This is your first graded homework. Your grade will solely be evaluated based on your answers, not your friends' answers, nor your look or gender ;). Try to write down your answers as clear as possible such that at least one of the assistants would be able to read them. :)

PROBLEM 1. Assume that we use the following code for encoding the source stream. Let $x = x_1 x_2 x_3, \ldots$ be a long sequence of source symbols where $x_i \in \{a, b, c, d\}$. Let $\mathbf{c} = c(x_1)c(x_2)c(x_3)\ldots$ be the corresponding long binary sequence which results when we encode x. Imagine picking one bit at random from c. What is the probability that this bit is a 1?

Symbol	Code	probability
a	0	1/2
b	10	1/4
c	110	1/8
d	111	1/8

PROBLEM 2. Find a probability distribution $\{p_1, p_2, p_3, p_3\}$ such that there are *two* optimal codes that assign different lengths l_i to the four symbols a, b, c, d.

Symbol	probability	length
a	p_1	l_1
b	p_2	l_2
c	p_3	l_3
d	p_4	l_4

PROBLEM 3. The World Series is a seven-game series that terminates as soon as either Nadal or Federe wins four games. Let X be the random variable that represents the outcome of a World Series between Nadal (N) and Federer (F). Possible values of X are FFFF, NFNFNFN, NNNFFFF, ... Let Y be the number of games played which ranges from 4 to 7. Assume that Nadal and Federer are equally matched and that the games are independent. Calculate $H(X)$ and $H(Y)$

PROBLEM 4. Which of these codes cannot be Huffman codes for any probability assignment?

- $\{0, 10, 11\}$.

- $\{00, 01, 10, 110\}$.

- $\{01, 10\}$.

PROBLEM 5. Consider the following source.

Symbol	probability
a	1/2
b	1/4
c	1/8
d	1/8

- What is the entropy of this source?

- Find a binary Huffman code. What is the average code length?

- Find a Shannon-Fano code. What is the average code length?

- What is the entropy of this source?

- Find a binary Huffman code. What is the average code length?