Algorithm Engineering – exercises
24 July 2023 – time 60 minutes

Name and Surname: #matricola:

Question #1 [scores 4+6+6] Given the string S = aabcbcbada, show the:

- LZ77 parsing of S.
- LZ78 parsing of S, along with the auxiliary data structure used to compute it.
- Canonical Huffman encoding of S.

Question #2 [scores 6] Given the two lists L_1 = (3, 5, 6, 8, 10, 13, 14, 16, 17) and L_2 = (5, 7, 8, 14, 15) compute their intersection using the:

- Two-level storage approach, with block size b = 3 for the list L_1.

Question #3 [scores 6] Given a set of 5 strings S = {AAB, ABA, BAA, BBA, BBB}, construct a Minimal Ordered Perfect Hash Function for S by assuming the two hash functions:

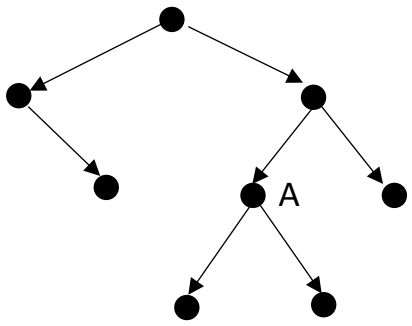
$h_1(xyz) = (x + y)*3 + z \bmod 11$

$h_2(xyz) = 2*x + 5*y + 2*z \bmod 11$

in which x, y, z are the codes of the first, second and third character of the argument to the hash function, where the character codes are A=1, and B=2.

As an example, the value of h_1 on ABB is $h_1(122) = [(1+2)*3 + 2] \bmod 11 = 0$

Question #4 [scores 4] Show the succinct representation of the following binary tree. Then show how to use this representation to navigate from the root to the node labelled A, and then back to the parent of A.



Algorithm Engineering – theory
24 July 2023 – time 60 minutes

Name and Surname: #matricola:

Question #1 [scores 4+3+3]

- Define what is the Suffix Array A built on a text T[1,n].
- Specify the pseudocode to construct A via qsort.
- State and comment the worst-case I/O-complexity of the construction of A via qsort.

Question #2 [scores 5+5]

Let us assume you are given a Patricia Trie built on a dictionary S of n strings.

- Describe how to use the Patricia Trie to search for the lexicographic position of a query pattern P[1,p] among the strings in S.
- State the time- and I/O-complexity of the above search algorithm.

Question #3 [scores 5+5]

- Define what is a Bloom Filter over a dictionary D of n keys, and discuss the operations it supports.
- State and prove the result on the false-positive error of a Bloom Filter.