Algorithm Engineering -- EXERCISES
21 June 2024

Name and Surname: **#matricola:**

Question #1 [score 5]. Given the strings S = {abbc, abra, bach, acat}, show the structure of a Ternary Search Tree built by inserting them in that order.

Question #2 [score 5+5]. Given the set S of pairs {<D,4>, <A,6>, <H, 10>, <B,3>, <G,7>, <F,2>}, where the first component is the key and the second component is the priority.
- Build a TREAP data structure by inserting the pairs in that order (you can assume that it is a MIN heap).
- Show the execution of the SPLIT operation on the key E

Question #3 [score 5] Given the set of n=4 strings, each consisting of two digits:

$$S = \{11, 22, 33, 44\}$$

Build (if possible) the Minimal Ordered Perfect Hash for S using the following two hash functions:

$$h_1(xy) = x+5*y \mod 7 \quad \text{and} \quad h_2(xy) = x+3*y \mod 7$$

in which x (resp., y) is the first (resp., second) digit of a string of S.
As an example, if the string is 11, then x=1 and y=1.

Question #4 [scores 3+2+2+3] Given the text T = BRABRA, apply the pipeline BWT+MTF+RLE0 (with Wheeler's code for the 0-runs) and finally apply Arithmetic coding on the first 3 numbers of the output of this pipeline.

Algorithm Engineering -- THEORY
21 June 2024

Name and Surname: **#matricola:**

Question #1 [score 5+3]
- Prove that the expected length of an ordered sequence produced by the algorithm Snow Plow is 2M.
- What is that expected length if the probability for an item to go in the "unsorted bucket" is 1/8 instead of ½ ?

Question #2 [rank 5+5]. Given two sorted lists of integers, say L1 and L2 of lengths n and m respectively:
- Describe the "two-level scan" algorithm to compute their intersection.
- State and prove the time complexity of the previous point.

Question #3 [score 4+4+4].
- Describe why we introduced the Canonical Huffman encoding algorithm
- Specify which data structures it keeps in the preamble of its compressed file
- Write its pseudo-code to decompress one symbol