**Algorithm Engineering [midterm]**

11 November 2020 – time 45 minutes

**Question #1 [ranks 4+3+1].** Consider the Snow Plow technique with memory M=2.
- Simulate Snow Plow over the sequence $S = (1, 3, 9, 10, 7, 6, 5, 4, 3, 8)$.
- Provide an example of sequence of length 10 that generates exactly 5 runs, when the memory has size M=2.
- Show the average length of the run produced by Snow Plow, if we assume that the probability that an item goes to the Heap is $\frac{3}{4}$ rather than $\frac{1}{2}$.

**Question #2 [rank 3].** Given a Universal class of hash functions that map keys from a universe $U$ to the range $[0,2^2]$. Compute the average number of collisions induced by a function $h$ drawn randomly from that class and mapping a set of 10 keys.

**Question #3 [rank 2].** Prove that the probability of having a 0 in a position of the binary array of the Bloom Filter is $\frac{1}{2}$, when the number of hash functions is set to the optimal one.

**Question #4 [rank 4].** Show the first 8 codewords of the (s,c)-code with s=3 and c=1, hence s+c = 4 (briefly explain your calculations).

**Question #5 [rank 4].** Decompress the 8th integer encoded via Elias-Fano in the two arrays:
$L = 0111000101001111001100$ and $H = 110 110 10 0 10 10 10 0 0 0 0 0 0$
where the original encoding of the integers is in 6 bits. (*hint: derive first the number of keys, and then the length of the low and high part)*

**Question #6 [rank 4].** Show the binary succinct encoding of the tree $T = \{ a \rightarrow b \ (\text{left child}); \ b \rightarrow c \ (\text{left child}); \ b \rightarrow e \ (\text{right child}); \ c \rightarrow d \ (\text{right child}) \}$ of root “a”.

**Question #7 [rank 5].** Given the set of strings (aa, ba, bb, bc, ca) and you wish to construct a minimal ordered perfect hash function where rank(a,b,c) = (2, 3, 4) and

$$h_1(c' c'') = 2 \times \text{rank}(c') \times \text{rank}(c'') \mod 11,$$
$$h_2(c' c'') = 5 \times \text{rank}(c') + \text{rank}(c'') + 1 \mod 11.$$

Construct the final $h(t)$.