Algorithm Engineering – exercises
7 September 2023 – time 60 minutes

Name and Surname: #matricola:

Question #1 [scores 4] Simulate the algorithm SnowPlow over the sequence 2,6,5,3,1,7,2, and show which sorted blocks it forms with a memory of size M=2.

Question #2 [scores 5+5] Given the ordered set of strings:

$$S = \{ \text{AABA, AACAAAC, AACAACC, BABAA, BABBB, BACA} \}$$

- Build the Patricia trie PT for S.
- Show the steps executed to lexicographically search for the pattern P = AACBACD in the set S by means of PT.

Question #3 [scores 3+3] Given the sequence a,b,c,d,e,f,g,h,i,l simulate:
- The sampling algorithm for m=2 which knows the sequence length n=10, assuming probabilities for the parameter p = [0.5, 0.5, 0.5, 1, 1, 0.1, 0.5, 1, 0.1, 1]
- The sampling algorithm for m=2 which does not know the sequence length, assuming values for the parameter h = [1, 3, 4, 2, 1, 5, 4, 6]

Question #4 [scores 5]. Given the string "ABABABC" compress it by using the pipeline BWT + MTF + RLE0 + Huffman, where MTF counts letter's positions from 0, and RLE0 uses the Wheeler's code.

Question #5 [scores 5]. You are given the binary array B = [0 0 0 1 0 1 1 1 0 0 0 1].
Build the data structure that solves the Rank query in constant time, by assuming that the big block has size 4 and the small block has size 2.

Algorithm Engineering – theory
7 September 2023 – time 60 minutes

Name and Surname: #matricola:

Question #1 [scores 6] Show the pseudo-code of the MultiKey-Quicksort algorithm, state and prove its time complexity.

Question #2 [scores 6+6]
- Define what is a class of universal hash functions
- Provide an example and prove its universality

Question #3 [scores 4+4+3+3]
- Define formally what is the suffix array SA of a text string T[1,n]
- Define formally what is its corresponding LCP array.
- Describe how to search for a pattern P[1,p] in SA
- State and prove the time complexity of the search operation above