. Answer in complete sentences!

The result from h) shows, that the total capacity is not changed by the polar encoder. However, due to encoding and successive cancellation decoding, u_2 faces a better channel than u_1. Hence, already for the small basic building block of a polar code, a polarization tendency is achieved.

j) Derive a generator matrix $\mathbf{G}_2$ for the basic encoder building block in Figure 1.

In general the encoding in vector-matrix form using the generator matrix $\mathbf{G}_2$ can be written as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{G}_2 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

In Figure 1 we see that $x_1 = u_1 \oplus u_2$ and $x_2 = u_2$. To achieve this using vector-matrix multiplication $\mathbf{G}_2$ has to equal

$$\mathbf{G}_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

k) Derive the code rate R_2 of the basic building block in Figure 1. Can error correction be achieved with this code? Give reasons and answer in complete sentences!

In Figure 1 two information bits, i.e. $K = 2$ are mapped onto two code bits, i.e. $N = 2$. Thus the rate R_2 can be found as

$$R_2 = \frac{K}{N} = \frac{2}{2} = 1$$

Consequently, the code does not add redundancy and, hence, no errors can be corrected.

A powerful Polar Code with strong polarization effect can be obtained by combining a large number of copies of the basic building block from Figure 1. The procedure is illustrated for a Polar Code of codeword length $N = 4$ in the following figure:

l) Derive a generator matrix $\mathbf{G}_4$ for the encoder in Figure 2.

In general the encoding in vector-matrix form using the generator matrix $\mathbf{G}_4$ can be written as

$$\mathbf{x} = \mathbf{G}_4\mathbf{u}$$

For each information bit u_i which influences the value of a certain x_j indicated by a connection in Figure 2 a one at position $G_{j,i}$ is added. Otherwise, if no connection exist a zero is added in $\mathbf{G}_4$ at the correct position.

$$\mathbf{G}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

m) Write the generator matrix $\mathbf{G}_4$ depending on the generator matrix $\mathbf{G}_2$ from j).

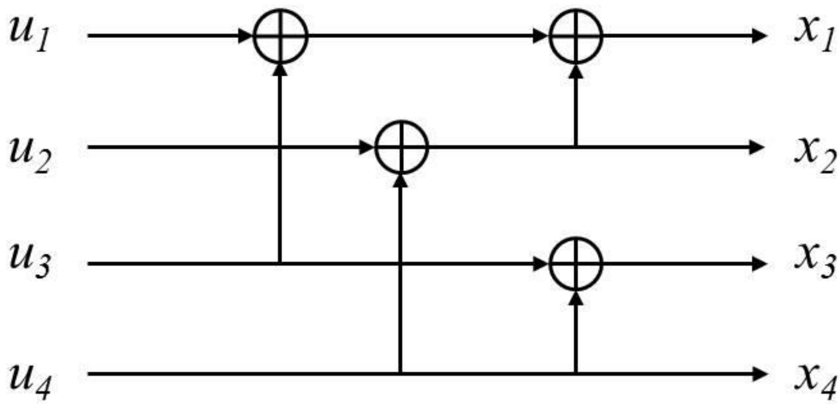


Figure 2: Encoder of Polar Code with codeword length $N = 4$.

$$\mathbf{x} = \mathbf{G}_4 \mathbf{u}$$

$$\mathbf{G}_4 = \begin{bmatrix} \mathbf{G}_2 & \mathbf{G}_2 \\ \mathbf{0} & \mathbf{G}_2 \end{bmatrix}$$

In order to obtain an error-correcting code with code rate R<1, some of the bits u_i are not used for transmission of information. Instead, they are replaced by fixed values, e.g. by 0. We call the fixed bits *frozen bits*. The polarization effect is exploited for choosing the positions of the frozen bits: Information is transmitted on the good effective channels, while the frozen bits are put to the bad (noisy) effective channels.

n) How many bits u_i need to be frozen in the encoder from Figure 2 in order to obtain a code rate of $R = 1/2$?

Since 4 code bits x_1, x_2, x_3, x_4 exist if $N = 4$. For a rate $R = 1/2$, $N \cdot R = K = 2$ bits are information bits. Hence, 2 out of the 4 u_i need to be frozen, i.e. they do not carry information.