

AD4DA – Final

17 December 2025 – time: 100 minutes

First and Last Name: Matricola:

Exercise 1 [2+2] Consider the LOUDS representation of a generic tree (ignoring the dummy root for simplicity).

1. Draw the tree corresponding to the binary string 11101110000011000 ,
2. Show a binary string with four 1 s and five 0 s which is not the LOUDS encoding of any tree.

Exercise 2 [2+3+3]

1. Give the γ code of the sequence 1, 2, 4, 8, 16, 32, 64
2. Provide an asymptotic estimate (in terms of n) of the cost of γ -encoding the sequence 1, 2, 4, 8, $\dots$ 2^n . If you cannot prove an exact asymptotic estimate, give upper and lower asymptotic estimates with O and Ω .
3. Provide an asymptotic estimate for the encoding of the same sequence using Elias-Fano encoding.

Exercise 3 [4+3+3] Given the BWT string ab\$bbrraaa

1. show how to compute the number of occurrences of the pattern arb using the backward search algorithm,
2. describe the computation for the inversion of the BWT and retrieve the original text
3. compute the SA and LCP array for the original text

Exercise 4 [3+3]

1. Compare advantages and disadvantages of Bloom filters with respect to ordinary hash tables
2. Discuss how the number r of hash functions used in a Bloom filter influences the false positive ratio.

Exercise 5 [4] Derive the complexity of the Skew algorithm (aka DC3) for the direct computation of the suffix array