WOLFGANG K. SEILER
Tel. 2515

B6, B4.08

October 5, 2023

Mathematics and Information, Exercise sheet 4

Problem 1: *(5 points)*

Let $A = \{a, b, c, d, e, f, g, h\}$ with $p(a) = p(b) = p(c) = \frac{1}{15}$, $p(d) = p(e) = p(f) = \frac{2}{15}$, and $p(g) = p(h) = \frac{1}{5}$.

a) What is the entropy of this source?

b) Construct a binary HUFFMAN code for it!

c) What is the average length of a code word?

d) Find a ternary HUFFMAN code and compare the averave code length to the entropy to base three!

Problem 2: *(10 points)*

Based on a letter count of JEAN PAUL's *Dr. Katzenbeisers Badereise,* the frequencies of letters in german plain text are as follows:

E	N	I	R	S	A	T	H	D	U	L	C	G
0,185	0,103	0,0735	0,0695	0,0681	0,0575	0,0546	0,0525	0,0481	0,0435	0,0369	0,0327	0,0279
O	M	B	W	F	K	Z	V	P	J	Y	X	Q
0,0275	0,0265	0,0203	0,0162	0,0158	0,0129	0,0123	0,00759	0,00567	0,00191	0,000497	0,000154	0,000108

Construct a binary HUFFMAN code for this alphabet!

Problem 3: *(4 points)*

Let A be an alphabet consisting of n letters occuring with equal probability each.

a) Construct a binary HUFFMAN code for $n = 3, 5, 6, 7$, and compare its average length with the entropy!

b) Determine the average length of a binary HUFFMAN code for $n = 2^k + 1$ and $n = 2^k - 1$ for $k \geq 2$!

c) ***True of false?*** If $C: A \to \mathcal{D}$ is a code with D different symbols, and if $\sum_{a \in A} D^{-\ell(a)} = 1$, then C is optimal.

Problem 4: *(1 points)*

Transmitting a message usually involves three coding steps:

- *Source coding* in order to adapt the message to the medium and possibly also compressing it
- *Channel coding* uses error correcting codes to safeguard against transmission errors
- *Cryptographic codes* prevent intelligent adversaries from reading or manipulating the message.

In which order should those three steps be applied for best results?

Can be submittel till Wednesday, October 11, 2023 at 15.30h